

Chapter 2

State-of-the-Art Survey on Color Constancy Methods

2.1 Introduction

Intuitively, color constancy is understood as the task of finding descriptors for the surfaces in a scene which are invariant to illumination changes. However, the job of correcting colors in an image can be considered as two separate stages. First, estimating the parameters of the scene illuminant which will be used in a second stage to compute the set of *illuminant-independent* descriptors from the image pixels [MW86].

Illumination-independent descriptors can be used afterwards in a wide span of problems in machine vision such as object recognition and image indexing. In these applications usually two images picturing the same scene or object may present some color misadjustments since the incident illumination may have varied from one image to the other. This way, any of the processes above can become even more prone to error. While some authors simply restrict the color constancy definition to the first stage, i.e., estimation of the scene illuminant, others are more interested in finding a transformation between image colors in order to keep them as resembling as possible to those under some reference (*canonic*) light conditions. This less restrictive definition of color constancy, introduced by Forsyth in [For90], is named by some authors also as *color correction*.

These two approaches are fully interchangeable once models of color formation and variation are specified. If the illuminant parameters are obtained somehow, it is then possible to compute the mapping between two different light conditions. Reversely, if a certain mapping is found between the canonic illumination and an unknown one, it is also viable to compute an approximation to the parameters of the latter light by applying the color mapping to the canonic illuminant. Depending on the change model it may be necessary to own further information to fully reverse these two color constancy conceptions, such as the information about the lights or the likeliest surfaces which will be found in the scene.

The diagonal model is by far the most used model of color change. This model transforms colors under an illuminant onto others taken under another

illuminant only by scaling every color channel independently. The efficiency of this restricted linear model is in great extent a function of the sensors in the vision system. Specifically, whether the sensors are narrow-band or their sensibility functions overlap [FDF93a, FDF93b, FDF94a, Fin95a, Fin95b, FF96]. In case the sensors are not completely narrow-band, it is claimed in [DF00, FDF94b, BF98] that the corresponding results can be slightly improved by applying what is known as *spectral sharpening*. This consists in simulating an alternative sensor functions as if they were a set of nonoverlapping narrow-band sensors.

Color constancy algorithms can also be classified by the total number of recovered parameters. Most algorithms try recovering two or three parameters, accordingly to the kind of color coordinates used, namely, *chromaticities* (normalized colors) or raw RGB, respectively. When considering the estimation of the camera response, it is natural to try to appraise the RGB color of the illumination¹. Nevertheless, we are usually more interested in the illuminant chromaticity than in its magnitude, discarding the intensity information. Others even endeavour the recovery of its spectral power distribution, expressed in some function basis.

The number of color constancy algorithms that exclusively work on a chromaticity space is relatively high [Fin95b, Fin96, FH97, FSC98, CFB98, FHH99, FCB99, FHH01]. The fact of ignoring the intensity of the illuminant and its effect on the solutions of these algorithms were extensively studied in [FHH97, FH98b, FH99, FH00]. The main conclusion is that there is no further advantage in using RGB algorithms if only the illuminant chromaticity is needed, since algorithms based on the chromaticity give analogous results, being at the same time more robust to shade and highlights. In such a case, an estimate of the chromaticity can not be turned into a complete RGB estimation unless an independent evaluation for the illuminant intensity is produced.

2.2 Outline of the Chapter

This chapter is divided into five sections. First, the color formation process is referred to in Section 2.3. Second, in Section 2.4 the hypothesis about the composition of lights and surfaces encountered in real scenes are reviewed. Third, in Section 2.5 the most important approaches on color constancy are surveyed. These methods are grouped into the following categories, namely, gray world methods, Retinex methods, gamut-mapping methods, Bayesian and correlation methods, neural network methods, and linear model methods. In Section 2.6 color constancy in the special context of object recognition and image indexing is examined. At the end, some conclusions are briefly summarized in Section 2.7.

2.3 Color Formation Model

Basically, there are two main phenomena occurring in the formation of color images, namely, light reflection on the object's surfaces and camera measurement of the light coming out from this reflection. For the first one, it is necessary

¹The color of an illumination is the response of the imaging device to an achromatic surface viewed under this light.

to describe the mutual interaction between light and surfaces as it is seen from a point in the image plane. Accounting for the second issue, the way a sensor integrates the light falling onto the image plane must be established. These two issues are concisely dealt in this Section.

2.3.1 Reflection Model

The alteration followed by a light beam from its birth to its fall into the camera sensor can be described as a set of successive reflections on the surfaces of the objects in a scene. This way, light changes its wavelength composition as it touches different surfaces. To explain this apparently chaotic process, it is usually defined a reflection model tying the light arriving onto and leaving from an infinitesimal surface element. This information is encompassed by the surface spectral reflectance function.

The diversity of reflection models found in literature is huge and their full review is out of the scope of our work. However, let us cite those of Lambert (1760), Torrance and Sparrow [TS67], Wolff [Wol94], Phong [Pho75], the Bidirectional Reflectance Distribution Function (BRDF), and the dichromatic model [Tom91] just as the most widespread in both computer vision and computer graphics.

Basically, all of them model the surface spectral reflectance function as a linear combination of different approximating functions each one describing a specific physical hint of the reflection phenomenon. In order to avoid being computationally expensive or excessively complex, most of the above models only take into account two aspects of the reflection phenomenon, namely, the *diffuse* reflection and the *specular* reflection.

The first one appears when a surface reflects the same proportion of incident light in all directions. This reflection component is known as the *Lambertian* component and is the basic constant component for all the surface reflectance models. The specular reflection shows up when the incident light is mainly reflected in a particular direction. In this case, the proportion of reflected light depends upon both the incident and the viewing directions and upon the physical structure of the surface (texture). This component causes highlights and the glossy aspect of objects.

If a Lambertian model is assumed for the surface reflectance, then the *image irradiance* I defined as the total light focused by the camera optical system onto the image plane, into a surface element forming an image pixel, will be the quantity $I(\lambda) = R(\lambda) E(\lambda)$, where $R(\lambda)$ is the *spectral reflectance function* of the piece of surface that optically corresponds to the pixel and $E(\lambda)$ is the *spectral power distribution* of the light beam falling into that pixel. Both I , E , and R are all *nonnegative* functions of the wavelength λ , but they depend neither on the incident nor on the viewing directions. Besides, the geometry of the object surface is merely a *scale factor* in function E .

2.3.2 Sensor Model

A general camera can be seen as an array of p sensors $S_h(\lambda)$, where $h = 1, \dots, p$ and usually $p = 3$ in color cameras. Each of these sensors measures the light $I(\lambda)$ arriving onto the image plane and giving rise to the h^{th} channel value y_h for that pixel in the image [VFTB97a, VFTB97b, VFTB97c, KAPH94, ST97].

This is generally modeled by the following expression

$$y_h = \int_{\lambda_0}^{\lambda_1} S_h(\lambda) I(\lambda) d\lambda, \quad h = 1, \dots, p \quad (2.1)$$

where $[\lambda_0, \lambda_1]$ is the interval where these sensors operates (400 ÷ 700 nm).

To go further into color research we advise first removing all the sensor nonlinearities that may appear. Vora et al. [VFTB97a, VFTB97b, VFTB97c] and Vrhel and Trussell [VT93] showed some feasible methods to calibrate color devices. Barnard [BF99] also suggested calibration as a first step in any color research. Nevertheless, former calibrations are difficult to carry out because the experimental set needed is expensive and pretty restrictive. Alternatively, there also exists a kind of calibration called *radiometric* [DM97, MN99] where the sensor response can be estimated varying the *exposure* under which the picture is taken. This observation permits estimating of a *linearized* response without any prior knowledge of the scene radiance and no special device required.

2.4 Hypotheses about the Physical World

A great extent of the research on color constancy assumes that the world consists in a collection of perfectly diffuse reflecting surfaces. An example of that can be found also in human color constancy research, where the norm is the use of solid matte color charts. In Retinex [LM71, LM77] a set of random collages of color charts, named Mondrian's² world, were used [FD88]. Color constancy has also tried to take advantage of specularities present in the image [Sha85, TW89, TW90], rather it is a pretty unusual approach to color constancy because the difficulty in identifying such regions as well as that of collecting the inevitable shape information.

Nevertheless, it is quite infrequent to handle a set of test images without any kind of specular reflection in it. Therefore, the Mondrian's world including specularities from *inhomogeneous dielectric materials*, such as plastic and paint, is a habitual context to evaluate color constancy algorithms despite its limited stretch. Besides, illumination is considered spatially uniform as a rule, both in chromaticity and intensity, although there exist some works which take into account the case of a smoothly varying light [Hor74, BFF96, BFF97].

Most of the color constancy algorithms take into account the diversity, and sometimes the distribution too, of surfaces and illuminants that can be easily found in the real world. As explained in the previous Section, information about surfaces and illuminants are generally provided from a collection of *reflectance functions* and *spectral power distributions*, respectively. The set of required data corresponding to the color of surfaces under these illuminants is computed using an appropriate camera model, e.g., the one in Section 2.3 [VFTB97a, VFTB97b, VFTB97c]. Nevertheless, it is neither always easy nor practical to collect all that indispensable data.

Empirical studies showed that both real illumination and reflectance are relatively smooth functions of the wavelength of light in the visual spectrum

²Pieter Cornelis Mondriaan (1872-1944): Dutch painter who developed *neoplastic* aesthetic involving reduction of paintings to elements of straight lines, primary colors, and noncolors.

(400 ÷ 700 nm). Hence, finite-dimensional linear models were introduced to describe spectral power distributions by Buchsbaum [Buc80], who extended previous works by Helson [Hel38] and Judd [Jud40, JMW64]. Similarly, reflectance functions of a great variety of materials have also been studied. Cohen [Coh64], Parkkinen et al. [PHJ86], Maloney [Mal86] studied the reflectance properties of the Munsell chips, whereas Vrhel et al. [VGI94], Krinov [Kri47], Wyszecki and Stiles [WS82] considered the reflectance of naturally occurring surface materials.

Data corresponding to all those sets was collected in an outstanding effort by Barnard et al. [BMFC02], furnishing a very interesting database of 1995 spectral reflectances for the color constancy research. This work also encompasses the spectral distributions of several lights that were picked for their ability to generate a range of chromaticities close both to natural and to artificial lights that may easily be found anywhere. Another similar data set can be found in [PHJ86] too.

In respect to daylight, research has established its spectral composition which can be reduced to the combination of a small set of estimated spectra [JMW64]. From these data, it is even possible to compute the daylight chromaticities and generate the geometric locus for all possible chromaticities as a function of the color temperature, which is a curve close to the one radiated by the *black body* at the same temperatures [BDRH95]. In [TEW01], e.g., the set of light chromaticities is recovered from the most highlighted regions from a set of evaluation images.

2.5 Color Constancy Methods

As said, the goal when applying color constancy to an image is to decrease the effect of a likely varying illumination in order to obtain color data describing more precisely the physical content in the image. Commonly, color constancy is characterized by estimating the color of the illuminant or a mapping relating it to the color of a canonic light. Once such information is computed, it can be directly employed to render an image of the scene as it would be seen under a reference light.

Many algorithms have been developed to recover such descriptors, of which we next describe the most principal ones. As it is an ill-conditioned problem, some additional assumptions about the physical world must be furnished in order to attain useful results. Algorithms can be classified depending on the aforementioned hypotheses and the considerations about the occasion where to be applied. An important axis of this classification is the degree of complexity of the illumination, being the most distinctive point the uniformity of the illumination.

Another axis is whether or not an algorithm is robust to specular reflections. Some need the presence of specular reflection, while others are neutral and some are severely degraded by the presence of such reflections. Most approaches suppose that illumination is uniform and no specularity exists, i.e., conditions of the Mondrian's world are assumed. Finally, other schemes are only concerned about the recovery of the illuminant chromaticity, avoiding the problem of determining its intensity.

2.5.1 Gray World Methods

Possibly, the simplest approach to color constancy is to compute one statistic from the whole scene to be used afterwards as an estimate of the illumination, which is assumed uniform for the entire interest region. A straightforward candidate for these statistics is the *mean* carrying the so called *gray world* hypothesis. In physical terms, such a hypothesis implies that the mean among all reflectances of the scene is referred to be *gray*³. Despite its simplicity, there exists a great number of variations.

A possible one dwells in the way gray color is specified. Alternatives include the wholly specification of the spectral power distribution, the set of their components on a function basis or just the RGB values under a known illuminant. Another notable choice concerns the selection of the reference gray color. Given a method to specify the gray, the best choice would be the real value of it in the world. Nevertheless, this is not always the case unless the data are synthetically generated.

An approximation consists in taking a real gray as the gray, i.e., a constant spectral reflectance of one half of the maximum spectral value or the corresponding RGB response as if the reflectance was uniformly that value, provided the diagonal model for the illumination change. Using this model, the algorithm normalizes the image using the relation between the gray RGB values under a canonic illumination and the mean RGB among image pixels. A related method computes the average spectral reflectance from a database of surfaces to recover the gray reflectance instead of using a uniform reflectance.

The work in [Buc80] uses this assumption to estimate a quantity analogous to the illumination matrix⁴. However, as it is stated in [GJT88], the weakest point of that method dwells in an *ad hoc* choice for the basis to express the reflectance functions, as well as the election of the gray, whose coefficients must be equal within that basis. The method is improved in [GJT88] by calculating the basis from a database of real reflectances, employing the average among the reflectance basis as the gray. The result is an estimation of the reflectance coefficients of the scene surfaces in respect to the chosen basis.

Additionally, in [GJT88] they realize that in order to get a more exact correspondence between the model and the world, a segmentation of the image is necessary to be able to compute mean values among surfaces instead of among pixels. In this model, two surfaces should have the same weight despite their different size. Segmentation confidence might seem a problem since segmentation of real images is a difficult process. Nevertheless, [Bar98] upholds the fact that this algorithm degrades slowly in front of inaccurate segmentations.

2.5.2 Retinex Methods

A capital contribution to color constancy is the historic Retinex theory by Land [LM71, LM77], which is analyzed and developed in a number of works [Hor74, Bla85, FDB92, McC97]. Their original goal is to have a model of computation for human vision, rather it has also been used and extended to include machine vision. Theoretically, most Retinex versions are robust to spatially varying

³All color channels are equal.

⁴A matrix where any element corresponds to the response of a sensor to one of the basis functions describing the set of surface reflectances [Bar98].

illuminations in a smooth manner, although the evaluation on real images has always been limited to scenes where illumination was controlled to be pretty uniform.

In Retinex-based methods, light variation is removed assuming that small spatial variations on the response are only due to illuminant variation, while big variations belong to changes in the surfaces. The goal of Retinex is to estimate the surface color by comparing the quantity of energy received for each pixel applying a statistic (maximum) computed on a certain region around a pixel. The relations between these quantities (or their logarithms) are used as interest descriptors in such a way that the method is implicitly assuming a diagonal model. Details of the methods varies depending on the version of Retinex taken into account.

In [LM77], the method follows random paths departing from the pixels of interest. While each path is tracked down, the relation between the channel responses in adjacent pixels is computed. If the ratio is close enough to one, it is assumed that the differences are produced by noise or illumination change and the ratio is accordingly changed into one. On the other hand, if the ratio is dissimilar enough from one, nothing happens. Ratios are combined to determine the ratio between responses in the interest pixels and the highest responses found along the path. Finally, the average among the results found in all the paths is calculated. To ease these computations, logarithms can be used instead since, this way, the identification of the color change is undergone by subtraction, whereas the recovery of luminosity is done by addition [Hor74, Bla85, FDB92].

Other strategies [Lan83, Lan86] exploit the differences between logarithms, thresholding the output to remove the effect of light variation along the random paths. In [Lan86] the estimate was simplified even further applying the logarithm of the ratio between the response of a given pixel and a weighted mean of the responses in a neighboring region around that pixel. If a uniform illumination is assumed, the first aforementioned Retinex version consists in scaling each channel by the maximum value found in the image. In a similar way, the latter method converges to a normalization using a geometric mean [BW86] and thus it is essentially a gray world algorithm. Therefore, Retinex can be simplified and implemented in a more powerful manner whenever a uniform illumination is assumed.

2.5.3 Gamut-Mapping Methods

The strategy of computing mappings between color gamuts was successfully introduced by Forsyth in [For90] and has also been the base for the extensive work of Finlayson [FDF94a, Fin95a, Fin95b, Fin96, FH98b, FH00] and others until the present day. The idea consists in explicitly restricting the set of applications between images of scenes taken under an unknown illumination and the canonic light. Despite the analysis in [For90] includes at the same time diagonal mappings as well as more general linear ones, the most successful algorithms and its posterior versions are all of them based on the diagonal model of color change.

The main source of constraints on the set of maps comes from the observed responses of sensors, i.e., the colors of pixels in the image. The set of all possible sensor responses due to all expected surface reflectances as they are observed under the known canonic light generates a convex set named *canonic color*

gamut. Assuming a diagonal model of color change, two color gamuts are only one diagonal mapping far from each other.

Despite the canonic gamut is known a priori, it is necessary to use colors in the input image as an estimate for the set of possible mappings since the real illuminant is unknown. This is known as the *image gamut*. Each of these mappings is a *feasible* solution and the main achievement of this algorithm is to propose a way to compute all these solutions. In a second stage of the algorithm, a sole solution is selected from the set of feasible mappings.

Since color gamuts are convex sets, they can be further reduced to their contours by a convex hull instead. This way, any color set is described as a list of vertexes representing the outer limits of the color gamut. In order to compute the set of feasible mappings transforming colors from a gamut under an unknown illuminant onto those of the canonic gamut, it is necessary to use the image gamut and its convex hull. Any mapping transforming the image convex hull onto a subset inside the canonic gamut also transforms any point inside that hull to a corresponding point inside the canonic gamut. Therefore, it is sufficient to consider mappings between colors on both convex hulls, namely, the image and the canonic convex hulls.

Thanks to the diagonal model, these mappings can be assumed to be points in a space of mappings where their convex hulls are extracted hence. This is a higher level of abstraction since geometric features are applied not only to color sets but also to mapping sets. This way, for each color in the measured convex hull it is possible to generate a convex hull of mappings by applying the diagonal model that exists between this color and all the colors in the canonic convex hull. As a logic consequence, the *correct* mapping should fall inside the room defined by the intersection of all those sets, since it must be able to take, at the same time, any of the colors in the image convex hull inside the canonic convex hull.

The works in [Fin95a, Fin95b, Fin96] apply the above procedure to a chromaticity space named *perspective color*, after proving that convexity in the 3D gamut constraint is preserved by the perspective projection. Gamut convexity is also required in 2D chromaticity space if the gamut constraint is to be efficiently exploited. Advantages of employing a chromaticity space are the obtaining of algorithms more robust to intensity variations such as those produced by shadings and specularities, as well as the reduction of computational complexity. Moreover, in [FH97, FH98b, FH99, FH00] it is verified that the set of feasible mappings computed in a chromaticity space is the same as the one computed in RGB coordinates and projected afterwards to perspective coordinates, meaning that no further advantage exists in using RGB coordinates if only the chromaticity of illuminant is to be found. Other papers using perspective coordinates are [FHH97, FHH99, FHH01].

Finlayson also introduces another helpful constraint related to the set of feasible illuminants. Likely enough, not all the theoretically possible illuminants will be encounter in reality. Therefore, the set of feasible mappings must be restricted by means of an *illuminant gamut*, that is, the convex hull of the set of expected light chromaticities. Unfortunately, this set is not convex, which hinders the intersection between other convex hulls. Nevertheless, Finlayson is able to apply these restrictions in the 2D case. In [Bar95, Bar99a] the convex hull of the nonconvex set is considered as a good approximation of a wider set of real illuminants.

Usually, the set of feasible diagonal mappings is still big enough so that the choice of a proper mapping represents an important stage of the algorithm. In [For90] the mapping maximizing the volume transformed onto the canonic gamut is selected. In [Fin95a] the same heuristic is maintained, but rather this time in the chromaticity space. In [Bar95] the *centroid* of the set of feasible mappings is used to compute the solution in both RGB and chromaticity spaces.

Choosing the centroid is optimum whenever feasible solutions are uniformly distributed. Nevertheless, in a chromaticity space this may not be a good heuristic since the projection might distort mapping uniformity. This way, in [FH97, FH98b, FH99, FH00, Bar00] there is an alternative procedure which computes the set of constraints in 2D while the mean of feasible mappings is calculated in the 3D space. This method is based on the earlier mentioned fact that the set of feasible mappings computed in 3D, after being projected onto a 2D space of chromaticities, is the same as the one obtained if this set would have been computed directly in the 2D space.

2.5.4 Bayesian and Correlation Methods

In [BF97] the Bayesian rule is applied to the problem of color constancy. This strategy assumes the existence of a certain knowledge about the prior distributions of lights and surfaces in the world. For each illuminant and combination of surfaces a set of possible observations arises. Given a set of colors, the Bayes's rule is used to count the posterior distribution for the illuminants and surfaces in the scene. The Bayes's law relates the probability for some parameters to describe the illuminant and the reflectances in the scene when certain colors are observed to the probability of these observations given those parameters.

Once this probability is computed from a set of observations, a likely value for the illuminant or the surface parameters is selected based on the value of that probability. There are two widely applied methods for achieving a single best estimate from a posterior distribution, namely, the *maximum a posteriori* and the *minimum mean squared error estimation*. Nevertheless, either a maximum or a minimum is sought, it may be necessary to fix upon one among some alternative extremes. A usual method to solve this problem is by means of a *cost* (loss) function as it is dealt in [BF97], where a new function accounting for the local loss of mass is claimed to better fit color constancy applications.

The method described in [BF97] presents some difficulties like the number of parameters being a function of the quantity of surfaces, which stands for a computationally expensive method. Computations of probability assume that some illuminants and surfaces are independent, which implies the corresponding image must be correctly segmented. If image pixels are directly used instead of segments, surfaces are not independent since neighbors tend to carry similar colors. Finally, PDFs for the real world are not fully known and, more than likely, there exist big discrepancies between simulations and real applications. For example, the authors in [BF97] only evaluate the algorithm using synthetic images fulfilling the proposed model, obtaining pretty good results as a consequence.

Some of these difficulties are also mentioned in the color-by-correlation approach [FHH97, FHH99, FHH01], although the probability density function must be known *a priori*. This algorithm is a direct application of the Bayes's concept. Besides, what is important is its freedom with respect to the complexity of the estimation of the surface parameters. In [FHH97, FHH99, FHH01] the

probability of observing a particular chromaticity is computed for each of a set of expected illuminants. Then, the probability matrix is employed along with the Bayes's framework to estimate the probability for each potential illuminant to be the real illuminant. Finally, the likeliest illuminant is chosen from this group by applying a cost function.

In addition, another aim of the above method is to comprehensively cope with most important color constancy methods by only interchanging the probability matrix accordingly. Therefore, it is proven in [FHH97, FHH99, FHH01] that color-by-correlation can accomplish other color constancy algorithms such as those of gray-world, gamut-mapping, and neural networks. As far as our knowledge reaches, this is the color constancy method achieving best results, for sure, because the true illuminant is selected from a discrete set of feasible lights.

2.5.5 Neural Network Methods

Recently, a series of some good results have been obtained using neural networks to recover the illuminant chromaticity [CFB98, Fun3, FCB99]. In these works a neural network is trained with synthetic images randomly generated from a database of illuminants and surfaces. Scenes likewise created can also include synthetic specularities. Furthermore, the chromaticity space is divided into discrete cells which represent the presence or absence of a certain color in the image. The neural network is afterwards fed with this binarized version of a chromaticity histogram. During the learning phase, inputs corresponding to the scenes synthetically generated are shown to the network along with the correct outputs. The learning phase is fulfilled using the *backpropagation* learning routine.

A further extension of the aforementioned works in [CF99, CFB99], carried out by the same authors, are basically centered on solving the color constancy problem whenever the origin of the images is unknown and no estimate of the illuminant is available. This is the case of images found in the Internet where a human observer is able to detect at once any color misadjustment. This is a general color constancy problem where neural networks show a good performance. What is more, in [CF99] these methods were even combined with other color constancy approaches to improve the results by means of a *voting* scheme.

Another type of network employed in [AHH96] are the Kohonen's *Self-Organizing Maps* (SOM) applied to evaluate the transformations between color gamuts. The main idea is to embed suitable 3D coordinate systems for each gamut by self-organization of a Kohonen's network and to be able to find corresponding points in both gamuts. The difference between the localization of these corresponding neurons in each SOM is then an approximation to the color change created by the difference between a canonic illuminant and the unknown one. Observed changes give rise to a look-up table in the color space which is posteriorly applied to correct colors in the images under the unknown illuminant.

2.5.6 Multiple Views Methods

It is quite common that the illumination of the scene spatially varies due to the interaction between different light sources, e.g., a white ball on a lawn under the

sun light receives a yellowish light direct from the sun and a more bluish light from the sky, while a greenish light is reflected from the grass. The problem of recovering the illumination can be partially solved if the surface under the varying illumination is extracted or, as it is done in [TO90], if two pairs of objects are identified under two different illuminations. However, despite the fact that varying illuminations are pretty common, there are still few methods capable of taking advantage of shape information to cope with the issue of a varying light.

As previously mentioned, the Retinex approach removes smooth spatial variations of light, without exploiting any specific shape information. In [BFF96] an algorithm is provided which is analogous to those in [For90, Fin96] but for the chromaticity case. Assuming that illuminant chromaticities are constrained, the authors show the chromaticity variation rate can be used to restrict the real illuminant chromaticity. In [Bar95, BFF97] a wider set of illuminants is employed. Here, the algorithm has also been modified in order to be used along with the gamut-mapping approaches developed for a uniform illumination in [For90, Fin96]. The idea is that once the varying illumination has been identified, the image can be rendered as if the scene had a uniform illumination, getting the constraints for the feasible illuminants from the surfaces in the scene as a consequence.

In [Bar95, BFF97] it is also introduced a method to identify a variable illumination in case this variation is slow. This procedure is based on the assumptions already used in the Retinex approach and only an approximate segmentation is needed. Despite too many spurious regions might degrade the performance of the algorithm, the precision of the segmentation process is not too significant in the final output of this approach. Anyway, once a segmentation is obtained, it is straightforward to determine the changing illumination within a region and then a method is proposed to combine in a robust manner all those lights into one single estimate for the whole illumination.

2.5.7 Linear Models Methods

A collection of color constancy methods slightly different to those reviewed so far are the algorithms based on the *physical* description of surface reflectances, light spectra, and sensors. A frequent way to describe these functions is by means of *finite-dimensional linear models* [MW92]. As insinuated in Section 2.3 while reviewing reflectance functions and the spectral power distributions, the strongest point of this approach is the general mathematical methodology where any function is given as a linear combination of functions from a suitable set forming a function basis. This kind of representation is very helpful since the information required to portray a function gets drastically reduced, spaces where these function span are this way easily handled and linear methods can be applied too.

A special and elegant method to compute color descriptors based on linear models is proposed in [Mal85, MW86, Wan86, Wan87]. Assuming illuminants are vectors in a p -dimensional space and surfaces have $(p-1)$ dimensions, where p is the number of sensors, any sensor response under a stable and unknown light belongs to a $(p-1)$ hyperplane at the origin. The orientation of this hyperplane depends on the illumination. Unfortunately, most of the times sensors only have 3 channels, being this approach clearly not good enough [FFB95], which

is not surprising at all since surfaces usually have a dimension higher than 2 and illumination easily surpasses 3 whenever a fluorescent light is used. An extension of this method when the same scene is captured under diverse lights can be found in [DI93].

Nevertheless, the description by virtue of linear models is very general and also includes a kind of algorithms named *spectral recovery*, which are capable of estimating at the same time functions corresponding to surfaces and lights. This is not precisely the color constancy concern and, as explained in [Fin98], there exists a substantial difference in recovering the whole light spectrum or just the illuminant chromaticity. The first one allows the computation of the second, rather the reverse path is under-constrained due to the dissimilar dimension. Color constancy is more connected to the study of surficial colors rather than the estimation of surface reflectances. However, a way to recover both reflectances and light spectra would greatly ease the solution to the color constancy problem.

Another interesting example of that kind is the spectral recovery method proposed in [HFD90] where the color signal is taken as an input that gives an estimate of the surface reflectance as the output. Light power distribution and surface reflectance function are both approximated minimizing the error of the generated color. Minimization is alternatively done at each step, first for the light and then for the reflectance. This approach has the advantage that from only a single measure an estimation of the reflectance is recovered, bringing color constancy independently for each pixel. In order to recover a whole light spectrum, a surprising method is proposed in [FH88, Ho88, HF88] based on the chromatic aberration of lenses in such a manner that, as the title of one of the previous papers states, light color can be obtained *from black and white*.

More recently, [CHY92b, CHY92a, CH95] have tried to improve the previous separation approach through a nonlinear least squares scheme. Nevertheless, very little progress has been made in this field up to the present, apart from some small improvements, as it is noticed in [Fin98]. Although the estimated reflectances are usually poorly accurate due to problems of numerical stability related to the sort of equations which are tried to be solved, linear models are still a good option to attack color constancy, at least theoretically. However, it is not clear whether the problem of separation is fully solvable and any progress here would have direct implications in the color constancy research.

In addition to the lot of approaches mentioned so far, there are other methods [TO00, FC84] not directly related to those of signal separation, but still based on linear models which is worth to be considered. Their goal is to recover the illuminant spectrum in order to remove its influence. Here, the principle of the *spectral maximum* is applied. This is pretty analogous to the idea of recovering the color of a light from regions with specular reflections. On the other hand, the work in [TO00] uses function basis from both reflectance and illumination spaces to compute a transformation between the colors observed under dissimilar illumination conditions as if a change of basis were carried out.

Lately, the approach in [Sap98, Sap99] tries to recover the parameters defining an illuminant using a generalized Hough transform. First, the color formation process is modeled by means of a *bilinear* scheme where it is possible to compute the observed color of a surface by a linear combination of matrices encompassing both light power distributions and surface reflectances. Then, this process is inverted to get an illuminant estimate from a color in the image. Since this is an under-determined mapping, a voting procedure is followed to discern

which is the likeliest illuminant to be the scene illuminant. It is also claimed that any available statistics about the distributions of colors as a function of the different illuminations can be provided to the method in order to attain better approximations.

2.6 Object Recognition and Image Indexing Methods

Two of the most remarkable tasks where color constancy is claimed to be helpful are *object recognition* and *image indexing* [FBM98]. While the goal in object recognition is to determine what is pictured in an image, the latter consists in finding a set of images in a database similar to the one given as a query. As widely known, these two problems are pretty sensitive to illumination changes since their performance greatly depends on the appearance. Therefore, algorithm efficiency should sensibly improve if those effects are removed. Despite any of the existing generic color constancy schemes could have been applied in most of the two kinds of tasks previously considered, there exist more specific approaches which take advantages of already known information about the objects in a given database.

In [MMK97] every object in the database is modeled depending on the expected illumination range. Modeling of known objects under a variety of expected light conditions is a strategy that has also been employed in [BD98]. Besides, it has been applied in the context of human faces pictured as gray level images in a number of works [BHK97, GKB98, BK98, GBK99, GBK00, GBK01]. Thus, in [MMK97] each surface belonging to an object is represented as a convex set of all possible chromaticities under the range of all possible illuminations. The occurrence of a particular chromaticity counts as a vote for the presence of the object. The authors in [MMK95] integrate the information about the adjacency given by the contour inside the object recognition scheme and the ratio between reflectance and intensity, defined in [NIK91, NB93, NB96], is exploited as a quantity invariant to the illumination for each color channel. The invariant taken into account is based upon the assumption that illumination is usually constant when crossing a region border, if it is coarsely taken. Whenever the diagonal model is taken into account, ratios between RGB are constant through the union of pairs of surfaces.

Image indexing is less restrictive than object recognition since generally the problem of background–foreground segmentation is obliterated as well as that of finding the exact identity. Indexing techniques can be after all employed to recognize objects and localize objects within an image by an exhaustive comparison of image regions. All those problems need the indexing to be fast and robust in respect to inclusions of the background during the computations, as well as to the scale and the object position within the image. However, the seminal work of [SB91] proposes a strategy for the recognition of objects which avoids most of the previously mentioned difficulties. This method compares images through their color histogram. As color histograms are greatly dependent on illumination, the posterior work in [FF95] tries to improve the precedent method putting forward an illumination–invariant approach based on comparing histograms of ratios of RGB channels through the borders of regions.

Another version in [CFF95, FCF96] exploits the angle formed by two colors, taken as if they were vectors, as an index to compare images. Indeed, the author [Fin00] goes a bit further and formulates a procedure to achieve one dimensional image from a RGB image where only one invariant value computed as a function of its color coordinates is assigned to each pixel. Similar techniques are those of [MD97, DWL98, DWL99] which are grounded on a sort of color normalization step, followed by the reduction of dimensionality by projecting colors onto a chromaticity space. Treating the chromaticity histograms as images, it is performed an effective low-pass filtering of the histogram by first reducing its resolution via a wavelet-based compression. This way, color constancy is carried out also in the compressed domain.

For a more complete review on these methods we suggest looking up into the noteworthy works in [BFM00, BFC00, BFMC00, BCF02, BMCF02], which consist of a compilation of color constancy and evaluating approaches within the context of image indexing. A previous alike review on this field can be found in [FSC98], where the question arisen in [FBM98] about whether color constancy is good enough to accomplish indexing and object recognition is put into play. At that time, the answer in [FBM98] was negative, despite the fact that posterior works such as [GS96, GS99, GBSG01], where different techniques to obtain kind of a color normalization by means of invariants, and those reported in [BFM00, BFMC00] have shown certain amount of improvement. Nevertheless, it seems now rather clear that applying some color constancy algorithm is usually better than doing nothing to improve indexing and recognition efficiency.

2.7 Conclusions

A large number of approaches are reviewed in this Chapter, being summarized in Table 2.1. Their main concern is that of recovering an estimate for the scene illuminant which can help to remove the effect of it on color images. As can be inferred from the amount of papers, this is a quite difficult problem and yet no general method has been proposed that fully solves it. However, there exist many partial solutions which are worth to be considered. First, it is a must to choose among the best algorithms. This towering endeavour was faced by Barnard in the series of works in [Bar95, Bar98, Bar99b, BFC00, BFM00, BFC00, BFMC00, BCF02, BMCF02] that studies, among other things, the practical application of most of the available color constancy algorithms.

Barnard finds that the most effective methods whenever synthetic images were used are those based on image statistics, such as color-by-correlation and neural networks [BFC00, BCF02]. Some variants of the Forsyth's 3D method of gamut-mapping also works pretty well with the advantage that these algorithms also estimate the illuminant intensity. Moreover, similar results are found in [BFMC00, BMCF02] for real images to those in the case of synthetic images. However, the results of algorithms based on statistics are slightly worse than expected, which highlights the difficulty of finding proper statistic models that be useful for real images.

Concerning the job of object recognition and image indexing by using the histogram comparison scheme in [SB91], the set of results in [BFMC00, BMCF02] are similar to those mentioned above, which indicates that the 3D gamut-mapping approaches outperform others. It was expected that some color con-

Color Constancy Techniques	Advantages	Disadvantages
Gray-World Methods	<ul style="list-style-type: none"> • They are simple • Many versions to compute the gray color 	<ul style="list-style-type: none"> • Usually they are too simple • Gray-world hypothesis does not always hold
Retinex Methods	<ul style="list-style-type: none"> • Close to human models of color constancy • Theoretically robust to spatially varying illuminations 	<ul style="list-style-type: none"> • Limited success on real images
Gamut-Mapping Methods	<ul style="list-style-type: none"> • Quite straightforward computation • Good performance 	<ul style="list-style-type: none"> • Difficulty to generalize to mappings other than the diagonal ones • Gathering reflectance and illumination may not be easy to carry out
Bayesian and Correlation Methods	<ul style="list-style-type: none"> • A sound way to choose among several solutions • Good performance 	<ul style="list-style-type: none"> • Lights and surfaces in real world are not fully known • Probability matrix is a discretization and may be too restrictive
Neural Network Methods	<ul style="list-style-type: none"> • They cope with unknown dynamics in a robust way • Pretty good results with synthetic images 	<ul style="list-style-type: none"> • Gathering training samples is not straightforward • Learning phase is slow
Multiple View Methods	<ul style="list-style-type: none"> • They account for spatially varying illuminations 	<ul style="list-style-type: none"> • Usually shape information is needed • Scarce number of applications to real scenes
Linear Models Methods	<ul style="list-style-type: none"> • Good theoretical foundations • Close to the physics of the color constancy issue 	<ul style="list-style-type: none"> • Difficulty in gathering all data needed in those models • Poor results reported with real images

Table 2.1: Color constancy techniques.

stancy process would have provided color descriptors that were precise enough so that indexing performance was close to that obtained in case no light change happened. Nevertheless, it was not the case and despite color constancy significantly improved the performance of such methods, the amendment was not enough to completely cope with the whole light variation.

Besides, the degree of improvement in the histogram comparison is almost linearly dependent on the prediction error for the light chromaticity, i.e., the better the illuminant is estimated, the more similar the histograms of images are. Therefore, this result makes perfectly clear that in order to attain any good result whenever trying to identify objects via their appearance or simply indexing images by their content it will be necessary to first estimate somehow the color of the light under which the image has been captured to render the image as it would be seen under a canonic illuminant.

Very recently, in [HF04] a reevaluation of previous experimental data for five different color constancy algorithms is presented. The most important point raised by this work is that the relative performance of such algorithms changes considerably depending on the criteria by which they are judged. In particular, if the Wilcoxon's Sign Test⁵ is used the conclusion is that gamut-mapping and color-by-correlation are statistically equivalent and perform significantly better than the three other algorithms tested (Max-RGB and two versions of gray-world).

Finally, we must reiterate that, as said in [Fin98], the problem of color constancy has conciliated lots of efforts from many researchers but not until recently a few of effective algorithms have been achieved being capable of working with true real images. The two basic omnipresent ideas to retain in mind are that, first, there exists a limited range for the colors observed by a camera which depends on the scene illumination and, second, that the feasible illuminations may only vary within some limits. This latter constraint is helpful in scenes where multiple incident lights exist. Nevertheless, most color constancy approaches will not work well unless the color diversity in the scene is wide enough, since it is statistically difficult to grasp the color of the scene light if objects apparently present very few variation among them.

⁵The Wilcoxon's Test is used to test the null hypothesis that the random variables A and B are such $Pr(A > B) = 0.5$, i.e., that the algorithms A and B have the same performance.