Computational Color Constancy: Survey and Experiments

Arjan Gijsenij, *Member, IEEE*, Theo Gevers, *Member, IEEE*, and Joost van de Weijer, *Member, IEEE*

***Abstract*—Computational color constancy is a fundamental prerequisite for many computer vision applications. This paper presents a survey of many recent developments and state-of-the-art methods. Several criteria are proposed that are used to assess the approaches. A taxonomy of existing algorithms is proposed and methods are separated in three groups: static methods, gamut-based methods, and learning-based methods. Further, the experimental setup is discussed including an overview of publicly available datasets. Finally, various freely available methods, of which some are considered to be state of the art, are evaluated on two datasets.**

***Index Terms*—Color constancy, illuminant estimation, performance evaluation, survey.**

I. Introduction

Color can be an important cue for computer vision or image processing related topics, like human–computer interaction [1], color feature extraction [2], and color appearance models [3]. The colors that are present in images are determined by the intrinsic properties of objects and surfaces *as well as* the color of the light source. For a robust color-based system, these effects of the light source should be filtered out. This ability to account for the color of the light source is called *color constancy*.

Human vision has the natural tendency to correct for the effects of the color of the light source, e.g., [4]–[8], but the mechanism that is involved with this ability is not yet fully understood. Early work resulted in the Retinex theory by Land and McCann [9]–[11], after which many computational models are derived that are based on this perceptual theory [12]–[14]. However, there still exists a discrepancy between human and computational color constancy. Computational models cannot fully explain the observed color constancy of human observers, as shown by Kraft and Brainard [15]. They tested the ability of several computational theories to account for human color constancy, but found that each theory leaves considerable residual constancy. In other words, without the specific cues corresponding to the computational models, humans are still to some extent color constant [15]. Alternatively, observations on human color constancy cannot be readily applied to computational models: Golz and MacLeod [16], [17] showed that chromatic scene statistics influence the accuracy of human color constancy, but when mapped to computational models, the influence was found to be very weak at best [18]. Therefore, the focus in this paper is on computational color constancy algorithms. As an example, consider the images in Fig. 1. These images depict the same scene, rendered under four different light sources. The goal of computational color constancy algorithms is to correct the (first three) target images (under three different colored light sources), so that they appear identical to the (fourth) canonical image (under a white light source).

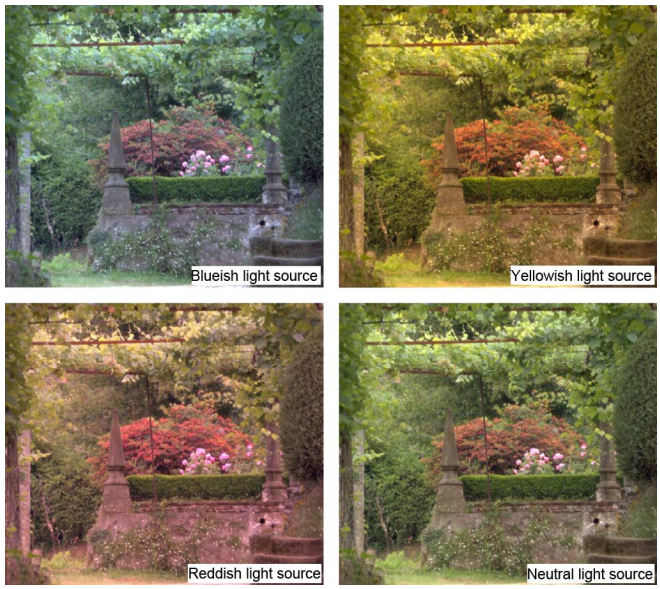

Fig. 1. Illustration of the influence of differently colored light sources on the measured image values. These images are adapted from [7] and show the same scene, rendered under four different light sources.

Often, computational models for color constancy are characterized by the estimation of the illuminant. The corresponding algorithms are based on the assumption that the color of the light source is spatially uniform across the scene. Hence, after globally estimating the color of the light source, color correction can be applied to the image to obtain a color constant image. Another line of research, not pursued in this paper, focuses on the invariance that can be obtained by applying various photometric

Manuscript received June 25, 2010; revised November 23, 2010; accepted February 13, 2011. Date of publication February 22, 2011; date of current version August 19, 2011. This work was supported by the EU projects ERGTS-VICI-224737, by the Spanish Research Program Consolider-Ingenio 2010: MIPRCV (CSD200700018), and by the Spanish projects TIN2009-14173 and the Ramon y Cajal fellowship. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Maya Gupta.

A. Gijsenij is with the University of Amsterdam, Amsterdam 1098 XG, The Netherlands (e-mail: a.gijsenij@uva.nl).

T. Gevers is with the University of Amsterdam, Amsterdam 1098 XG, The Netherlands, and also with the Computer Vision Center, Universitat Autònoma de Barcelona, Bellaterra, Barcelona 08193, Spain (e-mail: th.gevers@uva.nl).

J. van de Weijer is with the Computer Vision Center, Universitat Autònoma de Barcelona, Bellaterra, Barcelona 08193, Spain (e-mail: joost@cvc.uab.es).

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

Digital Object Identifier 10.1109/TIP.2011.2118224

transformations, sometimes also referred to as color constancy [19]–[21]. Such methods are often extended to incorporate other forms of invariance, like invariance to highlights or shadows, but do not result in output images that have any visual similarity to the original input image.

The main focus of this paper is on the estimation of the illuminant color, more specifically the estimation of the illuminant using a single image from a regular digital camera. Hence, methods using additional images, e.g., [22]–[27], physically different devices, e.g., [28], [29], or video sequences, e.g., [30], [31], are not included in this paper. When using a single image that is taken with a regular digital camera, illuminant estimation is an underconstrained problem; both the intrinsic properties of a surface and the color of the illuminant have to be estimated, while only the product of the two (i.e., the actual image) is known. Early solutions for estimating the illuminant tried to bridge this gap by adopting linear models of lights and surfaces [32]–[35]. Unfortunately, these approaches do not result in satisfactory results for real-world images.

In this paper, state-of-the-art approaches are divided into three types of algorithms: 1) static methods, 2) gamut-based methods; and 3) learning-based methods methods.[1] The first type of algorithms are methods that are applied to any image without the need for training. In other words, for a given dataset or application, the parameter setting is kept fixed (or *static*). For the second and third type of algorithms, a model needs to be trained before the illuminant can be estimated. This is an important distinction that partially determines the suitability of an algorithm for applicability to real-world systems. The criteria used in this paper to assess the computational methods are the following:

1) the requirement of training data;
2) the accuracy of the estimation;
3) the computational runtime of the method;
4) transparency of the approach;
5) complexity of the implementation;
6) number of tunable parameters.

For evaluation of computational color constancy methods, various datasets are currently available, ranging from high-quality hyperspectral scenes to large-scale real-world RGB images. Two large datasets are selected to analyze the performance of various publicly available methods. The results of these experiments are available online at http://www.color-constancy.com.

This paper is organized as follows. First, in Section II, a formal definition of the problem is discussed, together with the methodology explored in this paper. Then, in Sections III–V, current approaches as well as recent developments are described. Section VI describes the experimental setup, including commonly used error measures and databases with ground truth. In Section VII, two of these datasets are selected and extensive experiments are performed using a wide range of methods. Finally, in Section VIII, a discussion and future directions are presented.

[1]Note that the classification of the methods is not absolute, which means that some methods are, for example, both gamut-based and learning-based.

II. Color Constancy

Color constancy can be achieved by estimating the color of the light source, followed by a transformation of the original image values using this illuminant estimate. The aim of this transformation is not to scale the brightness level of the image, as color constancy methods only correct for the *chromaticity* of the light source. Section II-A will discuss the formation of an image, while more information on the transformation is discussed in Section II-B.

A. Image Formation

The image values $\mathbf{f} = (f_R,\ f_G,\ f_B)^T$ depend on the color of the light source $I(\lambda)$, the surface reflectance $S(\mathbf{x}, \lambda)$, and the camera sensitivity function $\rho(\lambda) = (\rho_R(\lambda), \rho_G(\lambda), \rho_B(\lambda))^T$, where λ is the wavelength of the light and $\mathbf{x}$ is the spatial coordinate [36], [37]

$$f_c(\mathbf{x}) = m_b(\mathbf{x}) \int_\omega I(\lambda)\rho_c(\lambda)S(\mathbf{x}, \lambda)d\lambda + m_s(\mathbf{x}) \int_\omega I(\lambda)\rho_c(\lambda)d\lambda \quad (1)$$

where $c = \{R, G, B\}$, ω is the visible spectrum, and m_b and m_s are scale factors that model the relative amount of body and specular reflectance that contribute to the overall light reflected at location $\mathbf{x}$. Under the Lambertian assumption, the specular reflection is ignored. This results in the following model:

$$f_c(\mathbf{x}) = m(\mathbf{x}) \int_\omega I(\lambda)\rho_c(\lambda)S(\mathbf{x}, \lambda)d\lambda \quad (2)$$

where $m(\mathbf{x})$ is the Lambertian shading. It is assumed that the scene is illuminated by one single light source and that the observed color of the light source $\mathbf{e}$ depends on the color of the light source $I(\lambda)$ as well as the camera sensitivity function $\rho(\lambda)$

$$\mathbf{e} = \begin{pmatrix} e_R \\ e_G \\ e_B \end{pmatrix} = \int_\omega I(\lambda)\rho(\lambda)d\lambda. \quad (3)$$

Without prior knowledge, both $I(\lambda)$ and $\rho(\lambda)$ are unknown, and hence, the estimation of $\mathbf{e}$ is an under-constrained problem that cannot be solved without further assumptions. Therefore, in practice, color constancy algorithms are based on various simplifying assumptions such as restricted gamuts (limited number of image colors which can be observed under a specific illuminant), the distribution of colors that are present in an image (e.g., white patch, gray world, etc.), and the set of possible light sources. This paper will give an overview on the assumptions and methods that are used for the estimation of the color of the light source.

B. Image Correction

The focus of this paper is on estimating the color of the light source. However, in many cases, the color of the light source is of less importance than the appearance of the input image under a reference light (called canonical light source). Therefore, the aim of most of the color constancy methods is to transform all colors of the input image, taken under an unknown light source,

to colors as they appear under this canonical light source. This transformation can be considered to be an instantiation of chromatic adaptation, e.g., [3]. Chromatic adaptation is often modeled using a linear transformation, which in turn can be simplified by a diagonal transformation when certain conditions are met [38]–[40]. Other possible chromatic adaptation methods include linearized Bradford [41] and CIECAT02 [42].

In this paper, the diagonal transform or *von Kries Model* [43] is used, without changing the color basis [44], [45] or applying spectral sharpening [46], [47]. These techniques have shown their merits of improving the quality of output images if the illuminant under which the original image was recorded is known. Since the discussed methods focus on estimation of this illuminant, changing color bases and spectral sharpening techniques are omitted in this paper for simplicity. The diagonal model that is used is given by

$$\mathbf{f}_t = \mathcal{D}_{u,t}\mathbf{f}_u \tag{4}$$

where $\mathbf{f}_u$ is the image taken under an unknown light source, $\mathbf{f}_t$ is the same image transformed, so it appears if it was taken under the canonical illuminant, and $\mathcal{D}_{u,t}$ is a diagonal matrix which maps colors that are taken under an unknown light source u to their corresponding colors under the canonical illuminant c

$$\begin{pmatrix} R_c \\ G_c \\ B_c \end{pmatrix} = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} \begin{pmatrix} R_u \\ G_u \\ B_u \end{pmatrix}. \tag{5}$$

Although this model is merely an approximation of illuminant change and might not accurately be able to model photometric changes, it is widely accepted as color correction model [38]–[40], [48], [49] and it underpins many color constancy algorithms (e.g., the gamut mapping in Section IV) and gray-world-based methods in Section III-A. The diagonal mapping is used throughout this paper to create output images after correction by a color constancy algorithm, where a perfect white light, i.e., $(1/\sqrt{3},\ 1/\sqrt{3},\ 1/\sqrt{3})^T$, is used as canonical illuminant.

III. Static Methods

The first type of illuminant estimation algorithms that is discussed in this paper are static methods, or methods that are applied to input images with a fixed parameter setting. Two subtypes are distinguished: 1) methods that are based on low-level statistics; and 2) methods that are based on the physics-based dichromatic reflection model.

A. Low-Level Statistics-Based Methods

The best-known and most often used assumption of this type is the gray-world assumption [50]: *the average reflectance in a scene under a neutral light source is achromatic*. It directly follows from this assumption that any deviation from achromaticity in the average scene color is caused by the effects of the illuminant. This implies that the color of the light source $\mathbf{e}$ can be estimated by computing the average color in the image

$$\int f_c(\mathbf{x})d\mathbf{x} = ke_c \tag{6}$$

where k is a multiplicative constant chosen such that the illuminant color, $\mathbf{e} = (e_R,\ e_G,\ e_B)^T$, has unit length. Alternatively, instead of computing the average color of all pixels, it has been shown that segmenting the image and computing the average color of all segments may improve the performance of the gray-world algorithm [51], [52]. This preprocessing step can lead to improved results because the gray world is sensitive to large uniformly colored surfaces, as this often leads to scenes where the underlying assumption fails. Segmenting the image before computing the scene average color will reduce the effects of these large uniformly colored patches. Related methods attempt to identify the intrinsic gray surfaces in an image, i.e., they attempt to find the surfaces under a colored light source that would appear gray if rendered under a white light source [53]–[55]. When *accurately* recovered, these surfaces contain a strong cue for the estimation of the light source.

Another well-known assumption is the white-patch assumption [10]: *the maximum response in the RGB-channels is caused by a perfect reflectance*. A surface with perfect reflectance properties will reflect the full range of light that it captures. Consequently, the color of this perfect reflectance is exactly the color of the light source. In practice, the assumption of perfect reflectance is alleviated by considering the color channels separately, resulting in the *max-RGB* algorithm. This method estimates the illuminant by computing the maximum response in the separate color channels

$$\max_{\mathbf{x}} f_c(\mathbf{x}) = ke_c. \tag{7}$$

Related algorithms apply some sort of smoothing to the image, prior to the illuminant estimation [56], [57]. This preprocessing step has similar effects on the performance of the white-patch algorithm as segmentation on the gray world. In this case, the effect of noisy pixels (with an accidental high intensity) is reduced, improving the accuracy of the white-patch method. An additional advantage of the local space average color method [57] is that it can provide a pixel-wise illuminant estimate. Consequently, it does not require the image to be captured under a spatially uniform light source. An analysis of the *max-RGB* algorithm is presented in [58] and [59], where it is shown that the dynamic range of an image, in addition to the preprocessing strategy, can have a significant influence on the performance of this method.

In [60], the white patch and the gray-world algorithms are shown to be special instantiations of the more general Minkowski framework

$$\mathcal{L}_c(p) = \left(\int f_c^p(\mathbf{x})d\mathbf{x} \right)^{1/p} = ke_c. \tag{8}$$

Substituting $p = 1$ in (8) is equivalent to computing the average of $\mathbf{f}(\mathbf{x})$, i.e., $\mathcal{L}(1) = (\mathcal{L}_R(1), \mathcal{L}_G(1), \mathcal{L}_B(1))^T$ equals the gray-world algorithm. When $p = \infty$, (8) results in computing the maximum of $\mathbf{f}(\mathbf{x})$, i.e., $\mathcal{L}(\infty)$ equals the white-patch algorithm. In general, to arrive at a proper value, p is tuned for the dataset at hand. Hence, the optimal value of this parameter may vary for different datasets.

The assumptions of the previous color constancy methods are based on the distribution of colors (i.e., pixel values) that are

present in an image. The incorporation of higher order image statistics (in the form of image derivatives) is proposed in [61], where a framework called gray edge is presented that incorporates the well-known methods like (8), as well as methods based on first- and second-order derivatives

$$\left(\int \left|\frac{\partial^n f_{c,\sigma}(\mathbf{x})}{\partial \mathbf{x}^n}\right|^p d\mathbf{x}\right)^{1/p} = k e_c^{n,p,\sigma} \tag{9}$$

where $|\cdot|$ indicates the Frobenius norm, $c = \{R, G, B\}$, n is the order of the derivative and p is the Minkowski-norm. Further, derivatives are defined as convolving the images by Gaussian derivative filters with scale parameter σ [62]

$$\frac{\partial^{s+t} f_{c,\sigma}}{\partial x^s y^t} = f_c * \frac{\partial^{s+t} G_\sigma}{\partial x^s \partial y^t} \tag{10}$$

where $*$ denotes the convolution and $s + t = n$. This method is enhanced with an illuminant constraint by Chen *et al.* [63]. Further, Chakrabarti *et al.* [64] explicitly model the spatial dependencies between pixels. The advantage of this approach compared to the gray edge is that it is able to learn the dependencies between pixels in an efficient way, but the training phase does rely on an extensive database of images. Finally, Gijsenij *et al.* [65] note that different types of edges might contain various amounts of information. They extend the gray-edge method to incorporate a general weighting scheme (assigning higher weights to certain edges), resulting in the weighted gray edge. Physics-based weighting schemes are proposed, concluding that specular edges are favored for the estimation of the illuminant. The introduction of these weighting schemes result in more accurate illuminant estimates, but at the cost of complexity (both in computation and implementation).

B. Physics-Based Methods

Most methods are based on the more simple Lambertian model following (2), but some methods adopt the dichromatic reflection model of image formation, following (1). These methods use information about the physical interaction between the light source and the objects in a scene and are called physics-based methods. These approaches exploit the dichromatic model to constrain the illuminants. The underlying assumption is that all pixels of one surface fall on a plane in RGB color space. If multiple of such planes are found, corresponding to various *different* surfaces, then the color of the light source is estimated using the intersection of those planes. Various approaches have been proposed that use specularities or highlights [66]–[69]. The intuition behind such methods is that if pixels are found where the body reflectance factor m_b in (1) is (close to) zero, then the color of these pixels are similar or identical to the color of the light source. However, all these methods suffer from some disadvantages: retrieving the specular reflections is challenging and color clipping can occur. The latter effectively eliminates the usability of specular pixels (which are more likely to be clipped than other pixels).

A different physics-based method is proposed by Finlayson and Schaefer [70]. This method uses the dichromatic reflection model to project the pixels of a single surface into chromaticity space. Then, the set of possible light sources is modeled by using the Planckian locus of black-body radiators. This planckian locus is intersected with the dichromatic line of the surface to recover the color of the light source. This method, in theory, allows for the estimation of the illuminant even when there is only one surface present in the scene. However, it does require all pixels in the image to be segmented, so that all unique surfaces are identified. Alternatively, the colors in an image can be described using a multilinear model consisting of several planes simultaneously oriented around an axis defined by the illuminant [71], [72]. This eliminates the problem of presegmentation, but does rely on the observation that a representative color of any given material can be identified. In [73], these requirements are relaxed, resulting in a two Hough transform voting procedure.

IV. Gamut-Based Methods

The gamut mapping algorithm has been introduced by Forsyth [74]. It is based on the assumption, that *in real-world images, for a given illuminant, one observes only a limited number of colors*. Consequently, any variations in the colors of an image (i.e., colors that are different from the colors that can be observed under a given illuminant) are caused by a deviation in the color of the light source. This limited set of colors that can occur under a given illuminant is called the *canonical gamut* $\mathcal{C}$ and it is found in a training phase by observing as many surfaces under one known light source (called the *canonical illuminant*) as possible.

The flow of the gamut mapping is illustrated in Fig. 2. In general, a gamut mapping algorithm takes as input an image taken under an unknown light source (i.e., an image of which the illuminant is to be estimated), along with the precomputed canonical gamut (see blocks 1 and 2 in Fig. 2). Next, the algorithm consists of three important steps.

1) Estimate the gamut of the unknown light source by assuming that the colors in the input image are representative for the gamut of the unknown light source. So, all colors of the input image are collected in the input gamut $\mathcal{I}$. The gamut of the input image is used as feature in Fig. 2.
2) Determine the set of *feasible mappings* $\mathcal{M}$, i.e., all mappings that can be applied to the gamut of the input image and that result in a gamut that lies completely within the canonical gamut. Under the assumption of the diagonal mapping, a unique mapping exists that converts the gamut of the unknown light source to the canonical gamut. However, since the gamut of the unknown light source is simply estimated by using the gamut of *one* input image, in practice, several mappings are obtained. Every mapping i in the set $\mathcal{M}$ should take the input gamut completely inside the canonical gamut

$$\mathcal{M}_i \mathcal{I} \in \mathcal{C}. \tag{11}$$

   This corresponds to block 4 in Fig. 2, where the learned model (e.g., the canonical gamut) together with the input features (e.g., the input gamut) is used to derive an estimate of the color of the light source.
3) Apply an estimator to select one mapping from the set of feasible mappings (see block 5 in Fig. 2). The selected

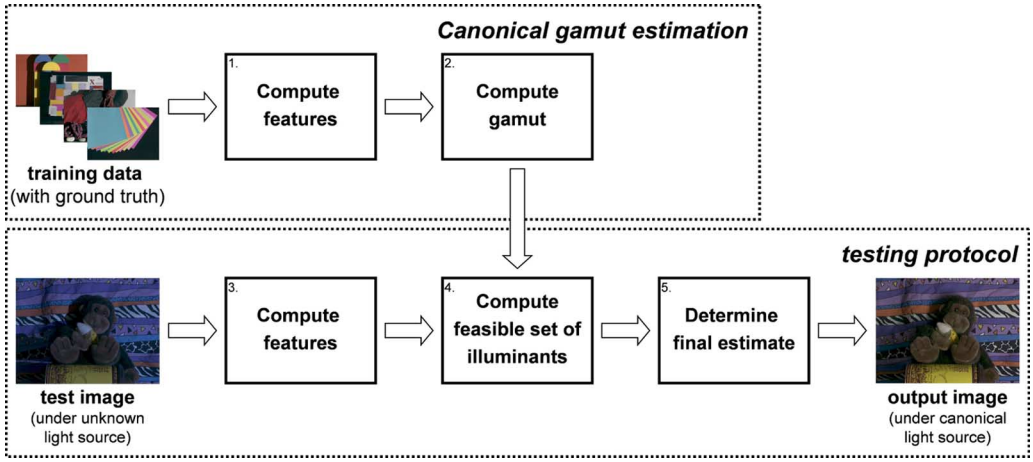

Fig. 2. Overview of gamut-based algorithms. The training phase consists of learning a model given the features of a wide variety of input images (see block 1), resulting in the canonical gamut (see block 2). The testing protocol consists of applying the learned model to the computed features of the input image (see blocks 3 and 4). Finally, one illuminant estimate is selected from the feasible set of illuminants (see block 5), and this estimate is used to correct the input image.

mapping can be applied to the canonical illuminant to obtain an estimate of the unknown illuminant. The original method [74] used the heuristic that the mapping resulting in the most colorful scene, i.e., the diagonal matrix with the largest trace, is the most suitable mapping. Simple alternatives are the average of the feasible set or a weighted average [75].

Several extensions have been proposed. First of all, difficulties in implementation are addressed in [76] and [77], where it is shown that the gamut mapping algorithm can also be computed in chromaticity space $(R/B, G/B)$. These modifications correspond to different feature computation in blocks 1 and 3 in Fig. 2. However, the performance of this 2-D approach is slightly worse than the performance of the 3-D approach. It is shown that this decrease in performance is caused by the perspective distortion of the possible set of illuminants (the set of feasible mappings in step 2) that is caused by the conversion of the original image to 2-D-chromaticity values. This can be solved by mapping the 2-D-feasible set back to three dimensions before selecting the most appropriate mapping [77], [78] (i.e., a slightly modified block 4 in Fig. 2). Alternatively, in [79], an efficient implementation is introduced using convex programming. This implementation is shown to result in similar performance as the original method. Finally, in [80], a simpler version of the gamut mapping is proposed using a simple cube rather than the convex hull of the pixel values.

Another extension of the gamut mapping algorithm deals with dependence on the diagonal model. One of the disadvantages of the original method is that a null solution can occur if the diagonal model fails. In other words, if the diagonal model does not fit the input data accurately, then it is possible that no feasible mapping can be found that maps the input data into the canonical gamut with one single transform. This results in an empty solution set. One heuristic approach to avoid such situations it to incrementally augment the input gamut until a nonempty feasible set is found [52], [81]. Another heuristic approach is to extend the size of the canonical gamut. Finlayson [76] increases the canonical gamut by 5%, while Barnard [75] systematically enlarges the canonical gamut by learning this gamut not only with surfaces that are illuminated by the canonical light source, but also with surfaces that are captured under different light sources which are mapped to the canonical illuminant using the diagonal model. Hence, possible failure of the diagonal model is captured by augmenting the canonical gamut. Another strategy is to simulate specularities during computation of the canonical gamut, potentially increasing the performance of the gamut mapping method even in situations where there is no null solution [82], [83]. Alternatively, to avoid this null solution, an extension of the diagonal model called diagonal-offset model is proposed [84]. This model allows for translation of the input colors in addition to the regular linear transformation, effectively introducing some slack into the model. All these modifications are implemented in block 5 of Fig. 2.

All these variations of the gamut mapping algorithm are restricted to the use of pixel values to estimate the illuminant. Gijsenij *et al.* [85] extended the gamut mapping to incorporate the differential nature of images. They analytically show that the gamut mapping framework is able to incorporate any linear filter output and that, if failure of the diagonal model can be prevented by adapting the diagonal-offset model [84], derivative-based gamut mapping will not result in null solutions. Further, they propose several combinations of different n-jet-based gamut mappings and show that the best performance is obtained by taking the intersection of feasible sets.

The fusion strategy proposed in [85] is based on the smaller set of possible light sources obtained when taking the intersection of multiple feasible sets. Another method to constrain the feasible set is proposed by Finlayson *et al.* [86] and is called gamut-constrained illuminant estimation. This method effectively reduces the problem of illuminant estimation to *illuminant classification* by considering only a limited number of possible light sources, similar to color by correlation. One canonical gamut is learned for every possible light source. Then, the unknown illuminant of the input image is estimated by matching the input gamut to each of the canonical gamuts, selecting the best match as final estimate.

V. Learning-Based Methods

The third type of algorithms estimate the illuminant using a model that is learned on training data. Indeed, gamut-based methods in Section IV can be considered learning based too, but since this approach has been quite influential in color constancy research, it has been discussed separately.

Initial approaches using machine learning techniques are based on neural networks [87]. The input to the neural network consists of a large binarized chromaticity histogram of the input image, the output is two chromaticity values of the estimated illuminant. Although this approach, when trained correctly, can deliver accurate color constancy even when only a few distinct surfaces are present, the training phase requires a large amount of training data. Similar approaches apply support vector regression [88]–[90] or linear regression techniques like ridge regression and kernel regression [91]–[93] to the same type of input data. Alternatively, thin-plate spline interpolation is proposed in [94] to interpolate the color of the light source over a nonuniformly sampled input space (i.e., training images).

A. Methods Using Low-Level Statistics

Color by correlation [95] is generally considered to be a discrete implementation of the gamut mapping, but it is actually a more general framework which includes other low-level statistics-based methods like gray world and white patch as well. The canonical gamut is replaced with a correlation matrix. The correlation matrix for a known light source $\mathbf{e}_i$ is computed by first partitioning the chromaticity space into a finite number of cells, followed by computation of the probabilities of occurrence of the coordinates under illuminant $\mathbf{e}_i$. One correlation matrix is computed for every possible illuminant that is considered. Then, the information that is obtained from the input image matched to the information in the correlation matrices to obtain a probability for every considered light source. The probability of illuminant $\mathbf{e}_i$ indicates the likelihood that the current input image was captured under this light source. Finally, using these probabilities, one light source is selected as scene illuminant, e.g., using maximum likelihood [95] or Kullback–Leibler divergence [96].

Other methods using low-level statistics are based on the Bayesian formulation. Several approaches are proposed that model the variability of reflectance and light source as random variables. The illuminant is then estimated from the posterior distribution conditioned on the image intensity data [97]–[99]. However, the assumptions of independent reflectance that is Gaussian distributed proved to be too strong (unless learned for and applied to a specific application like outdoor object recognition [100]). Rosenberg *et al.* [101] replace these assumptions with nonparametric models, using the assumption that nearby pixels are correlated. Further, Gehler *et al.* [102] show that competitive results to state of the art can be obtained when precise priors for illumination and reflectance are used.

B. Methods Using Medium- and High-Level Statistics

Despite the large variety of available methods, none of the color constancy methods can be considered as universal. All algorithms are based on error-prone assumptions or simplifications and none of the methods can guarantee satisfactory results for all images. To still be able to obtain good results on a full set of images rather than on a subset of images, multiple algorithms can be combined to estimate the illuminant. The outline of such approaches is illustrated using Fig. 3. The first attempts of combining color constancy algorithms are based on combining the output of multiple methods [103]–[105]. In [103], three color constancy methods are combined using both linear (a weighted average of the illuminant estimates) and nonlinear (a neural network based on the estimates of the considered methods) fusion methods are considered. It is shown that a weighted average, optimizing the weights in a least-mean-square sense, results in the best performance, outperforming the individual methods that are considered. In [104], a statistics-based method is combined with a physics-based method. Both methods return likelihoods for a predefined set of light sources and by combining these likelihoods *a posteriori*, more accurate results are obtained. Finally, in [105], several different combination strategies are employed. These strategies include the mean value of all estimates, the mean value of the two closest estimates and the mean value of all methods excluding the N most remote estimates (i.e., excluding the estimates with the largest distance to the other estimates). This latter strategy, excluding two out of six estimates, resulted in the best performance. All these approaches use fixed fusion weights in blocks 3 and 6 in Fig. 3 and the features in blocks 1 and 4 can be seen as the illuminant estimates themselves.

Instead of combining the output of multiple algorithms into one more accurate estimate, a different strategy is proposed by Gijsenij and Gevers [106], [107]. They use the intrinsic properties of natural images to select the most appropriate color constancy method for every input image. Characteristics of natural images are captured using the Weibull parameterization (e.g., grain size and contrast), and they show that the corresponding parameters (β and γ) are related to image attributes to which color constancy methods using low-level features (e.g., gray world, white patch, and gray edge) are sensitive to. In other words, they select the most appropriate color constancy algorithm for every image, depending on the contents of the image. For instance, if an image contains only a few edges (corresponding to a low signal-to-noise ratio), then pixel-based methods like gray world and white patch are preferred. On the other hand, edge-based methods (e.g., first- and second-order gray edge) are preferred when the signal-to-noise ratio is medium or high. Instead of using Weibull parameterization, various other features are explored in [108]–[111] to predict the most appropriate algorithm for a given image. The most notable differences between these approaches is in block 1 of Fig. 3.

C. Semantic Information

Recently, several methods have been proposed that estimate the illuminant using some sort of semantic information. Gijsenij and Gevers [106], [107], [112] propose to dynamically determine which color constancy algorithm should be used for a specific image, depending on the scene category. They do not discuss the actual classification of the images and how to use the uncertainty in the classification results, but merely assume that the scene category of an image is known. Bianco *et al.* [113] propose an indoor–outdoor classifier and use the uncertainty of the classifier to introduce an "unsure" class. Then, they learn the

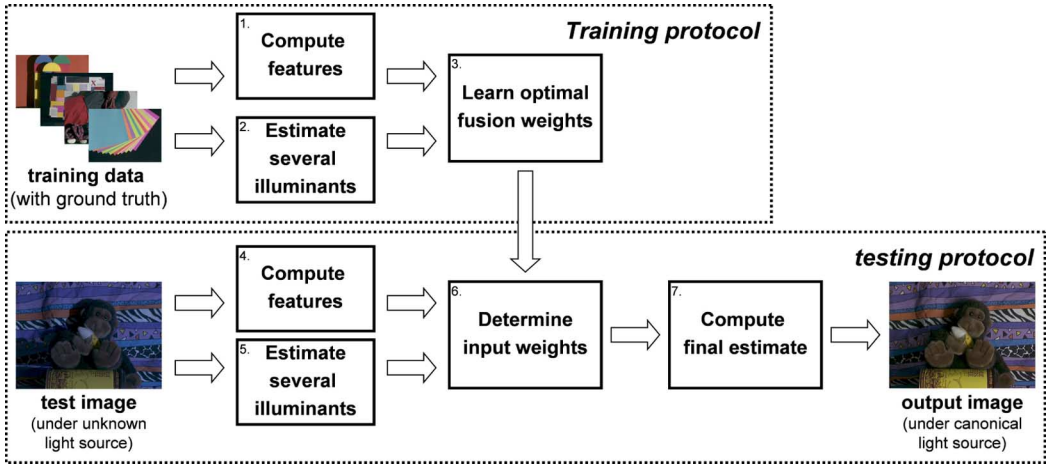

Fig. 3. Overview of several learning-based algorithms. After the input features are compared to the training features, the optimal fusion weights are determined. These weights can be static (fixed for all input images) or dynamic (dependent on the features of the input images). After the fusion weights for the current input image are determined, the estimated illuminants are combined into one final estimate. This combination can either by hard (one of the illuminants is selected as final estimate) or soft (a weighted average of estimates is computed).

appropriate color constancy algorithm for each of these three classes. However, the distinction between indoor and outdoor classes is rather arbitrary. Therefore, Lu *et al.* [114], [115] propose to use a stage classifier that distinguishes medium-level semantic classes [116]. This results in a color constancy method that explicitly uses 3-D scene information for the estimation of the color of the light source.

A different approach uses high-level visual information. Rather than classifying images into a specific class and applying different color constancy methods depending on the semantic category, van de Weijer *et al.* [117] propose to cast illuminant hypotheses that are generated and evaluated based on the likelihood of semantic content. Using prior knowledge about the world, an illuminant estimate is selected that results in colors that are consistent with the learned model of the world. In other words, an illuminant estimate is selected that will generate plausible images, e.g., images with a blue rather than purple sky and green rather than reddish grass. A similar approach is proposed in [118], where the term *memory color* is used to refer to color that are specifically associated with object categories. These object-specific colors are used to refine the estimated illuminants.

VI. Experimental Setup

Evaluation of illuminant estimation algorithms requires images with a scene illuminant that is known (ground truth). The general experimental setup is as follows. First, a part of the data is used for training, if the algorithm requires this. Then, the color of the light source is estimated for every remaining image of the database and compared to the ground truth. The comparison requires some similarity or distance measure; an often used measure is the angular error

$$d_{\text{angle}}(\mathbf{e}_e, \mathbf{e}_u) = \cos^{-1}\left(\frac{\mathbf{e}_e \cdot \mathbf{e}_u}{\|\mathbf{e}_e\| \cdot \|\mathbf{e}_u\|}\right) \tag{12}$$

where $\mathbf{e}_e \cdot \mathbf{e}_u$ is the dot product of the estimated illuminant $\mathbf{e}_e$ and the ground truth $\mathbf{e}_u$ and $\|\cdot\|$ is the Euclidean norm of a vector. Alternate setups exist, depending on the application. For instance, Funt *et al.* [119] describe an experiment to evaluate the usefulness of color constancy algorithms as preprocessing step in object recognition.

In most situations, for instance when the application is to obtain an accurate reproduction of the image under a white light source, the distance measure should be an accurate reflection of the quality of the output image. In [120], several distance measures are analyzed with respect to this requirement, and it is shown that the often used angular error correlates *reasonably well* with the perceived quality of the output images. However, to optimize this correlation, a dataset specific measure, called perceptual Euclidean distance, should be adopted.

Multiple algorithms are typically compared using a large number of images, so the performance of every algorithm needs to be summarized over all images. An intuitive measure would be to simply compute the average error over the full database. However, the error measures are often not normally distributed but rather skewed resulting in a nonsymmetric distribution. Hence, the mean value of the errors is a poor summary statistic [120], [121]. More appropriate measures to summarize the distribution are the median [121] or the trimean [120]. The median gives an indication of the performance of the method on the majority of the images, while the trimean also gives an indication of the extreme values of the distribution.

In addition to these summarizing statistics, more insight into the performance of the algorithms can be obtained by showing box plots or by performing significance tests [120], [121]. A box plot is used to visualize the underlying distribution of the error metric of one color constancy method. A significance test, like the Wilcoxon sign test, is usually performed between two methods to show that the difference between two algorithms is statistically significant [121]. Further, the obtained improvement can only be considered to be perceptually significant if the relative difference between two methods is at least 5–6%. Below that, the difference is not noticeable to human observers [120].

A. Datasets

Two types of data can be distinguished that are used to evaluated color constancy methods: hyperspectral data and RGB images. Databases containing hyperspectral datasets are often smaller (less images) and contain less variation than datasets with RGB images. The main advantage of hyperspectral data is that many different illuminants can be used to realistically render the same scene under various light sources, and consequently, a systematic evaluation of the methods is possible. However, the simulation of illuminants generally does not include real-world effects like interreflections and nonuniformity. Consequently, the evaluation on RGB images results in more realistic performance evaluations. Ideally, both types of data should be used for a thorough evaluation of color constancy methods [52], [81].

An often used hyperspectral database was composed by Barnard *et al.* [122]. This set consists of 1995 surface reflectance spectra and 287 illuminant spectra. These reflectance and illuminant spectra can be used to generate an extensive range of surfaces (i.e., RGB values), allowing for a systematic evaluation of color constancy performance. Another database that is specifically useful for the evaluation of color constancy algorithm is created by Foster *et al.* [123], [124]. These two sets each contain eight natural scenes that can be converted into an arbitrary number of images using various illuminant spectra (not provided). Finally, a database by Párraga *et al.* [125] contains 29 hyperspectral images with low resolution (256×256 pixels).

Databases with RGB images are more informative on the performance of the algorithms under realistic circumstances. The first step toward realistic evaluation of color constancy methods involves isolated compositions of objects that are illuminated by 11 different light sources [122]. The 11 different lights include three different fluorescent lights, four different incandescent lights, and four incandescent lights combined with a blue filter and are selected to span the range of natural and man-made illuminants as best as possible. The complete database contains 22 scenes with minimal specularities, 9 scenes with dielectric specularities, 14 scenes with metallic specularities, and 6 scenes with at least one fluorescent surface. Often, for illuminant estimation evaluation, a subset of 31 scenes is used that only consists of the scenes with minimal and with dielectric specularities. Even though these images encompass several different illuminants and scenes, the variation of the images is limited.

A more varied database is composed by Ciurea and Funt [126]. This dataset contains over 11 000 images, extracted from 2 h of video recorded under a large variety of imaging conditions (including indoor, outdoor, desert, cityscape, and other settings). In total, the images are divided into 15 different clips taken at different locations. The ground truth is acquired by attaching a gray sphere to the camera, that is displayed in the bottom right corner of the image. Obviously, this gray sphere should be masked during experiments to avoid biasing the algorithms. Some examples of images that are in this dataset are shown in Fig. 4(a). The main disadvantage of this set is the correlation that exists between some of the images. Since the images are extracted from video sequences, some images are rather similar in content. This should especially be taken into account when dividing the images into training and test sets. Another issue of this dataset is that an unknown postprocessing procedure is applied to the images by the camera, including gamma correction and compression. A similar dataset has been recently proposed in [127]. Although the number of images in this set (83 outdoor images) is not comparable to the previous set, the images are not correlated and are available in XYZ format and can be considered to be of better quality. Further, an extension of the dataset is proposed in [128], where an additional 126 images with varying environments (e.g., forest, seaside, mountain snow, and motorways) are introduced. Gehler *et al.* [102] introduced a database, consisting of 568 images, both indoor and outdoor. The ground truth of these images is obtained using a MacBeth color checker that is placed in the scene. The main advantage of this database is the quality of the images (which are free of correction), but the variation of the images is not as large as the dataset containing over 11 000 images. Some examples of images that are in this dataset are shown in Fig. 4(b). Finally, Shi and Funt generated a set of 105 high-dynamic-range images [58], [59]. These images use four color checkers to capture the ground truth and are constructed from multiple exposures of the same scene.

(a)

(b)

Fig. 4. Some examples of the two datasets that are used for the experiments. (a) Example images of SFU dataset. (b) Example images of color-checker-set.

A summary of available datasets is presented in Table I. Generally, a distinction can be made between real-world RGB images and images with controlled illumination conditions. The latter type of data, including hyperspectral images, should mainly be used to aid the development of new algorithms and for the systematic analysis of methods. Conclusions about the performance with respect to existing methods based on such datasets should be avoided as much as possible, since it is relatively easy to tune any algorithm to obtain a high performance on such datasets. The real-world RGB images are more suited to compare algorithms, as such data are probably the target data of the intended application of most color constancy algorithms.

TABLE I
SUMMARY OF DATASETS WITH ADVANTAGES AND DISADVANTAGES

Data set	Pros	Cons
SFU hyperspectral set [122] (1,995 surface spectra)	+ Large variety + Allows for systematic evaluation	− Best-case assessment of performance
Foster et al. [123, 124] (8 + 8 images)	+ High quality hyperspectral images + Real-world natural scenes	− Limited amount of data
Bristol set [125] (28 images)	+ Hyperspectral images + Real-world natural scenes	− Low quality images
SFU set [122] (223 + 98 + 149 + 59 images)	+ Scenes with varying characteristics + Captured with calibrated camera	− Laboratory setting
Grey-ball SFU-set [126] (11,346 images)	+ Largest set available + Large variety of images	− Correlation exists between images − Images are post-processed
Barcelona set [127] (83 + 126 images)	+ Uncorrelated images + High-quality XYZ-data available	− Few images − Short time-frame
Color-checker-set [102] (568 images)	+ High quality images + Uncorrected data	− Medium variety
HDR images [58, 59] (105 images)	+ High Dynamic Range images + Uncorrected data	− Few images

VII. EVALUATION

This section consists of a large-scale evaluation of various color constancy algorithms.[2] Most methods selected are evaluated using publicly available source code, ensuring the repeatability of these results. Two different datasets are used for evaluation: the gray-ball SFU set and the color-checker-set (note that the data used in this paper are obtained from [129]). These sets are selected because of their size (they are the two largest sets available to date), their nature (the sets consist of real-world images in an unconstrained environment), and their benchmark status (the gray-ball SFU set is widely used; the more recent color-checker-set has the potential to become widely used).

A. Experimental Details

Both datasets contain a marker used to obtain the ground truth. This marker is masked during all experiments. Further, all experiments are performed on linear RGB images, as the color formation model in Section II-A is based on linear images. Moreover, color constancy is generally implemented on a digital camera *prior to* the conversion of the raw data to device-dependent RGB images. Hence, using linear RGB images is basically the only logical option. The color-checker-set is available in $sRGB$-format, but Shi and Funt [129] reprocessed the raw data to obtain linear images with a higher dynamic range (14 bits as opposed to standard 8 bits). The ground truth of the gray-ball SFU set is obtained using the original images (color model of these images is $NTSC$-RGB). Therefore, we recomputed the ground truth by converting the images from $NTSC$-RGB to linear RGB assuming $\gamma = 2.2$. It is important to note that recomputing the ground truth from the gamma-correction images is different from applying gamma correction to the originally provided ground truth values of the illuminants. Due to the gamma correction, the illuminant estimation of the scenes are more chromatic, and consequently, this leads to higher angular errors.[3] This paper is the first to apply color constancy to the linear images of the gray-ball SFU set and the obtained results are, therefore, not comparable to previously published results.

All algorithms are trained using the same setup, based on cross validation. Training on the gray-ball SFU set is performed by dividing the data into 15 parts, where we ensure that the correlated images (i.e., the images of the same scene) are grouped in the same part. Next, the method is trained on 14 parts of the data and tested on the remaining part. This procedure is repeated 15 times, so every image is in the test set exactly once and all images from the same scene will either be in the training set or in the test set at the same time. The color-checker-set adopts a simpler threefold cross validation. The threefolds are provided by the authors of the dataset and to ensure repeatability of the results we did not diverge from this. This cross-validation-based procedure is also adapted to learn the optimal parameter setting for the static algorithms (optimizing p and σ) and the gamut-based algorithms (optimizing the filter size σ). Further, the regression-based method is implemented using LIBSVM [130] and is optimized for number of bins of the binary histogram and for the SVR parameters. Finally, all combination-based methods are applied to a selected set of static methods: using (9), we systematically generated nine methods using pixel values, eight methods using first-order derivatives, and seven methods using second-order derivatives. Based on the details of the corresponding methods, the following strategies are deployed. The No-N-Max combination method [105] is applied to a subset of six methods (finding the optimal combination of six methods using the same cross-validation-based procedure), the method using high-level visual information [117] is applied to the full set of methods (setting the number of semantic topics to 30), and the method using natural image statistics [106], [107] is applied to a subset of three methods (one pixel-based, one edge-based, and one second-order derivative-based method, finding the optimal combination using the same cross-validation procedure).

B. Gray-Ball SFU Set

The results[4] on the SFU set are shown in Table II, and statistical significance is demonstrated in Table IV(a). Some example results are shown in Fig. 5. Pixel-based gamut mapping performs similar to the gray-edge method, but judging from these results, simple methods like the white patch and the

[2]All estimated illuminants can be downloaded from http://www.colorconstancy.com.

[3]The new ground truth can be downloaded from http://www.colorconstancy.com.

[4]Bayesian color constancy is omitted from this table because we did not obtain satisfactory results on this dataset.

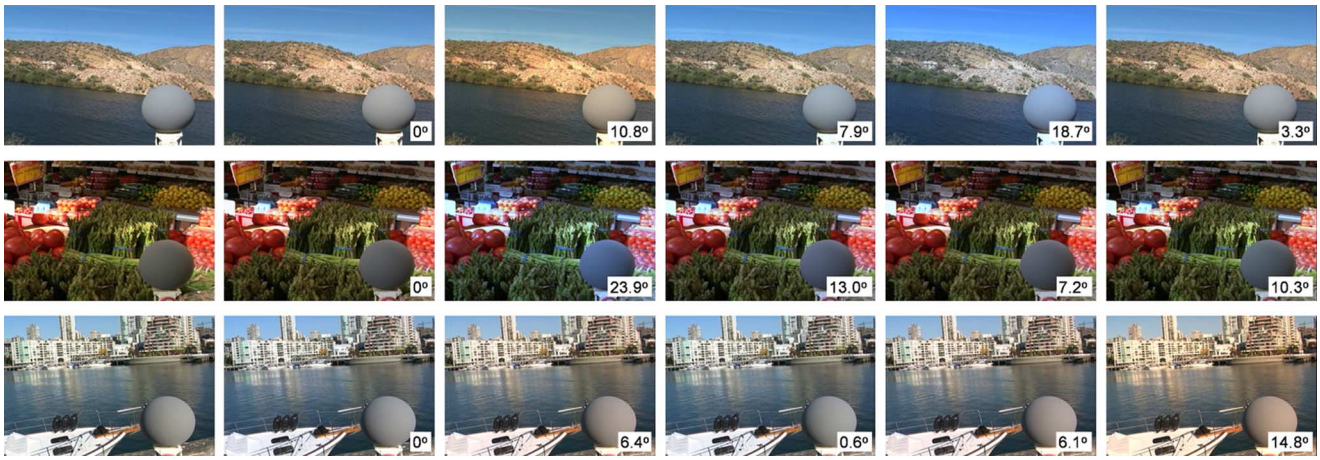

Fig. 5. Some example results of various methods applied to several test images. The angular error is shown in the bottom right corner of the images. The methods used are, from left to right, perfect color constancy using ground truth, gray world, second-order gray edge, inverse intensity chromaticity space, and high-level visual information.

TABLE II
PERFORMANCE OF SEVERAL METHODS ON THE *LINEAR* GRAY-BALL SFU SET (11 346 IMAGES)

Method	**Mean μ**	**Median**	**Trimean**	**Best-25% (μ)**	**Worst-25% (μ)**
Do Nothing	15.6°	14.0°	14.6°	2.1°	33.0°
White-Patch ($\mathbf{e}^{0,\infty,0}$)	12.7°	10.5°	11.3°	2.5°	26.2°
Grey-World ($\mathbf{e}^{0,1,0}$)	13.0°	11.0°	11.5°	3.1°	26.0°
general Grey-World ($\mathbf{e}^{0,p,\sigma}$)	12.6°	11.1°	11.6°	3.8°	23.9°
1st−order Grey-Edge ($\mathbf{e}^{1,p,\sigma}$)	11.1°	9.5°	9.8°	3.2°	21.7°
2nd−order Grey-Edge ($\mathbf{e}^{2,p,\sigma}$)	11.2°	9.6°	10.0°	3.4°	21.7°
Spatial Correlations (without reg.)	12.7°	10.8°	11.5°	2.4°	26.0°
Spatial Correlations (with reg.)	12.7°	5.3°	5.7°	1.2°	16.1°
Using Inverse Intensity Chromaticity Space	14.7°	11.0°	11.6°	3.2°	32.7°
Pixel-based Gamut Mapping	11.8°	8.9°	10.0°	2.8°	24.9°
Edge-based Gamut Mapping	13.7°	11.9°	12.3°	3.7°	26.9°
Intersection: Complete 1-jet	11.8°	8.9°	10.0°	2.8°	24.9°
Regression (SVR)	13.1°	11.2°	11.8°	4.4°	25.0°
Statistical Combination (No−N−Max)	10.3°	8.2°	8.8°	2.7°	21.2°
Using High-level Visual Information	9.7°	7.7°	8.2°	2.3°	20.6°
Using Natural Image Statistics	9.9°	7.7°	8.3°	2.4°	20.8°

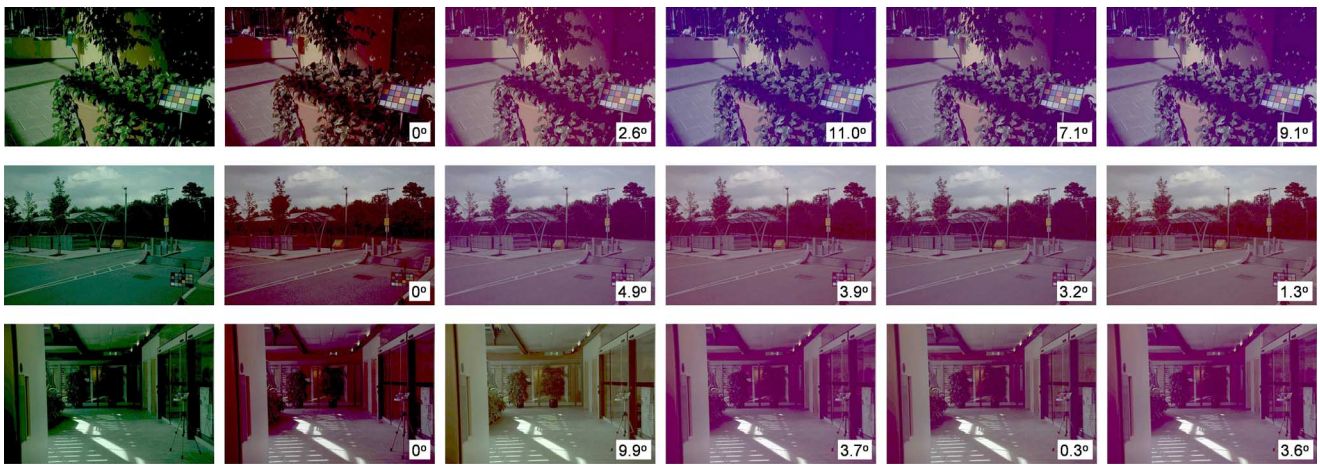

Fig. 6. Some example results of various methods applied to several test images. The angular error is shown in the bottom right corner of the images. The methods used are, from left to right, perfect color constancy using ground truth, white patch, first-order gray edge, pixel-based gamut mapping, and natural image statistics.

gray world are not suited for this dataset with the current pre-processing strategy. As expected, combination-based methods outperform single algorithms, where the difference between illuminant estimation using high-level visual information and using natural image statistics is negligible (i.e., not statistically significant).

C. *Color-Checker-Set*

The results on this dataset are shown in Table III (see Table IV(b) for statistical significance) and some example results are shown in Fig. 6. On this dataset, the edge-based methods, i.e., gray edge, spatial correlations, and edge-based

TABLE III
PERFORMANCE OF SEVERAL METHODS ON *LINEAR* COLOR-CHECKER-SET (568 IMAGES)

Method	**Mean μ**	**Median**	**Trimean**	**Best-25% (μ)**	**Worst-25% (μ)**
Do Nothing	13.7°	13.6°	13.5°	10.4°	17.2°
White-Patch ($\mathbf{e}^{0,\infty,0}$)	7.5°	5.7°	6.4°	1.5°	16.2°
Grey-World ($\mathbf{e}^{0,1,0}$)	6.4°	6.3°	6.3°	2.3°	10.6°
general Grey-World ($\mathbf{e}^{0,p,\sigma}$)	4.7°	3.5°	3.8°	1.0°	10.1°
$1^{\text{st}}-$order Grey-Edge ($\mathbf{e}^{1,p,\sigma}$)	5.4°	4.5°	4.8°	1.9°	10.0°
$2^{\text{nd}}-$order Grey-Edge ($\mathbf{e}^{2,p,\sigma}$)	5.1°	4.4°	4.6°	1.9°	10.0°
Spatial Correlations (without reg.)	5.9°	5.1°	5.4°	2.4°	10.8°
Spatial Correlations (with reg.)	4.0°	3.1°	3.3°	1.1°	8.5°
Using Inverse Intensity Chromaticity Space	13.6°	13.6°	13.5°	9.5°	18.0°
Pixel-based Gamut Mapping	4.1°	2.5°	3.0°	0.6°	10.3°
Edge-based Gamut Mapping	6.7°	5.5°	5.8°	2.1°	13.7°
Intersection: Complete 1-jet	4.1°	2.5°	3.0°	0.6°	10.3°
Bayesian	4.8°	3.5°	3.9°	1.3°	10.5°
Regression (SVR)	8.1°	6.7°	7.2°	3.3°	14.9°
Statistical Combination (No$-N-$Max)	4.3°	3.4°	3.7°	1.4°	8.5°
Using High-level Visual Information	3.5°	2.5°	2.6°	0.8°	8.0°
Using Natural Image Statistics	4.2°	3.1°	3.5°	1.0°	9.2°

TABLE IV
WILCOXON SIGN TEST ON THE TWO DATASETS. A POSITIVE VALUE (1) AT LOCATION (i, j) INDICATES THE MEDIAN OF METHOD i IS SIGNIFICANTLY LOWER THAN THE MEDIAN OF METHOD j AT THE 95% CONFIDENCE LEVEL. A NEGATIVE VALUE (-1) INDICATE THE OPPOSITE AND A ZERO (0) INDICATES THERE IS NO SIGNIFICANT DIFFERENCE BETWEEN THE TWO METHODS (A) *LINEAR* GRAY-BALL SFU SET (B) *LINEAR* COLOR-CHECKER-SET

	1. Do Nothing	2. White-Patch	3. Grey-World	4. general Grey-World	5. $1^{\text{st}}-$order Grey-Edge	6. $2^{\text{nd}}-$order Grey-Edge	7. Spatial Correlations (without reg.)	8. Spatial Correlations (with reg.)	9. Using Inverse Intensity Chromaticity Space	10. Pixel-based Gamut Mapping	11. Edge-based Gamut Mapping	12. Intersection: Complete 1-jet	14. Regression (SVR)	15. Statistical Combination (No$-N-$Max)	16. Using High-level Visual Information	17. Using Natural Image Statistics
1.	0	-1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2.	1	0	1	0	0	0	-1	1	1	-1	1	-1	1	-1	-1	-1
3.	-1	-1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	-1	-1	-1
4.	1	0	1	0	-1	-1	-1	-1	-1	-1	1	-1	1	-1	-1	-1
5.	1	0	1	1	0	1	-1	1	1	0	1	0	1	-1	-1	-1
6.	1	0	1	1	-1	0	-1	1	1	-1	1	-1	1	-1	-1	-1
7.	1	1	1	1	1	1	0	1	1	1	1	1	1	0	-1	-1
8.	1	-1	1	1	-1	-1	-1	0	0	-1	1	-1	1	-1	-1	-1
9.	1	-1	1	1	-1	-1	-1	0	0	-1	1	-1	1	-1	-1	-1
10.	1	1	1	1	0	1	-1	1	1	0	1	-1	1	-1	-1	-1
11.	1	-1	1	-1	-1	-1	-1	-1	-1	-1	0	-1	0	-1	-1	-1
12.	1	1	1	1	0	1	-1	1	1	1	1	0	1	-1	-1	-1
14.	1	-1	0	-1	-1	-1	-1	-1	-1	-1	0	-1	0	-1	-1	-1
15.	1	1	1	1	1	1	0	1	1	1	1	1	1	0	-1	-1
16.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
17.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0

(a)

	1. Do Nothing	2. White-Patch	3. Grey-World	4. general Grey-World	5. $1^{\text{st}}-$order Grey-Edge	6. $2^{\text{nd}}-$order Grey-Edge	7. Spatial Correlations (without reg.)	8. Spatial Correlations (with reg.)	9. Using Inverse Intensity Chromaticity Space	10. Pixel-based Gamut Mapping	11. Edge-based Gamut Mapping	12. Intersection: Complete 1-jet	13. Bayesian	14. Regression (SVR)	15. Statistical Combination (No$-N-$Max)	16. Using High-level Visual Information	17. Using Natural Image Statistics
1.	0	-1	-1	-1	-1	-1	-1	-1	0	-1	-1	-1	-1	-1	-1	-1	-1
2.	1	0	0	-1	0	0	0	-1	1	-1	0	-1	-1	1	-1	-1	-1
3.	1	0	0	-1	-1	-1	-1	-1	1	-1	0	-1	-1	1	-1	-1	-1
4.	1	1	1	0	1	0	1	0	1	0	1	0	1	1	0	-1	-1
5.	1	0	1	-1	0	-1	1	-1	1	-1	1	-1	-1	1	-1	-1	-1
6.	1	0	1	-1	1	0	1	-1	1	-1	1	-1	-1	1	-1	-1	-1
7.	1	0	1	-1	-1	-1	0	-1	1	-1	0	-1	-1	1	-1	-1	-1
8.	1	1	1	0	1	1	1	0	1	-1	1	-1	1	1	1	-1	0
9.	0	-1	-1	-1	-1	-1	-1	-1	0	-1	-1	-1	-1	-1	-1	-1	-1
10.	1	1	1	0	1	1	1	1	1	0	1	0	1	1	1	0	0
11.	1	0	0	-1	-1	-1	0	-1	1	-1	0	-1	-1	1	-1	-1	-1
12.	1	1	1	0	1	1	1	1	1	0	1	0	1	1	1	-1	0
13.	1	1	1	-1	1	1	1	-1	1	-1	1	-1	0	1	0	-1	-1
14.	1	-1	-1	-1	-1	-1	-1	-1	1	-1	-1	-1	-1	0	-1	-1	-1
15.	1	1	1	0	1	1	1	-1	1	-1	1	-1	0	1	0	-1	0
16.	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1
17.	1	1	1	1	1	1	1	0	1	0	1	0	1	1	0	-1	0

(b)

gamut mapping, perform significantly worse than pixel-based methods like gamut mapping and general gray world. However, it can be observed that the error on "difficult" images (i.e., images on which the method estimates an inaccurate illuminant, the *Worst*-25% column) for both types of algorithms is similar. This indicates that the performance of methods using low-level information (either static algorithms or learning-based methods) is bounded by the information that is present. Using multiple algorithms is required to decrease the error of these "difficult" images, as can be seen by the performance of combination-based methods. Even though the increase in overall performance is not very high, methods using high-level visual information and natural image statistics are statistically similar to the pixel-based gamut mapping; the largest improvement in accuracy is obtained on these difficult images (the mean angular error on the worst 25% of the images drops from 10.3° to 8.0° and 9.2°, respectively). Hence, to arrive at a robust color constancy algorithm that is able to accurately estimate the illuminant on any type of image, it is necessary to combine several approaches.

VIII. DISCUSSION AND FUTURE DIRECTIONS

In this paper, an overview of often used approaches to illuminant estimation is presented, together with recent developments. Criteria that are important for computational color constancy algorithms are the requirement of training data, the accuracy of the estimation, the computational runtime of the method, the

TABLE V
SUMMARY OF METHODS WITH ADVANTAGES AND DISADVANTAGES

Method	Ex.	Pros	Cons
Static (using low-level statistics)	[49] – [62]	+ Simple to implement + Accurate for adequate parameters + Fast execution	– Opaque parameter selection – Inaccurate for inferior parameters
Static (physics-based)	[63] – [69]	+ No training phase + Fast execution + Few parameters	– Difficult to implement – Mediocre performance
Gamut-based	[70] – [81]	+ Elegant underlying theory + Potentially high accuracy	– Requires training data – Difficult to implement – Requires proper preprocessing
Learning-based (using low-level statistics)	[82] – [96]	+ Tunable for specific data set + Simple to implement	– Requires training data – Non-intuitive – Slow execution
Learning-based (using higher-level statistics)	[97] – [104]	+ Potentially high accuracy + Intuitive	– Requires training data – Inherently slower than single methods – Difficult to implement
Learning-based (using semantics)	[105] – [111]	+ Potentially high accuracy + Incorporates semantics	– Requires training data – Difficult to implement – Slow execution

transparency of the approach, the complexity of the implementation, and the number of tunable parameters. A summary of the discussed methods is presented in Table V. In Section III-A, methods that are based on low-level information are presented. These methods are not dependent on training data and the parameters are not dependent on the input data and are, therefore, called *static*. Existing methods include the gray world and the white patch and recent developments extended these methods to incorporate higher order statistics. Advantages of such methods are a simple implementation (often, merely a few lines of code are required) and fast execution. Further, the accuracy of the estimations can be quite high, provided the parameters are selected appropriately. On the other hand, inaccurate parameter selection can severely reduce the performance. Moreover, the selection of the optimal parameters is quite opaque, especially without prior knowledge on the input data. Physics-based methods discussed in Section III-B suffer less from the parameter selection but are also less accurate (even for properly selected parameters).

In Section IV, the gamut-based methods and an extension to incorporate the differential nature of images are described. The main advantages of gamut-based methods are the elegant underlying theory and the potential high accuracy. However, proper implementation requires some effort and appropriate preprocessing can severely influence the accuracy.

Finally, in Section V, methods that cannot operate without training phase are discussed. Section V-A discusses methods that learn low-level statistics, like regression techniques and Bayesian approaches. Advantages of such methods are that they are (relatively) simple to implement and that they can be tuned toward specific data (like indoor or outdoor images). Disadvantages are that the output is often rather nonintuitive since the model that is learned is quite opaque. On the other hand, methods using higher level statistics and semantics discussed in Sections V-B and V-C, like the selection algorithm using natural image statistics, are often quite intuitive since it can be predicted beforehand which method will be selected for a specific input image. Moreover, the accuracy of such approaches has been proven to be state of the art. However, the use of multiple single algorithms means it is inherently slower than the single algorithms themselves.

A. Future Directions

As explained in Section I, explanations for human color constancy and computational approaches are diverging. It would be interesting to bring the recent advances in human color constancy closer to the computational level, to map the computational advancements to human explanations. For instance, Gijsenij and Gevers [106], [107] suggest that the optimal computational approach that is taken for a specific image is based on the statistics of the scene. It is unknown to what extent human observers use a similar approach; if multiple cues are available then in what order are they processed and what weight is given to each cue? On the other hand, recent developments in human color constancy suggest that color memory, possibly in addition to contextual clues, could play an important role [131]–[134]. It is worth exploring the corresponding computational approaches, if they exist.

A first step toward convergence of human and computational color constancy might be to adapt a new image correction model. It is suggested that human color constancy is relational rather than absolute and recently experiments are performed that indicate that human observers do not explicitly judge the color of the illuminant [135]. Mapping this to computational methods would imply that the two-stage setup (illuminant estimation followed by image correction) is not appropriate. According to this theory, methods closer to human color constancy are methods that learn correspondences between images and possible illuminants [136]. This implies that the experimental setup be changed, as the illuminant is not explicitly estimated.

Finally, all methods discussed so far are based on the assumption that the illuminant in the image is spatially uniform. However, in real-life scenarios, this assumption is easily violated: indoor images can depict multiple rooms in the same image, while all rooms can have spectrally different light sources. Furthermore, outdoor images can show parts of the scene in shadow and other parts in bright sunlight. For simplification, such examples are ignored by most current approaches. Only a few methods have been proposed that consider the presence of multiple light sources. For instance, Finlayson *et al.* [137] and Barnard *et al.* [138] propose a Retinex-based approach that explicitly assumes

that surfaces exist in the scene that are illuminated by multiple light sources. Another Retinex-based approach [25] uses stereo images to derive 3-D information on the surfaces that are present in the images, to be able to distinguish material transitions from local light color changes. In [139], human interaction is employed to specify locations in images that are illuminated by different light sources. Finally, Ebner [57] proposes a gray-world-like method that is based on the assumption that the light source smoothly varies across the scene. An additional difficulty of this line of research is the lack of ground truth. In Section VI, several databases are described, all based on the assumption that there is only one light source in the scene. Consequently, before proposing new methods that depart from the uniform assumption, new proper databases for evaluation have to be designed.

To conclude, interesting future research directions include investigation of the relation between human and computational color constancy theories, adopting alternate image correction models besides the two-stage approach used in this paper and departure of the uniform light source assumption.

ACKNOWLEDGMENT

The authors acknowledge the advice of J. Vazquez-Corral on the calibration of the gray-ball SFU set.

REFERENCES

[1] J. Yang, R. Stiefelhagen, U. Meier, and A. Waibel, "Visual tracking for multimodal human computer interaction," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 1998, pp. 140–147.

[2] T. Gevers and A. Smeulders, "Pictoseek: Combining color and shape invariant features for image retrieval," *IEEE Trans. Image Process.*, vol. 9, no. 1, pp. 102–119, Jan. 2000.

[3] M. Fairchild, *Color Appearance Models (Wiley-IS&T Series in Imaging Science and Technology)*, 2nd ed. Chichester, U.K.: Wiley, 2005.

[4] L. Arend, A. Reeves, J. Schirillo, and R. Goldstein, "Simultaneous color constancy: Papers with diverse munsell values," *J. Opt. Soc. Amer. A*, vol. 8, no. 4, pp. 661–672, 1991.

[5] D. Brainard, J. Kraft, and P. Longere, "Color constancy: Developing empirical tests of computational models," in *Colour Perception: From Light to Object*, R. Mausfeld and D. Heyer, Eds. London, U.K.: Oxford Univ. Press, 2003, pp. 307–334.

[6] P. Delahunt and D. Brainard, "Does human color constancy incorporate the statistical regularity of natural daylight?," *J. Vis.*, vol. 4, no. 2, pp. 57–81, 2004.

[7] D. Foster, K. Amano, and S. Nascimento, "Color constancy in natural scenes explained by global image statistics," *Vis. Neurosci.*, vol. 23, no. 3–4, pp. 341–349, 2006.

[8] D. Foster, "Color constancy," *Vis. Res.*, vol. 51, pp. 674–700, 2010.

[9] E. Land and J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Amer. A*, vol. 61, pp. 1–11, 1971.

[10] E. Land, "The retinex theory of color vision," *Sci. Amer.*, vol. 237, no. 6, pp. 108–128, Dec. 1977.

[11] E. Land, "Recent advances in retinex theory," *Vis. Res.*, vol. VR-26, pp. 7–21, 1986.

[12] D. Jobson, Z. Rahman, and G. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. Image Process.*, vol. 6, no. 3, pp. 451–462, Mar. 1997.

[13] D. Jobson, Z. Rahman, and G. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–976, Jul. 1997.

[14] E. Provenzi, C. Gatta, M. Fierro, and A. Rizzi, "A spatially variant white-patch and gray-world method for color image enhancement driven by local contrast," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 10, pp. 1757–1770, Oct. 2008.

[15] J. Kraft and D. Brainard, "Mechanisms of color constancy under nearly natural viewing," in *Proc. Nat. Acad. Sci.*, 1999, vol. 96, pp. 307–312.

[16] J. Golz and D. MacLeod, "Influence of scene statistics on colour constancy," *Nature*, vol. 415, pp. 637–640, 2002.

[17] J. Golz, "The role of chromatic scene statistics in color constancy: Spatial integration," *J. Vis.*, vol. 8, no. 13, pp. 1–16, 2008.

[18] F. Ciurea and B. Funt, "Failure of luminance-redness correlation for illuminant estimation," in *Proc. Color Imag. Conf.*, 2004, pp. 42–46.

[19] B. Funt and G. Finlayson, "Color constant color indexing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 5, pp. 522–529, May 1995.

[20] T. Gevers and A. Smeulders, "Color-based object recognition," *Pattern Recognit.*, vol. 32, no. 3, pp. 453–464, 1999.

[21] J. Geusebroek, R. van den Boomgaard, A. Smeulders, and T. Gevers, "Color constancy from physical principles," *Pattern Recognit. Lett.*, vol. 24, no. 11, pp. 1653–1662, 2003.

[22] G. Finlayson, S. Hordley, and P. Morovic, "Colour constancy using the chromagenic constraint," in *Proc IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 1079–1086.

[23] G. Fischer and M. Sajjaa, "Whitebalpr—A new method for automatic white balance," in *Proc. IS&T's Eur. Conf. Color Graph., Imag. Vis.*, Jun. 2008, pp. 202–207.

[24] C. Fredembach and G. Finlayson, "The bright-chromagenic algorithm for illuminant estimation," *J. Imag. Sci. Technol.*, vol. 52, no. 4, pp. 040906-1–040908-11, 2008.

[25] W. Xiong and B. Funt, "Stereo retinex," *Image Vis. Comput.*, vol. 27, no. 1–2, pp. 178–188, 2009.

[26] M. Lecca and S. Messelodi, "Computing von kries illuminant changes by piecewise inversion of cumulative color histograms," *Electron. Lett. Comput. Vis. Image Anal.*, vol. 8, no. 2, pp. 1–17, 2009.

[27] M. Sajjaa and G. Fisher, "Automatic white balance: Whitebalpr using the dichromatic reflection model," in *Proc. SPIE*, 2009, vol. 7250, pp. 72500D–.

[28] J. Nieves, C. Plata, E. Valero, and J. Romero, "Unsupervised illuminant estimation from natural scenes: An RGB digital camera suffices," *Appl. Opt.*, vol. 47, no. 20, pp. 3574–3584, 2008.

[29] F. Zaraga and G. Langfelder, "White balance by tunable spectral responsivities," *J. Opt. Soc. Amer. A*, vol. 27, no. 1, pp. 31–39, 2010.

[30] J. Renno, D. Makris, T. Ellis, and G. Jones, "Application and evaluation of colour constancy in visual surveillance," in *Proc. 14th Int. Conf. Comput. Commun. Netw.*, 2005, pp. 301–308.

[31] R. Malik and P. Bajcsy, "Achieving color constancy across multiple cameras," in *Proc. ACM Int. Conf. Multimedia*, 2009, pp. 893–896.

[32] B. Funt, M. Drew, and J. Ho, "Color constancy from mutual reflection," *Int. J. Comput. Vis.*, vol. 6, no. 1, pp. 5–24, 1991.

[33] M. D'Zmura and G. Iverson, "Color constancy. I. Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces," *J. Opt. Soc. Amer. A*, vol. 10, no. 10, pp. 2148–2165, 1993.

[34] M. D'Zmura and G. Iverson, "Color constancy. II. Results for two-stage linear recovery of spectral descriptions for lights and surfaces," *J. Opt. Soc. Amer. A*, vol. 10, no. 10, pp. 2166–2180, 1993.

[35] L. Maloney and B. Wandell, "Color constancy: A method for recovering surface spectral reflectance," *J. Opt. Soc. Amer. A*, vol. OSA-3, no. 1, pp. 29–33, 1986.

[36] S. Shafer, "Using color to separate reflection components," *Color Res. Appl.*, vol. 10, no. 4, pp. 210–218, 1985.

[37] G. Klinker, S. Shafer, and T. Kanade, "A physical approach to color image understanding," *Int. J. Comput. Vis.*, vol. 4, no. 1, pp. 7–38, 1990.

[38] G. West and M. Brill, "Necessary and sufficient conditions for von kries chromatic adaptation to give color constancy," *J. Math. Biol.*, vol. 15, no. 2, pp. 249–258, 1982.

[39] G. Finlayson, M. Drew, and B. Funt, "Color constancy: Generalized diagonal transforms suffice," *J. Opt. Soc. Amer. A*, vol. 11, no. 11, pp. 3011–3019, 1994.

[40] B. Funt and B. Lewis, "Diagonal versus affine transformations for color correction," *J. Opt. Soc. Amer. A*, vol. 17, no. 11, pp. 2108–2112, 2000.

[41] K. Lam, "Metamerism and Colour Constancy," Ph.D. dessertation, Univ. Bradford, Bradford, U.K., 1985.

[42] N. Moroney, M. Fairchild, R. Hunt, C. Li, M. Luo, and T. Newman, "The ciecam02 color appearance model," in *Proc. IS&T/SID's Color Imag. Conf.*, 2002, pp. 23–27.

[43] J. von Kries, "Influence of adaptation on the effects produced by luminous stimuli," in *Sources of Color Vision*, D. MacAdam, Ed. Cambridge, MA: MIT Press, 1970, pp. 109–119.

[44] B. Funt and H. Jiang, "Non-diagonal colour correction," in *Proc. IEEE Int. Conf. Image Process.*, Barcelona, Spain, Sep. 2003, pp. 539–544.

[45] H. Chong, S. Gortler, and T. Zickler, "The von kries hypothesis and a basis for color constancy," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2007, pp. 1–8.

[46] G. Finlayson, M. Drew, and B. Funt, "Spectral sharpening: Sensor transformations for improved color constancy," *J. Opt. Soc. Amer. A*, vol. 11, no. 5, pp. 1553–1563, 1994.

[47] K. Barnard, F. Ciurea, and B. Funt, "Sensor sharpening for computational color constancy," *J. Opt. Soc. Amer. A*, vol. 18, no. 11, pp. 2728–2743, 2001.

[48] J. Worthey and M. Brill, "Heuristic analysis of von kries color constancy," *J. Opt. Soc. Amer. A*, vol. OSA-3, no. 10, pp. 1708–1712, 1986.

[49] M. Brill, "Minimal von kries illuminant invariance," *Color Res. Appl.*, vol. 33, no. 4, pp. 320–323, 2008.

[50] G. Buchsbaum, "A spatial processor model for object colour perception," *J. Franklin Inst.*, vol. FI-310, no. 1, pp. 1–26, Jul. 1980.

[51] R. Gershon, A. Jepson, and J. Tsotsos, "From [r, g, b] to surface reflectance: Computing color constant descriptors in images," in *Proc. Int. Joint Conf. Artif. Intell.*, Milan, Italy, 1987, pp. 755–758.

[52] K. Barnard, L. Martin, A. Coath, and B. Funt, "A comparison of computational color constancy algorithms—Part II: Experiments with image data," *IEEE Trans. Image Process.*, vol. 11, no. 9, pp. 985–996, Sep. 2002.

[53] W. Xiong, B. Funt, L. Shi, S. Kim, B. Kang, S. Lee, and C. Kim, "Automatic white balancing via gray surface identification," in *Proc. IS&T/SID's Color Imag. Conf.*, 2007, pp. 143–146.

[54] W. Xiong and B. Funt, "Cluster based color constancy," in *Proc. IS&T/SID's Color Imag. Conf.*, 2008, pp. 210–214.

[55] B. Li, D. Xu, W. Xiong, and S. Feng, "Color constancy using achromatic surface," *Color Res. Appl.*, vol. 35, no. 4, pp. 304–312, 2010.

[56] A. Gijsenij and T. Gevers, "Color constancy by local averaging," in *Proc. Comput. Color Imag. Workshop, Conjunct. Int. Conf. Image Anal. Process. 2007*, Modena, Italy, Sep. 2007, pp. 1–4.

[57] M. Ebner, "Color constancy based on local space average color," *Mach. Vis. Appl.*, vol. 20, no. 5, pp. 283–301, 2009.

[58] B. Funt and L. Shi, "The rehabilitation of maxRGB," presented at the presented at the IS&T/SID's Color Imag. Conf., San Antonio, TX, 2010.

[59] B. Funt and L. Shi, "The effect of exposure on maxRGB color constancy," presented at the presented at the Proc. SPIE, San Jose, CA, 2010 vol. 7527.

[60] G. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Proc. IS&T/SID's Color Imag. Conf.*, 2004, pp. 37–41.

[61] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2207–2214, Sep. 2007.

[62] W. Freeman and E. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 9, pp. 891–906, Sep. 1991.

[63] H. Chen, C. Shen, and P. Tsai, "Edge-based automatic white balancing linear illuminant constraint," in *Proc. Vis. Commun. Image Process.*, 2007, p. 64081D.

[64] A. Chakrabarti, K. Hirakawa, and T. Zickler, "Color constancy beyond bags of pixels," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.

[65] A. Gijsenij, T. Gevers, and J. v. d. Weijer, "Physics-based edge evaluation for improved color constancy," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 581–588.

[66] H. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," *J. Opt. Soc. Amer. A*, vol. 3, no. OSA-10, pp. 1694–1699, 1986.

[67] S. Tominaga and B. Wandell, "Standard surface-reflectance model and illuminant estimation," *J. Opt. Soc. Amer. A*, vol. 6, no. 4, pp. 576–584, 1989.

[68] G. Healey, "Estimating spectral reflectance using highlights," *Image Vis. Comput.*, vol. 9, no. 5, pp. 333–337, 1991.

[69] R. Tan, K. Nishino, and K. Ikeuchi, "Color constancy through inverse-intensity chromaticity space," *J. Opt. Soc. Amer. A*, vol. 21, no. 3, pp. 321–334, 2004.

[70] G. Finlayson and G. Schaefer, "Solving for colour constancy using a constrained dichromatic reflection model," *Int. J. Comput. Vis.*, vol. 42, no. 3, pp. 127–144, 2001.

[71] J. Toro and B. B. V. Funt, "A multilinear constraint on dichromatic planes for illumination estimation," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 92–97, Jan. 2007.

[72] J. Toro, "Dichromatic illumination estimation without pre-segmentation," *Pattern Recognit. Lett.*, vol. 29, no. 7, pp. 871–877, 2008.

[73] L. Shi and B. Funt, "Dichromatic illumination estimation via hough transforms in 3D," in *Proc. IS&T's Eur. Conf. Color Graph., Imag. Vis.*, Barcelona, Spain, 2008.

[74] D. Forsyth, "A novel algorithm for color constancy," *Int. J. Comput. Vis.*, vol. 5, no. 1, pp. 5–36, 1990.

[75] K. Barnard, "Improvements to gamut mapping colour constancy algorithms," in *Proc. Eur. Conf. Comput. Vis.*, 2000, pp. 390–403.

[76] G. Finalyson, "Color in perspective," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 10, pp. 1034–1038, Oct. 1996.

[77] G. Finlayson and S. Hordley, "Improving gamut mapping color constancy," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1774–1783, Oct. 2000.

[78] G. Finalyson and S. Hordley, "Selection for gamut mapping colour constancy," *Image Vis. Comput.*, vol. 17, no. 8, pp. 597–604, 1999.

[79] G. Finlayson and R. Xu, "Convex programming color constancy," in *Proc. IEEE Workshop Color Photometric Methods Comput. Vis., Conjunct. ICCV'03*, 2003, pp. 1–8.

[80] M. Mosny and B. Funt, "Cubical gamut mapping colour constancy," presented at the presented at the IS&T's Eur. Conf. Color Graph., Imag. Vis., Joensuu, Finland, 2010.

[81] K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms—Part I: Methodology and experiments with synthesized data," *IEEE Trans. Image Process.*, vol. 11, no. 9, pp. 972–984, Sep. 2002.

[82] K. Barnard and B. Funt, "Color constancy with specular and non-specular surfaces," in *Proc. IS&T/SID's Color Imag. Conf.*, 1999, pp. 114–119.

[83] S. Tominaga, S. Ebisui, and B. Wandell, "Scene illuminant classification: Brighter is better," *J. Opt. Soc. Amer. A*, vol. 18, no. 1, pp. 55–64, 2001.

[84] G. Finlayson, S. Hordley, and R. Xu, "Convex programming colour constancy with a diagonal-offset model," in *Proc. IEEE Int. Conf. Image Process.*, 2005, pp. 948–951.

[85] A. Gijsenij, T. Gevers, and J. van de Weijer, "Generalized gamut mapping using image derivative structures for color constancy," *Int. J. Comput. Vis.*, vol. 86, no. 1–2, pp. 127–139, 2010.

[86] G. Finlayson, S. Hordley, and I. Tastl, "Gamut constrained illuminant estimation," *Int. J. Comput. Vis.*, vol. 67, no. 1, pp. 93–109, 2006.

[87] V. Cardei, B. Funt, and K. Barnard, "Estimating the scene illumination chromaticity using a neural network," *J. Opt. Soc. Amer. A*, vol. 19, no. 12, pp. 2374–2386, 2002.

[88] B. Funt and W. Xiong, "Estimating illumination chromaticity via support vector regression," in *Proc. IS&T/SID's Color Imag. Conf.*, 2004, pp. 47–52.

[89] W. Xiong and B. Funt, "Estimating illumination chromaticity via support vector regression," *J. Imag. Sci. Technol.*, vol. 50, no. 4, pp. 341–348, 2006.

[90] N. Wang, D. Xu, and B. Li, "Edge-based color constancy via support vector regression," *IEICE Trans. Inf. Syst.*, vol. E92-D, no. 11, pp. 2279–2282, 2009.

[91] V. Agarwal, A. Gribok, A. Koschan, and M. Abidi, "Estimating illumination chromaticity via kernel regression," in *Proc. IEEE Int. Conf. Image Process.*, 2006, pp. 981–984.

[92] V. Agarwal, A. Gribok, and M. Abidi, "Machine learning approach to color constancy," *Neural Netw.*, vol. 20, no. 5, pp. 559–563, 2007.

[93] V. Agarwal, A. Gribok, A. Koschan, B. Abidi, and M. Abidi, "Illumination chromaticity estimation using linear learning methods," *J. Pattern Recognit. Res.*, vol. 4, no. 1, pp. 92–109, 2009.

[94] W. Xiong, L. Shi, B. Funt, S. Kim, B. Kan, and S. Lee, "Illumination estimation via thin-plate spline interpolation," in *Proc. IS&T/SID's Color Imag. Conf.*, 2007, pp. 25–29.

[95] G. Finlayson, S. Hordley, and P. Hubel, "Color by correlation: A simple, unifying framework for color constancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1209–1221, Nov. 2001.

[96] C. Rosenberg, M. Hebert, and S. Thrun, "Color constancy using KL-divergence," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2001, pp. 239–246.

[97] M. D'Zmura, G. Iverson, and B. Singer, "Probabilistic color constancy," in *Geometric Representations of Perceptual Phenomena*. Mahwah, NJ: Lawrence Erlbaum Associates, 1995, pp. 187–202.

[98] D. Brainard and W. Freeman, "Bayesian color constancy," *J. Opt. Soc. Amer. A*, vol. 14, pp. 1393–1411, 1997.

[99] G. Sapiro, "Color and illuminant voting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 11, pp. 1210–1215, Nov. 1999.

[100] Y. Tsin, R. Collins, V. Ramesh, and T. Kanade, "Bayesian color constancy for outdoor object recognition," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2001, pp. 1132–1139.

[101] C. Rosenberg, T. Minka, and A. Ladsariya, "Bayesian color constancy with non-Gaussian models," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2003.

[102] P. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp, "Bayesian color constancy revisited," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8 [Online]. Available: http://www.kyb.tuebingen.mpg.de/bs/people/pgehler/colour/

[103] V. Cardei and B. Funt, "Committee-based color constancy," in *Proc. IS&T/SID's Color Imag. Conf.*, 1999, pp. 311–313.
[104] G. Schaefer, S. Hordley, and G. Finlayson, "A combined physical and statistical approach to colour constancy," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Washington, DC, 2005, pp. 148–153.
[105] S. Bianco, F. Gasparini, and R. Schettini, "Consensus-based framework for illuminant chromaticity estimation," *J. Electron. Imag.*, vol. 17, no. 2, pp. 023013-1–023013-9, 2008.
[106] A. Gijsenij and T. Gevers, "Color constancy using natural image statistics," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Minneapolis, MN, Jun. 2007, pp. 1–8.
[107] A. Gijsenij and T. Gevers, "Color constancy using natural image statistics and scene semantics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 687–698, Apr. 2011.
[108] B. Li, D. Xu, and C. Lang, "Colour constancy based on texture similarity for natural images," *Coloration Technol.*, vol. 125, no. 6, pp. 328–333, 2009.
[109] S. Bianco, F. Gasparini, and R. Schettini, "Region-based illuminant estimation for effective color correction," in *Proc. Int. Conf. Image Anal. Process.*, 2009, pp. 43–52.
[110] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Automatic color constancy algorithm selection and combination," *Pattern Recognit.*, vol. 43, no. 3, pp. 695–705, 2010.
[111] M. Wu, J. Sun, J. Zhou, and G. Xue, "Color constancy based on texture pyramid matching and regularized local regression," *J. Opt. Soc. Amer. A*, vol. 27, no. 10, pp. 2097–2105, 2010.
[112] A. Gijsenij and T. Gevers, "Color constancy using image regions," in *Proc. IEEE Int. Conf. Image Process.*, San Antonio, TX, , Sep. 2007, pp. III-501–III-504.
[113] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Improving color constancy using indoor-outdoor image classification," *IEEE Trans. Image Process.*, vol. 17, no. 12, pp. 2381–2392, Dec. 2008.
[114] R. Lu, A. Gijsenij, T. Gevers, K. van de Sande, J. Geusebroek, and D. Xu, "Color constancy using stage classification," in *Proc. IEEE Int. Conf. Image Process.*, 2009, pp. 685–688.
[115] R. Lu, A. Gijsenij, T. Gevers, D. Xu, V. Nedovic, and J. Geusebroek, "Color constancy using 3D stage geometry," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2009, pp. 1756–1749.
[116] V. Nedovic, A. Smeulders, A. Redert, and J. Geusebroek, "Stages as models of scene geometry," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1673–1687, Sep. 2010.
[117] J. van de Weijer, C. Schmid, and J. Verbeek, "Using high-level visual information for color constancy," in *Proc. IEEE Int. Conf. Comput. Vis.*, Rio de Janeiro, Brazil, 2007, pp. 8–1.
[118] E. Rahtu, J. Nikkanen, J. Kannala, L. Lepistö, and J. Heikkilä, "Applying visual object categorization and memory colors for automatic color constancy," in *Proc. Int. Conf. Image Anal. Process.*, 2009, pp. 873–882.
[119] B. Funt, K. Barnard, and L. Martin, "Is machine colour constancy good enough?," in *Proc. Eur. Conf. Comput. Vis.*, 1998, pp. 445–459.
[120] A. Gijsenij, T. Gevers, and M. Lucassen, "A perceptual analysis of distance measures for color constancy algorithms," *J. Opt. Soc. Amer. A*, vol. 26, no. 10, pp. 2243–2256, 2009.
[121] S. Hordley and G. Finlayson, "Reevaluation of color constancy algorithm performance," *J. Opt. Soc. Amer. A*, vol. 23, no. 5, pp. 1008–1020, May 2006.
[122] K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for color research," *Color Res. Appl.*, vol. 27, no. 3, pp. 147–151, 2002.
[123] S. Nascimento, F. Ferreira, and D. Foster, "Statistics of spatial cone-excitation ratios in natural scenes," *J. Opt. Soc. Amer. A*, vol. 19, no. 8, pp. 1484–1490, 2002.
[124] D. Foster, S. Nascimento, and K. Amano, "Information limits on neural identification of colored surfaces in natural scenes," *Vis. Neurosci.*, vol. 21, pp. 331–336, 2004.
[125] C. Párraga, G. Brelstaff, T. Troscianko, and I. Moorehead, "Color and luminance information in natural scenes," *J. Opt. Soc. Amer. A*, vol. 15, no. 3, pp. 563–569, 1998.
[126] F. Ciurea and B. Funt, "A large image database for color constancy research," in *Proc. IS&T/SID's Color Imag. Conf.*, 2003, pp. 160–164 [Online]. Available: http://www.cs.sfu.ca/~colour/data/gray_ball/
[127] J. Vazquez-Corral, C. Párraga, M. Vanrell, and R. Baldrich, "Color constancy algorithms: Psychophysical rvaluation on a new dataset," *J. Imag. Sci. Technol.*, vol. 53, no. 3, pp. 1–9, 2009.
[128] C. Párraga, J. Vazquez-Corral, and M. Vanrell, "A new cone activation-based natural images dataset," *Perception*, vol. 36, no. Suppl, pp. 180–180, 2009.
[129] L. Shi and B. V. Funt., Re-Processed Version of the Gehler Color Constancy Database of 568 Images Nov. 1, 2010 [Online]. Available: http://www.cs.sfu.ca/~colour/data/
[130] C.-C. Chang and C.-J. Lin., LIBSVM: A Library for Support Vector Machines 2001 [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm
[131] P. Bradley, "Constancy, categories and bayes: A new approach to representational theories of color constancy," *Philosoph. Psychol.*, vol. 21, no. 5, pp. 601–627, 2008.
[132] V. de Almeida and S. Nascimento, "Perception of illuminant colour changes across real scene," *Perception*, vol. 38, no. 8, pp. 1109–1117, 2009.
[133] T. Lin and C. Sun, "Representation or context as a cognitive strategy in colour constancy?," *Perception*, vol. 37, no. 9, pp. 1353–1367, 2008.
[134] M. Hedrich, M. Bloj, and A. Ruppertsberg, "Color constancy improves for real 3D objects," *J. Vis.*, vol. 9, no. 4, pp. 1–16, 2009.
[135] J. Granzier, E. Brenner, and J. Smeets, "Can illumination estimates provide the basis for color constancy?," *J. Vis.*, vol. 9, no. 3, pp. 1–11, 2009.
[136] T. Moerland and F. Jurie, "Learned color constancy from local correspondences," in *Proc. IEEE Int. Conf. Multimedia Expo.*, 2005, pp. 820–823.
[137] G. Finlayson, B. Funt, and K. Barnard, "Color constancy under varying illumination," in *Proc. IEEE Int. Conf. Comput. Vis.*, Washington, DC, 1995, pp. 720–725.
[138] K. Barnard, G. Finlayson, and B. Funt, "Color constancy for scenes with varying illumination," *Comput. Vis. Image Understanding*, vol. 65, no. 2, pp. 311–321, 1997.
[139] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand, "Light mixture estimation for spatially varying white balance," in *Proc. ACM SIGGRAPH*, 2008, pp. 1–7.

Arjan Gijsenij (M'07) received the M.Sc. degree in computer science from the University of Groningen, Groningen, The Netherlands, in 2005, and the Ph.D. degree from the University of Amsterdam, Amsterdam, The Netherlands, in 2010.

He is currently a Postdoctoral Researcher at the University of Amsterdam. He has published papers in high-impact journals and conferences and served on the Program Committee of several conferences and workshops. His research interests include the field of color in computer vision and psychophysics.

Theo Gevers (M'02) received the Ph.D. degree in 1996 from the University of Amsterdam, Amsterdam, The Netherlands.

He is currently an Associate Professor of computer science and a Teaching Director of the M.Sc. in artificial intelligence at the University of Amsterdam, Amsterdam, The Netherlands. He is also a Full Professor at the Computer Vision Center, Universitat Autònoma de Barcelona, Barcelona, Spain. His research interests include the fundamentals of color image processing, image understanding, and human behavior analysis. He is the author or coauthor of more than 100 papers. He is a Program Committee Member of a various number of conferences and an invited speaker at major conferences. He is a Lecturer of postdoctoral courses given at various major conferences (CVPR, ICPR, SPIE, CGIV).

Prof. Gevers is a recipient of the VICI-Award (for excellent researchers) from the Dutch Organisation for Scientific Research.

Joost van de Weijer (M'00) received the M.Sc. degree from the Delft University of Technology, Delft, The Netherlands, in 1998 and the Ph.D. degree from the University of Amsterdam, Amsterdam, The Netherlands, in 2005.

He was a Marie Curie Intra-European Fellow in INRIA Rhone-Alpes. Since 2008, he has been with the Computer Vision Center, Universitat Autònoma de Barcelona, Barcelona, Spain.

Dr. van de Weijer was a recipient of Ramon y Cajal Research Fellowships in computer science funded by the Spanish Ministry of Science and Technology.