Optimum sensors for color constancy in scenes illuminated by daylight

Sivalogeswaran Ratnasingam,[1,*] Steve Collins,[1] and Javier Hernández-Andrés[2]

[1]*Department of Engineering Science, University of Oxford, OX1 3PJ, Oxford, UK*
[2]*Department of Optics, Sciences Faculty, University of Granada, Granada 18071, Spain*
[*]*Corresponding author: siva@robots.ox.ac.uk*

Received February 10, 2010; revised August 11, 2010; accepted August 17, 2010;
posted August 17, 2010 (Doc. ID 123908); published September 20, 2010

The apparent color of an object within a scene depends on the spectrum of the light illuminating the object. However, recording an object's color independent of the illuminant spectrum is important in many machine vision applications. In this paper the performance of a blackbody-model-based color constancy algorithm that requires four sensors with different spectral responses is investigated under daylight illumination. In this investigation sensor noise was modeled as Gaussian noise, and the responses were quantized using different numbers of bits. A projection-based algorithm whose output is invariant to illuminant is investigated to improve the results that are obtained. The performance of both of these algorithms is then improved by optimizing the spectral sensitivities of the four sensors using freely available CIE standard daylight spectra and a set of lightness-normalized Munsell reflectance data. With the optimized sensors the performance of both algorithms is shown to be comparable to the human visual system. However, results obtained with measured daylight spectra show that the standard daylights may not be sufficiently representative of measured daylight for optimization with the standard daylight to lead to a reliable set of optimum sensor characteristics.

OCIS codes: 330.0330, 330.1690, 330.1720, 330.1730.

1. INTRODUCTION

In naturally illuminated scenes direct sunlight and shadow can create a scene with a wide dynamic range that can then lead to saturation and underexposure of parts of a scene. These large variations in intensity together with changes in the spectral power distribution of daylight can also cause unwanted variations in the apparent color of the surfaces in a scene. It is these variations in the recorded colors of surfaces that make it difficult to use color as a reliable source of information when creating machine vision systems. In contrast reliable chromaticity information can be obtained from the sunlight or skylight illuminated scenes using an algorithm based on the blackbody model of the spectrum of the illuminant [1]. Marchant and Onyango [2] proposed an algorithm for solving color constancy under daylight by taking ratios of sensor responses. This algorithm is based on the assumptions that the power spectrum of daylight can be approximated by the blackbody model, and that the spectral width of the image sensors is infinitely narrow. In addition, the large variation of intensity and power spectrum of illuminant in a daylight scene can be easily separated from the reflectance by taking the logarithm of the sensor responses. In solving color constancy the advantage of using logarithmic responses was proposed in the Retinex algorithm [3] and in Horn's algorithm [4]. Based on the blackbody assumption Finlayson and Hordley [5] proposed an algorithm based on the logarithm of three sensor responses to find a one-dimensional (1D) solution to the color constancy problem. However, it was shown that finding a 1D solution to the color constancy problem leads to confusion of perceptually different colors [6]. Finlayson and Drew [7] applied this 1D color constancy algorithm to four sensor responses to form an illuminant-independent two-dimensional (2D) space [7]. Based on the blackbody model Romero *et al.* [8] proposed an algorithm for color constancy in scenes illuminated by natural light. Recently, Ratnasingam and Collins [1] proposed a model-based algorithm that extracts two illuminant-independent features that represent the surface reflectance using data from four sensors. For mathematical convenience Ratnasingam and Collins [1] assumed that the sensors respond to only a single wavelength, and that the illuminant spectrum can be modeled by a blackbody spectrum. This model-based algorithm estimates the illuminant effect on one sensor response using the responses of two other sensors to create an illumination-independent feature [1]. However, sensors with such extremely narrow spectral responses are both difficult to manufacture and would require long exposure times. Ratnasingam and Collins [1] have shown that narrow spectral responses are not essential to the algorithm.

In this paper two methods of improving the quality of illuminant-independent reflectance descriptors (referred to as features) are investigated. In Section 2 the performance of the model-based algorithm proposed by Ratnasingam and Collins [1] is investigated for sensors with different spectral bandwidths and levels of both sensor noise and quantization noise. A projection-based approach to obtaining features that are independent of the illuminant spectrum is then investigated as a method of improving the quality of the features that can be obtained from the responses of four sensors with different spectral responses. These spectral responses are modeled using a

Gaussian function, and the response of each sensor is simulated by numerically integrating the image equation [1]. The resulting sensor responses are then used as the data from which illuminant-independent features are obtained, and the quality of these features is then assessed by using them to distinguish between perceptually similar colors. Another approach to obtaining better results is to optimize the spectral responses of the four different types of sensors for this particular application. In particular it may be possible to extend the developing interest in organic photodetectors integrated onto silicon substrates [9] to use families of different organic chromophores [10] to create cameras whose pixels have application-specific spectral responses. In Section 3 the effect of optimizing the wavelength at which the response of each sensor is maximal is investigated using gradient descent with particular sets of reflectances and daylight spectra. The performance of the optimized algorithms is investigated by changing reflectances and illuminants in Sections 4 and 5, respectively.

2. PERFORMANCE EVALUATION

A simple method of extracting chromaticity features from logarithmic sensors with four different spectral responses was described recently by Ratnasingam and Collins [1]. In the derivation of this method it was assumed that the spectral width of each sensor is infinitely narrow, and the illuminant spectrum can be approximated by a blackbody spectrum. If these assumptions are valid then it is possible to create two illuminant independent features (F_1 and F_2) from the responses of four sensors using the equations

$$F_1 = \log(R_2) - \{\alpha \log(R_1) + (1-\alpha)\log(R_3)\}, \quad (1)$$

$$F_2 = \log(R_3) - \{\gamma \log(R_2) + (1-\gamma)\log(R_4)\}, \quad (2)$$

where R_1, R_2, R_3, and R_4 are the sensor responses, and α and γ are two coefficients that will be referred to as channel coefficients. If λ_1, λ_2, λ_3, and λ_4 are the wavelengths at which the four sensors have their maximum responses then the variations in both the illuminant intensity and power spectrum can be removed if the two channel coefficients satisfy the following two equations [1]:

$$\frac{1}{\lambda_2} = \frac{\alpha}{\lambda_1} + \frac{1-\alpha}{\lambda_3}, \quad (3)$$

$$\frac{1}{\lambda_3} = \frac{\gamma}{\lambda_2} + \frac{1-\gamma}{\lambda_4}. \quad (4)$$

In the previous investigation of the feature space formed by F_1 and F_2 [1], the reflectance data used were those Munsell reflectances that correspond to colors taken from a thin plane in the CIELab color space, and the relative spread of each of the reflectance samples on the feature space was assessed using a Mahalanobis distance metric. The sensitivity functions of the sensors were modeled using Gaussian functions with a sensible choice of parameters to cover the entire visible spectrum evenly.

Previously Ratnasingam and Collins investigated the feature space formed by F_1 and F_2 using CIE standard daylight spectra with correlated color temperatures (CCTs) between 5000 K and 9000 K. However, measurements of actual daylight spectral power distributions [11,12] show that the CCT of measured daylight can occasionally fall outside this range. Therefore, in this study the algorithm's performance with the entire CCT range defined by the International Commission on Illumination (CIE) was used. In particular a set of spectra of CIE standard daylight was chosen with CCTs varying between 4000 K and 25000 K [13]. The particular CCTs used could be chosen so that they are spaced evenly along the mired scale (given by 10^6 K^{-1}) [14]. However our overall aim is to differentiate surfaces illuminated by daylight. The CCTs used have therefore been chosen to have a similar distribution of CCTs as the actual measured daylight [11,12]. This set of 20 daylight spectra is referred to as the CIE standard test daylights in the rest of this paper.

In the initial study of the model-based algorithm [1] the features were obtained from the responses of evenly spread Gaussian sensors in the visible spectrum (400 nm to 700 nm) with spectral peak positions 437.5 nm, 512.5 nm, 587.5 nm, and 662.5 nm (FWHM 80 nm). Figure 1 shows the illuminant-independent feature space formed with the features (F_1 and F_2) obtained from this sensor combination. In generating the sensor responses for this feature space 204 Munsell reflectances [15] with similar relative lightness were illuminated by the 20 spectra of CIE standard test daylights. In Fig. 1 the color of each reflectance is used to represent the points in the space that correspond to the surface. The figure shows that, in typical spaces such as the one shown in this figure, color in the space varies smoothly across the space. However, there is a small gap in the upper right hand corner of the feature space, and on one side of this gap dissimilar colors appear as neighbors. These features occur in all the feature spaces that we have observed and arise because of the metamer problem that occurs whenever the responses of a small number of detectors are used to distinguish between different reflectance spectra.

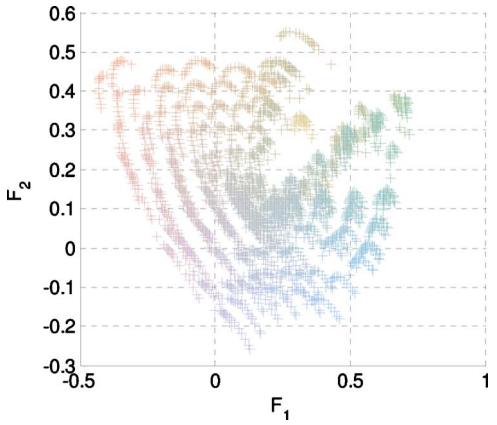

Fig. 1. (Color online) Chromaticity space formed by the model-based algorithm with unquantized responses of evenly spread Gaussian sensors of FWHM of 80 nm. In this space 204 Munsell samples are projected when illuminated by 20 spectra of CIE standard test daylights.

A closer inspection of the feature space shows that the residual dependence on the illuminant means that each of the Munsell reflectances creates a small cluster of responses in the feature space. The size of these clusters depends on several factors including the width of the sensor responses, the amount of noise in the sensor responses, and the difference between the spectrum of the light source and that of a blackbody. To determine the spectral bandwidth of the sensors needed to obtain useful features a method was proposed to determine the significance of the area occupied by each cluster [1].

The cluster of points formed by each of the Munsell test reflectances when illuminating with CIE standard test daylights form a non-uniform distribution of points (shown in Fig. 1). To account for this observed non-uniform distribution it is appropriate to characterize the size of each cluster of points using a distance metric that takes into account this non-uniform spread of points. Therefore, the Mahalanobis distance was applied to determine a boundary that ideally encloses all points in each cluster of responses corresponding to the same Munsell reflectance. For n-dimensional normally distributed data, the Mahalanobis distance between the center of the distribution C and a point in the distribution P is defined as

$$D_M^2 = (P - C)' \Sigma^{-1} (P - C), \tag{5}$$

where Σ is the covariance matrix of the distribution. For a pair of surface reflectances representing colors separated by a known distance in CIELab space the Mahalanobis distance can be used to determine a boundary around each cluster. To determine the Mahalanobis distance boundary of a particular reflectance pair the first step is finding the center of each cluster of responses. Then the Mahalanobis distance from the center of the respective clusters to the boundary was increased from a small value until the boundaries formed by both members of a pair touched each other. To assess the dependency of the feature space on the illuminant the number of responses that fell inside the respective boundary in the pair was then counted. This test was repeated on all the 100 pairs of reflectances in the test data set, and the percentage of points falling within the correct boundary was recorded.

The lightness component of a color and the brightness component of an illuminant are inseparable [5,16]. In Ratnasingam and Collins' model-based algorithm [1] removing the brightness component of the illuminant to deal with potential variations in brightness also removes the lightness component of a color. This is an advantage because when changing the viewing illuminant of a color surface the variation in luminance is large compared to its chrominance [17]. The model-based algorithm preserves only the chromaticity descriptors of a surface. In the previous paper [1] a thin plane of Munsell samples with CIELab L values between 47.8 and 50.2 was used in generating the test sets. However, in this paper, to more accurately assess the algorithm for chromaticity constancy, the 1269 Munsell reflectance spectra were normalized in such a way that all these samples have a luminance L value in CIELab color space of 50 units. This normalization ensured that the only differences in the colors of all the reflectance spectra was in their chromaticity. The particular value of L (L=50) was chosen because it is the mid range of the 'L' axis, and it is also used in defining the CIE standard color difference model (E94) [18]. In the new test sets applied in this paper the reflectance pairs differ only by their chromaticity. In CIELab space there are several qualitative descriptors defined depending on the application [11]. One example of a set of qualitative descriptors is defined by Abrardo *et al.* [19], who describe colors that differ by between 1.0 to 3.0 CIELab units as very good matches to each other and 3.0 to 6.0 CIELab units as good color matches to each other [19,20]. From this normalized Munsell data set two sets of test reflectances were chosen. Each of these test reflectance sets has 100 pairs of reflectances with pairwise distances of 2.99 to 3.01 and 5.995 to 6.005 CIELab units, respectively.

The reflectance data and the daylight spectra were sampled at 1 nm intervals, and the response of each sensor to the different reflectances was obtained by integrating the product of Munsell reflectance, the CIE standard daylight spectra, and a Gaussian function representing the spectral sensitivity of the sensors. For an image sensor with spectral sensitivity $F(\lambda)$ imaging a scene with reflectance $S(\lambda)$ the noisy response R is given by

$$R = N(1, \sigma^2) \int_{400\ \mathrm{nm}}^{700\ \mathrm{nm}} S(\lambda) E(\lambda) F(\lambda) d\lambda, \tag{6}$$

where $E(\lambda)$ is the power spectral distribution of the light source. $N(1, \sigma^2)$ is a normal distribution with a mean value of one and a variance that determines the signal-to-noise ratio (SNR) of the response.

For each sensor response the sensor noise ($N(1, \sigma^2)$) was simulated using 100 normally distributed random numbers. The final step in the model was to represent the effects of using an analog-to-digital converter (ADC) to convert the sensor responses to digital quantities. In representing the quantizer effect the first stage is to determine the maximum sensor response. A white standard reflectance and the CIE standard daylight illuminant (6500 K) were used to determine this maximum response. As the optimization process shifts the peak position of each of the sensors the maximum sensor response was calculated by shifting the sensor's peak position in the visible spectrum in 1 nm steps. This way the maximum sensor response corresponding to different sensor spectral bandwidths was calculated. This maximum sensor response was then divided by 2^n, where n is the number of bits applied in the quantizer, and each sensor response was then approximated to the nearest one of these quantized levels. In this investigation 8 and 10 bits were applied to quantize the sensor responses.

When capturing an image of a scene with an image sensor some parts of the scene are well exposed and generate the maximum sensor response, and some other parts of the scene will be underexposed. To avoid either underexposure or overexposure of any of the modeled sensor response the spectral power distribution of the daylight illuminant was scaled in such a way that the sensor responses are all near the middle of the ADC range. Finally, the features were obtained from the noisy sensor responses using the method described in Eqs. (1) and (2).

The model-based algorithm was tested with different levels of sensor noise. The SNR of data available from any

camera depends on several factors including the charge storage capacity of the pixel, the noise introduced by the readout electronics, and the photon shot noise [21]. Fowler [21] modeled the expected variations of the SNR of digital cameras and showed that good quality cameras that have pixels with a large charge storage capacity give an SNR of larger than 30 dB for all the photocurrents that can be detected when a 10 bit ADC is used to represent the response from each pixel. Based on their visual psychophysical experiments Xiao *et al.* [22] report that an imaging device should be able to achieve a SNR of 30 dB or above across the whole dynamic range to render the photon noise invisible [22]. Imagers are available with SNRs larger than 40 dB [23]. We have therefore investigated the performance of the model-based algorithm with sensor noise of 30 dB and 40 dB. The sensor noise with these two SNR values was simulated by generating normally distributed random numbers (100 samples) with standard deviations of 3% and 1%, respectively.

The Mahalanobis distance boundary was drawn for both members of each test reflectance pair on the feature space using the method described above. Two typical Mahalanobis distance boundaries drawn around the responses from one pair of Munsell reflectances when they are illuminated by 20 CIE standard daylight spectra and with 100 samples of noise that represent 40 dB Gaussian noise are shown in Fig. 2. The number of points falling within the correct boundary was counted for all pairs of reflectances in a test set, and the percentage of points that fell within the correct boundary was recorded. Figure 3 shows the test results of the model-based algorithm when applying the responses generated by evenly spread sensors with different FWHM (20 nm to 200 nm). In this test the 3- and 6-unit Munsell test data sets were illuminated with CIE standard daylight spectra. As expected the performance of the algorithm degrades when the noise level is increased. This is because as the noise level increases the variability in the responses increases as well, and this variability leads to increases in the size of the clusters, and more points fall outside the correct Mahalanobis distance boundary. Therefore the performance of the algorithm drops with decreasing SNR. The other observation is that the overall performance of the algorithm drops with the sensor width. The reason is that as the width increases the overlap between the adjacent sensitivity function increases. This increase in overlap between spectral sensitivities leads to an increase in correlation between the sensor responses and results in a degradation in performance.

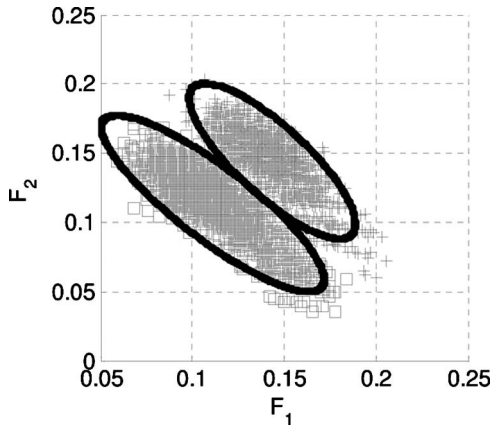

Fig. 2. Typical Mahalanobis distance boundaries for a pair of Munsell samples when illuminated with 20 spectra of CIE standard daylights. Noise was simulated by generating 100 values of random numbers that represent Gaussian noise of 40 dB.

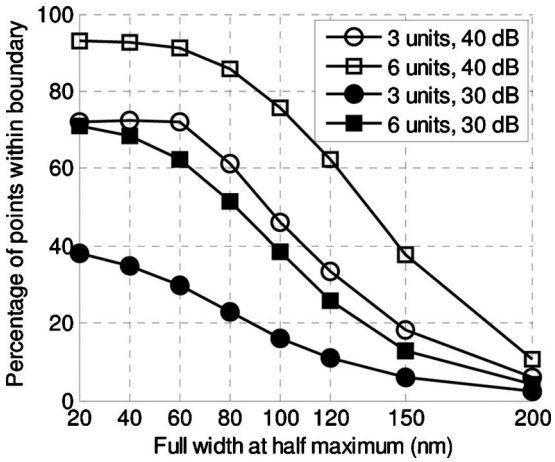

Fig. 3. Results of the model-based algorithm when testing with evenly spread sensor responses. Gaussian noise of 30 dB and 40 dB was applied to the sensor responses. The resulting linear responses were quantized to 10 bits. Munsell 3- and 6-units test sets were illuminated with the CIE standard test daylights.

The model-based algorithm relies on the blackbody model of the illuminant spectrum in calculating the correct channel coefficients to discount the illuminant effect from the sensor responses. The performance of this algorithm might be improved by adapting the algorithm to estimate illuminant effects with real illuminants instead of using the blackbody model. To estimate the illuminant effect on the sensor responses an approach proposed by Finlayson and Drew [7] was applied. This algorithm was also developed using the assumptions that the power spectral distribution of the illuminant can be approximated by a blackbody spectrum, and the image sensors respond to a single wavelength. In Finlayson and Drew's [7] approach the ratios of sensor responses were taken to remove any dependency on the illuminant intensity and scene geometry. These normalized responses are then projected in the direction of illuminant-induced variation on the sensor responses. This projection results in a 2D space that is approximately independent of illuminant. In finding the direction of illuminant-induced variations the CIE standard illuminants were used instead of blackbody illuminants as used by Finlayson and Drew [7]. As this algorithm normalizes the lightness component of a color the 2D space formed by this approach also represents the chromaticity of a color.

Finlayson and Drew's algorithm [7] and the model-based algorithm [1] were tested with Munsell reflectance data and CIE standard daylight. The results obtained with Gaussian sensors with peak spectral responses at 437.5 nm, 512.5 nm, 587.5 nm, and 662.5 nm when the SNR of the data from the sensors is 40 dB are shown in Fig. 4. Applying Finlayson and Drew's [7] algorithm to remove the illuminant effect gives only a slight improvement compared to the model-based algorithm. The performance of both of these algorithms might be improved by

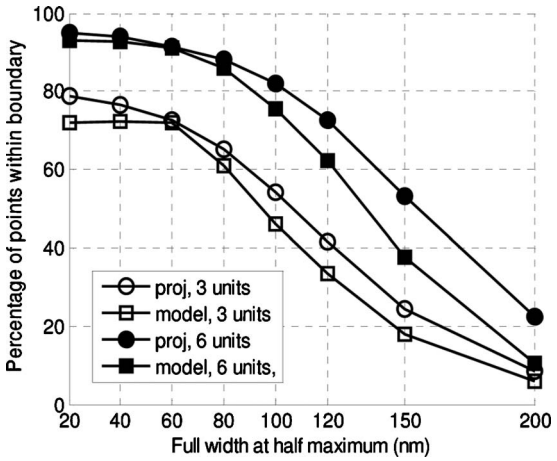

Fig. 4. Test results of model-based algorithm and Finlayson and Drew's [7] algorithm with evenly spread sensors. Gaussian noise of 40 dB was applied to the sensor responses and the resulting linear responses were quantized to 10 bits. Munsell data (3 and 6 units) were illuminated by the CIE standard test daylights.

optimizing the sensor characteristics applied to capture the data required for the algorithms.

3. OPTIMIZATION

The sensor parameters of both the model-based and Finlayson and Drew's [7] algorithms were optimized separately in such a way that the ability to identify perceptually similar colors is improved. In the rest of the discussion the optimized version of Finlayson and Drew's [7] approach is referred as a projection-based algorithm. For this initial study the number of variables was limited by assuming that the spectral bandwidth of all the sensors was the same and the parameters of the algorithms were optimized with different sensor spectral bandwidths. In optimizing the sensor parameters 100 pairs of Munsell reflectance samples with members separated by 1 CIELab unit were illuminated by 20 spectra of CIE standard daylight. Again this set of spectra was chosen so that it had a distribution of CCTs similar to the measured daylight. However, the individual CCTs of this set of spectra (CIE standard training daylights) are different from the CCTs of the test spectra. This set of CIE standard training daylights was used to determine the number of points that fall inside the correct boundaries using the method described in Section 2. The inverse of this number was then used as the error measure in the optimization process. In particular if the gradient of this error measure for a set of independent parameters p is $\nabla G(p)$ then the new parameters

$$p_{n+1} = p_n - \varepsilon \nabla G(p_n), \tag{7}$$

where ε is the parameter step size. If this change in parameters was found to increase the error measure then the change in parameters must be too large. In these circumstances the step size was reduced by a factor of 0.9 until a set of parameters was found that reduced the error measure. The gradient at this new set of parameters was then calculated and a new set of parameters determined until it was impossible to reduce the error measure.

For the model-based algorithm the two outer sensors' peak positions and the channel coefficients were taken as the independent parameters in this optimization. The two inner sensor positions can be calculated from Eqs. (3) and (4). The outer sensor positions and the two channel coefficients were taken as the independent parameters to make sure that the sensitivity functions cover the entire visible region and that both the neighboring sensors contribute approximately equally in estimating the illuminant effect on the inner sensors, respectively. For the projection-based algorithm the four peak sensor positions were taken as the independent parameters in the optimization. The optimization was performed separately by normalizing the sensor responses by the responses generated by sensors 1 to 4 and for different spectral bandwidths of the Gaussian sensors interested between 20 nm to 200 nm. Two sets of sensible starting sensor positions were chosen for the investigation. The first is the evenly spread sensors used to generate the chromaticity space in Fig. 1 ; the second set is the equal-weight sensors with peak positions at 437.5 nm, 493.5 nm, 565.5 nm, and 662.5 nm. For the model–based algorithm, for a given sensor position the initial channel coefficients can be cal-

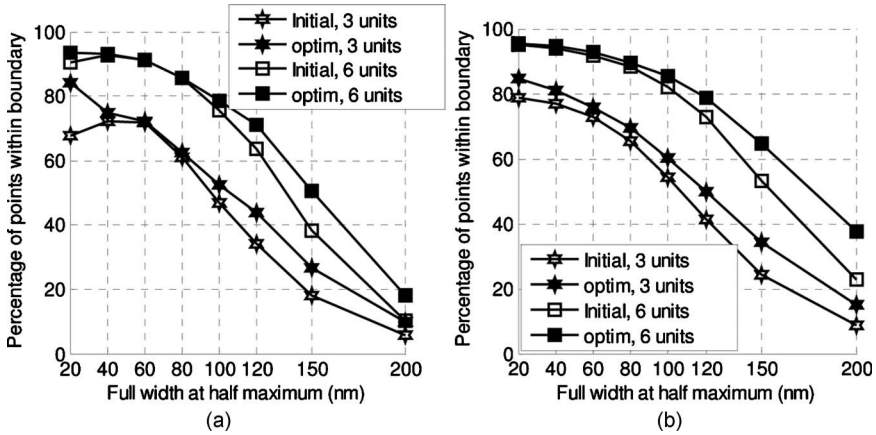

Fig. 5. Initial and optimized performance of the model-based and projection-based algorithm. Gaussian sensors were optimized with Munsell data (1 unit) and the CIE standard training daylights. Munsell data (3 and 6 units) and the CIE standard test daylights are applied in testing both algorithms. Gaussian noise of 40 dB was applied to the sensor responses and the resulting linear responses were quantized to 10 bits. (a) Model-based algorithm. (b) Projection-based algorithm.

Table 1. Optimized Parameters of the Model-Based Algorithm

Sensor ID	1	2	3	4
Peak position (nm)	434.4	513.6	593.3	676.9
Channel coefficient		α=0.425, γ=0.444		

Table 2. Optimized Parameters of the Projection-Based Algorithm When Normalizing the Sensor Responses by the Response of Sensor 3

Sensor ID	1	2	3	4
Peak position (nm)	400.0	487.6	556.0	675.1
Vector 1		−0.3488 0.7706 −0.5332		
Vector 2		0.8407 0.5087 0.1852		

culated from Eqs. (3) and (4). For the projection-based algorithm the unit vectors can be calculated by applying eigen vector decomposition on the space formed by the ratio of sensor responses (this was done in logarithmic scale) generated by the 20 spectra of CIE standard daylight and the 1-unit Munsell reflectance data set. Test results with initial and final parameters of model-based and projection-based algorithm for 40 dB sensor noise are shown in Fig. 5. It can be seen that the performance of the algorithms has improved slightly when testing with Munsell and CIE standard test daylights. The improvement in performance when testing with 3-unit reflectance data is larger than that obtained with the 6-unit data. This is particularly true of the model-based algorithm. The reason for this is probably that although both algorithms have been derived assuming sensors that respond at a single wavelength the model-based algorithm lacks the flexibility to deal with the effects of breaking this assumption that arise from the data-based projection operation in the alternative algorithm. Consequently the performance of the model-based algorithm is more dependent on sensors with narrow spectral responses. However, sensors with narrow spectral responses are difficult to manufacture and, since they are starved of photons, they would require long exposure times. A possible limit to the narrowest spectral responses that might be both possible and practical is suggested by the Sony DXC930 camera, which has three types of pixels sensors with spectral widths of approximately 80 nm. Of all the different sets of optimized sensors that have been obtained those with a FWHM of 80 nm are therefore particularly interesting. Their parameters are therefore listed in Tables 1 and 2, and their responses are shown in Fig. 6. The parameters listed in Tables 1 and 2 were obtained when starting the optimization with evenly spread sensors. For the projection-based algorithm typical optimized sensors listed in Table 2 were obtained when normalizing the sensor responses by the response of sensor 3. An important conclusion from the parameters such as those in Tables 1 and 2 is that the spectral responses of optimum sensor sets can be quite different. In view of the considerable investment needed to develop a sensor with a particular peak spectral response it is important to ensure that the correct spectral responses are specified.

4. ROBUSTNESS OF THE CONCLUSIONS TO THE CHANGE IN REFLECTANCE DATA

In Section 3 the performance of the projection-based and model-based algorithms was investigated with the Munsell reflectance data set. However, this data set is believed to be generated with a limited number of basis functions [24]. Therefore the performance of the algorithms was investigated with real-world measured reflectances. The reflectances used are flower reflectances measured around the world [25]. These 2211 measured reflectances were normalized as described in Section 2. From these normalized reflectances two sets of 100 pairs of reflectances with member separations from 2.998 to 3.002 and from 5.998 to 6.001 units in the CIELab color space were obtained. Both of the algorithms were compared with evenly spread sensors, and the results are shown in Fig. 7. Similar to the results presented in Section 3 these results also show that the improvement achieved by applying the projection method is small.

As both the algorithms were optimized with Munsell data and CIE standard daylight illuminant the algorithms were tested with floral data to ensure that the process used to find the optimum sensor combination has not overfit the sensor responses to the Munsell data. The re-

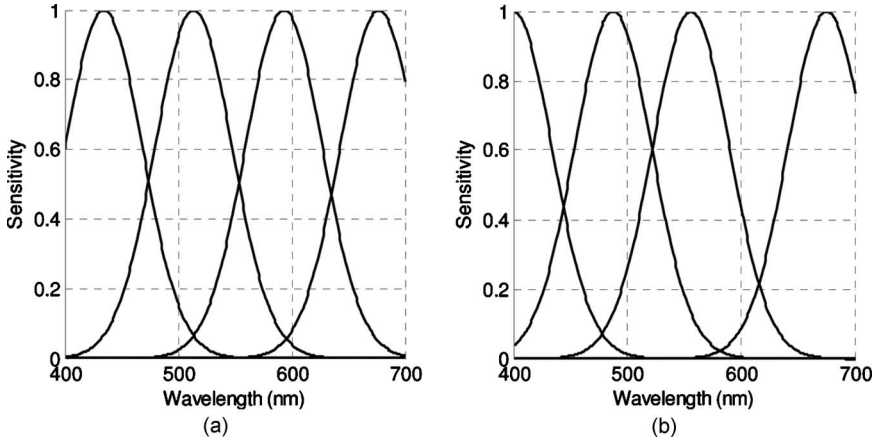

Fig. 6. Sensitivity functions of optimized Gaussian sensors for (a) model-based algorithm and (b) projection-based algorithm. In optimizing the Gaussian sensors Munsell data (1 unit) and the CIE standard training daylights are applied.

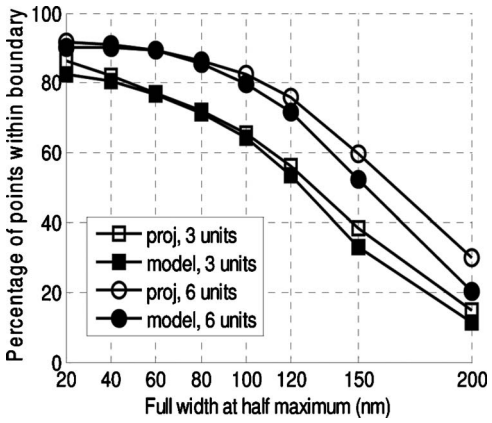

Fig. 7. Performance of model-based and projection-based algorithms with evenly spread sensors. In this test floral reflectances were illuminated by the CIE standard test daylights. Gaussian noise of 40 dB was applied to the sensor responses and the resulting linear responses were quantized to 10 bits. (a) Model-based algorithm. (b) Projection-based algorithm.

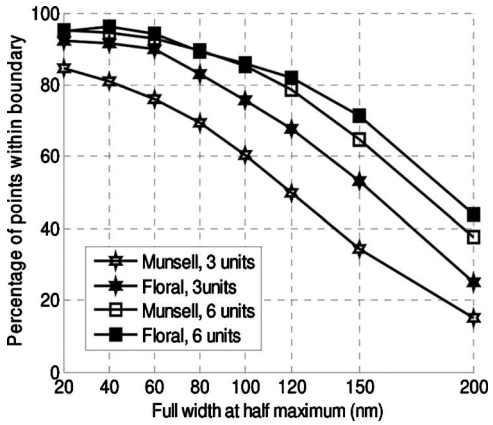

Fig. 9. Test results of the projection-based algorithm with optimized sensor responses. Gaussian sensors were optimized with Munsell data (1 unit) and the CIE standard training daylights. The CIE standard test daylights were applied in illuminating the reflectances. Gaussian noise of 40 dB was applied to the sensor responses and the resulting linear responses were quantized to 10 bits.

sults in Fig. 8 show that the optimization has found a sensor combination that also improves the performance of the algorithm with measured reflectances, and the results for both algorithms suggest that the optimization has not overfit to the Munsell data set. More importantly, these results suggest that using projection to extract features from the responses of optimized sensors leads to an improvement in the quality of the features. However, a comparison between the results obtained with the different reflectance data sets, Fig. 9, shows that the similarity between the Munsell and floral data set is not good enough for the Munsell to accurately predict the performance with floral data.

5. ROBUSTNESS TO THE CHANGE IN ILLUMINANT

In Sections 3 and 4 the performances of the model-based and the projection-based algorithms have been investigated with CIE standard daylight illuminants. However, the CIE standard daylight model was developed with three basis vectors. These basis vectors were generated using daylight measurements taken in three places (Ottawa, Canada; Rochester, New York, USA, and Middlesex, England). It is known that the daylight spectra vary with location, time of day, time of year, and weather conditions. Figure 10 shows the spectrum of CIE standard daylight (6500 K) and three of the measured daylight spectra with CCTs close to 6500 K [11]. From this figure it can be seen that even though the CIE standard daylight spectrum models the overall shape of the measured spectra, it does not represent the fine details caused by atmospheric absorption. This atmospheric absorption varies with weather conditions and the elevation of the Sun, and this elevation of the Sun varies with time and location on the Earth [18]. To ensure that any conclusions are independent of the data, the performance of the model-based and the projection-based algorithms was tested with measured daylight. In the test 146 daylight spectra measured on a day in the first week of each month in 1997 was applied [11]. The measurements were taken between 5:30 am and 7:30 pm under different types of weather conditions in Granada, Spain. Test results of both algorithms with evenly spread sensors are shown in Fig. 11. These results show that both of the algorithms are comparable

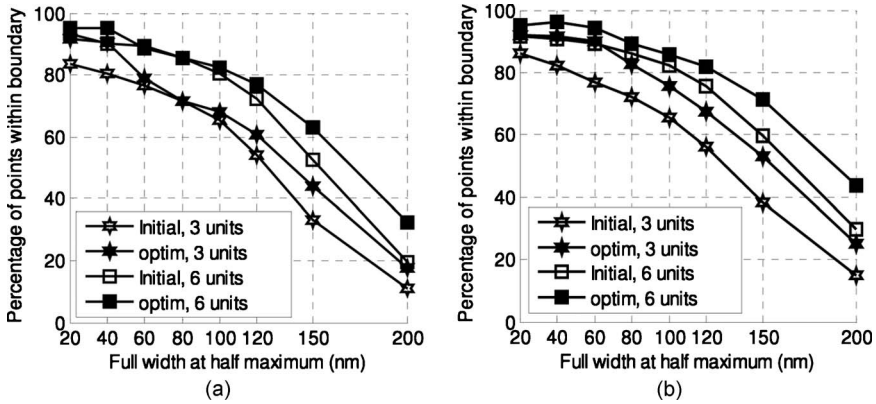

Fig. 8. Initial and optimized performance of the model-based and projection-based algorithms. In optimizing the Gaussian sensors Munsell data (1 unit) and the CIE standard training daylights are applied. Floral data (3 and 6 units) and the CIE standard test daylights are applied in testing both algorithms. Gaussian noise of 40 dB was applied to the sensor responses. The resulting linear responses were quantized to 10 bits.

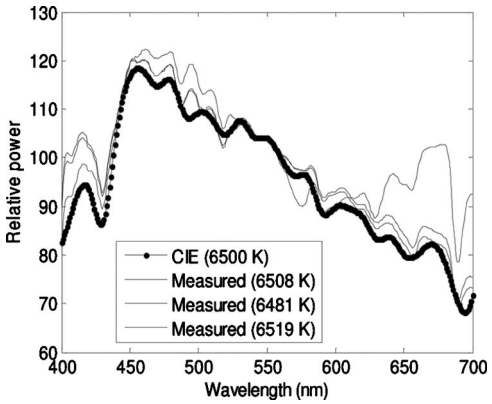

Fig. 10. Spectra of CIE (6500 K) and measured daylight with correlated color temperature 6508 K, 6481 K, and 6519 K.

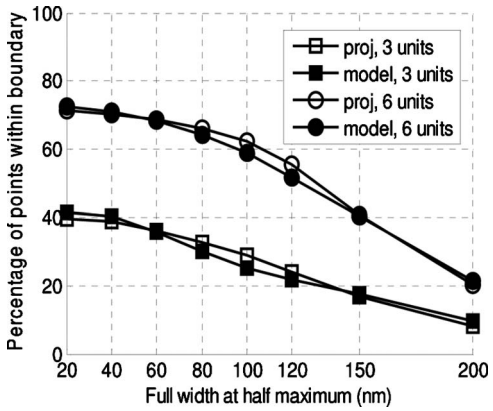

Fig. 11. Test results of model-based and projection-based algorithms when testing with responses generated by evenly spread sensor responses. Gaussian noise of 40 dB was applied to the sensor responses and the resulting linear responses were quantized to 10 bits. Munsell reflectances were illuminated by measured daylight.

in performance with evenly spread sensors. These results and the previous results suggest that the performance improvement achieved by applying the projection-based approach is small.

As the model-based and projection-based algorithms were optimized with CIE standard daylight both algorithms were tested with measured daylight. Figure 12 illustrates the performance of both algorithms with initial and optimized sensor responses when applying Munsell reflectance and measured daylight spectra. These results show that the optimization of both algorithms has not always improved the performance of the algorithms when tested with Munsell reflectances and measured daylight spectra. This suggests that the CIE standard daylight might not be a good representative data set for measured daylight at a particular location.

The performance of the projection-based algorithm was compared with CIE standard daylight and measured daylight. Figure 13 illustrates the performance of the algorithm with Munsell reflectances illuminated by CIE standard daylight and measured daylight. It can be seen that there is a significant performance drop when changing the illuminant from CIE standard daylight to the measured daylight. The reason could be that the CIE standard daylight model was developed using the daylight measurements taken in three places, and the measured illuminants applied in our investigation were measured in Granada, Spain, under different types of weather conditions over a longer period of time (two years). As the measurements used in the CIE standard model and our test illuminants are taken from different parts of the world the CIE standard daylight might not adequately represent the measured daylight used in this paper. This leads to the performance difference when changing the illuminants from CIE standard daylight to the measured daylight. The results presented in Figs. 12 and 13 suggest that the CIE standard daylight model is not a good representative spectra that could be used to optimize the sensor characteristics or to predict the exact performance of a color constancy algorithm.

6. CONCLUSIONS

The performance of a blackbody-model-based algorithm for color constancy under daylight illuminant was investigated and compared with the performance obtained using an alternative algorithm proposed by Finlayson and Drew [7]. Both of these algorithms are based on the assumptions that four types of sensors are used that each respond to a different single wavelength, and that the il-

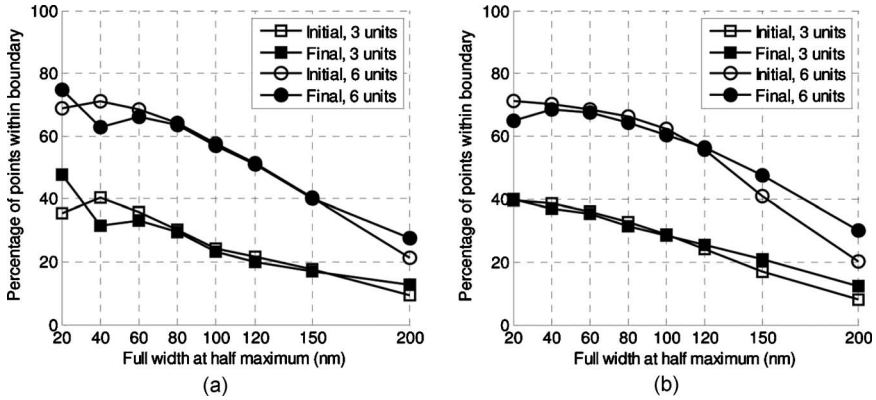

Fig. 12. Initial and optimized performance of (a) model-based and (b) projection-based algorithm with Munsell and measured daylight spectra. Gaussian sensors were optimized with Munsell data (1 unit) and the CIE standard training daylights. Both the algorithms were tested with Munsell reflectance data (3, 6 units) and 146 spectra of measured daylight. Gaussian noise of 40 dB was applied to the sensor responses and the resulting linear responses were quantized to 10 bits.

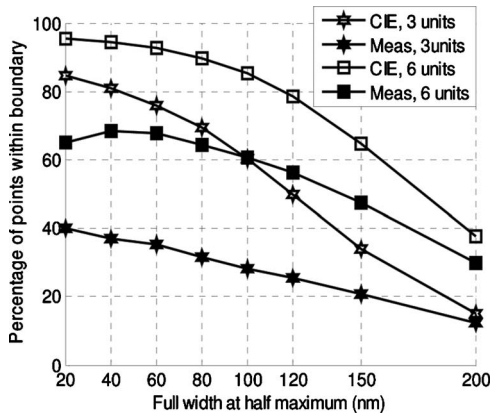

Fig. 13. Test results of the projection-based algorithm with optimized sensors when illuminating the Munsell reflectance samples with the CIE standard test daylight spectra and measured daylight spectra. Gaussian noise of 40 dB was applied to the sensor responses and the resulting linear responses were quantized to 10 bits.

luminant spectral power density can be approximated by that of a blackbody. Results that have been presented show that either of the algorithms can be used to obtain useful information even when each of the sensors responds to a relatively wide range of wavelengths and there is significant overlap between the sensor responses. The difference between the algorithms is that the model-based algorithm only requires the calculation of two coefficients, while the alternative algorithm requires a projection. While the coefficient calculation is very easy the best projection can only be calculated using the measured sensor responses to several colors under different illuminants. Although this makes the projection algorithm less convenient for the user, the projection algorithm might be expected to give better results for real illuminant spectra and sensor responses. Somewhat surprisingly the results that have been obtained suggest that the performance of the two algorithms when used to process noisy data from a sensible choice of sensors is almost identical. Using the more sophisticated algorithm to process the sensor data is therefore not a reliable way to achieve better results.

An alternative approach to obtaining better results from a set of sensors is to exploit the flexibility offered by families of different organic chromophores [10] to optimize the spectral responses of the four different types of sensors. The effect of optimizing the spectral responses of the four sensors on the results that could be obtained using the two algorithms has been investigated using a steepest descent algorithm. The optimization metric that was chosen for this investigation was the separability of perceptually similar colors. The data used during optimization were the freely available Munsell reflectance data and the CIE standard daylight spectra. As expected this optimization improved the separability of perceptually similar Munsell samples illuminated by CIE standard daylight spectra, and the results obtained were comparable to those achieved by the human visual system. However, it was observed that the optimization did not always improve the performance of the model-based and projection-based algorithms when measured illuminant spectra were used. This shows that optimizing the sensor parameters with CIE standard daylight does not necessarily result in an optimum set of sensors for real daylight. In addition, the results obtained with CIE standard daylight are significantly better than those obtained with measured daylight. These two observations suggest that the CIE standard daylight spectra do not represent the measured daylight spectra in enough detail for the CIE standard daylight spectra to be used to reliably optimize a system designed to distinguish a quite subtle difference between two colors. Although the optimization could be performed using the data measured in Granada it is not clear that the resulting system would perform well in different locations. Before a set of optimum sensors can be reliably defined, daylight spectra from different parts of the world are required.

REFERENCES

1. S. Ratnasingam and S. Collins, "Study of the photodetector characteristics of a camera for color constancy in natural scene," J. Opt. Soc. Am. A **27**, 286–294 (2010).
2. J. A. Marchant. and C. M. Onyango, "Shadow-invariant classification for scenes illuminated by daylight," J. Opt. Soc. Am. A **17**, 1952–1961 (2000).
3. E. H. Land and J. J. McCann, "Lightness and retinex theory," J. Opt. Soc. Am. **61**, 1–11 (1971).
4. B. K. P. Horn, "Determining lightness from an image," Comput. Graph. Image Process. **3**, 277–299 (1974).
5. G. D. Finlayson and S. D. Hordley, "Color constancy at a pixel," J. Opt. Soc. Am. A **18**, 253–264 (2001).
6. M. Ebner, *Color Constancy*, Wiley Series in Imaging Science and Technology (Wiley, 2007).
7. G. D. Finlayson and M. S. Drew, "4-sensor camera calibration for image representation invariant to shading, shadows, lighting, and specularities," in *Proceedings of IEEE International Conference on Computer Vision* (IEEE, 2001), pp. 473–480.
8. J. Romero, J. Hernández-Andrés, J. L. Nieves, and E. M. Valero, "Spectral sensitivity of sensors for a color-image descriptor invariant to changes in daylight conditions," Color Res. Appl. **31**, 391–398 (2006).
9. S. Takada, M. Ihama, M. Inuiya, T. Komatsu, and T. Saito, "CMOS color image sensor with overlaid organic photoconductive layers having narrow absorption band," Proc. SPIE **6502**, 650207 (2007).
10. R. Jansen van Vuuren, K. D. Johnstone, S. Ratnasingam, H. Barcena, P. C. Deakin, A. K. Pandey, P. L. Burn, S. Collins, and I. D. W. Samuel, "Determining the absorption tolerance of single chromophore photodiodes for machine vision," Appl. Phys. Lett. **96**, 253303 (2010).
11. J. Hernández-Andrés, J. Romero, J. L. Nieves, and R. L. Lee, Jr, "Color and spectral analysis of daylight in southern Europe," J. Opt. Soc. Am. A **18**, 1325–1335 (2001).
12. S. T. Hendersons and D. Hodgkiss, "The spectral energy distribution of daylight," Br. J. Appl. Phys. **14**, 125–133 (1963).
13. Munsell Color Science Laboratory, "Daylight spectra," http://mcsl.rit.edu/.
14. S. Tominaga and B. A. Wandell, "Natural scene-illuminant estimation using the sensor correlation," in Proc. IEEE **90**, 42–56 (2002).
15. Database—"Munnsell Colours Matt," ftp://ftp.cs.joensuu.fi/pub/color/spectra/mspec/.
16. R. Kawakami, J. Takamatsu, and K. Ikeuchi, "Color constancy from blackbody illumination," J. Opt. Soc. Am. A **24**, 1886–1893 (2007).
17. J. Yang, W. Lu, and A. Waibel, "Skin-color modeling and adaptation," Lect. Notes Comput. Sci. **1352**, 687–694(1997).
18. H. C. Lee, *Introduction to Color Imaging Science* (Cambridge University Press, 2005), pp. 46–47, 138–141, 450–459.

19. A. Abrardo, V. Cappellini, M. Cappellini, and A. Mecocci, "Art-works color calibration using the VASARI scanner," in *Proceedings of IS&T and SID's 4th Color Imaging Conference: Color Science, Systems and Applications* (IS&T, 1996), pp. 94–97.
20. J. Y. Hardeberg, "Acquisition and reproduction of color images: colorimetric and multispectral approaches," Ph.D. dissertation (Ecole Nationale Supérieure des Télécommunications, 1999).
21. B. Fowler, "High dynamic range image sensor architectures," in *High Dynamic Range Imaging Symposium and Workshop*, Stanford University, California, 2009, http://scien.stanford.edu/HDR/HDR_files/Conference%20Materials/Presentation%20Slides/Fowler_WDR_sensor_architectures_9_8_2009.pdf.
22. F. Xiao, J. E. Farrell, and B. Wandell, "Psychophysical thresholds and digital camera sensitivity: The thousand photon limit," Proc. SPIE **5678**, 75–84 (2005).
23. S. Winkler and S. Susstrunk, "Visibility of noise in natural images," Proc. SPIE **5292**, pages 121–129 (2004).
24. J. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen, "Characteristic spectra of Munsell colors," J. Opt. Soc. Am. A **6**, 318–322 (1989).
25. S. E. J. Arnold, V. Savolainen, and L. Chittka , "FReD: The floral reflectance spectra database," Nature, http://dx.doi.org/10.1038/npre.2008.1846.1.