Spectral-daylight recovery by use of only a few sensors

Javier Hernández-Andrés, Juan L. Nieves, Eva M. Valero, and Javier Romero

Departamento de Óptica, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

Received April 20, 2003; revised manuscript received July 28, 2003; accepted August 28, 2003

Linear models have already been proved accurate enough to recover spectral functions. We have resorted to such linear models to recover spectral daylight with the response of no more than a few real sensors. We performed an exhaustive search to obtain the best set of Gaussian sensors with a combination of optimum spectral position and bandwidth. We also examined to what extent the accuracy of daylight estimation depends on the number of sensors and their spectral properties. A set of 2600 daylight spectra [J. Opt. Soc. Am. A **18**, 1325 (2001)] were used to determine the basis functions in the linear model and also to evaluate the accuracy of the search. The estimated spectra are compared with the original ones for different spectral daylight and skylight sets of data within the visible spectrum. Spectral similarity, colorimetric differences, and integrated spectral irradiance errors were all taken into account. We compare our best results with those obtained by using a commercial CCD, revealing the CCD's potential as a daylight-estimation device. © 2004 Optical Society of America

OCIS codes: 150.0150, 150.2950, 040.0040.

1. INTRODUCTION

Daylight spectral power distribution (SPD) data are important to many scientific disciplines and areas of industry and technology; thus a complete knowledge of daylight SPDs in the visible, ultraviolet, and near-infrared ranges of the spectrum, sometimes at different sites and for varying atmospheric conditions, is often called for. Ideally a high-resolution daylight SPD is measured in different spectral bands with a spectroradiometer, but either practical or economic restraints sometimes prevent us from using complex and expensive instrumentation.

Because of the high temporal and spatial variability of daylight, the concept of multispectral images and the design of multispectral devices have received considerable attention from researchers during the past decade. A multispectral image is an image in which each pixel contains information about the spectral reflectance of the image scene or else, as in the present paper, about the spectral power distribution of the illumination impinging on the acquisition device. A multispectral device to recover the daylight or skylight spectra in each pixel of the image may be relevant to several fields of application, such as remote sensing or astronomy, where the spectral information from the illumination is spatially and temporally variable.

Several authors have devoted their research to estimating solar spectral radiation, using different techniques.[1,2] In 1984 Michalsky and Kleckner[1] described two methods of estimating the low-resolution spectral distribution of direct solar radiation by using seven broadband sensors spanning a wavelength range of between 360 and 1030 nm. In another study,[2] intended to demonstrate the practicality of these techniques, Michalsky showed that his integrated estimated spectra had a margin of error within 1% but that agreement was poorer in their spectral detail. In fact, the average fractional deviation (i.e., the sum of the products of the absolute values of the differences in spectral irradiance and the wavelength interval normalized by the integrated spectral irradiance) was 0.052 for the seven test spectra in question. He also found that the results were affected to a greater extent by the spectral position of the filters than by their bandwidths.

Different techniques for estimating spectral-reflectance curves have also been published.[3–8] They all have in common recovery of a continuous function by using only a small number of samples of that function. Connah *et al.* used linear models to recover spectral reflectances[6–8] and studied the effect of different parameters, such as the number of Gaussian sensors, the sensors' spectral properties, and the choice of illuminant and sensor noise, on the performance of the recovery algorithm. They found that the optimum spectral sensitivity of the sensor depended on the accuracy pursued and the spectral properties of the illuminant used, concluding that "all these parameters interact with each other in a complex way and therefore the optimum set of parameters cannot easily be determined" (Ref. 7, p. 619).

For very low spectral-resolution detail, Michalsky's techniques[1,2] are perfectly adequate. Daylight spectra, however, unlike spectral reflectances, are not smoothly varying functions, and so a normal resolution spectrum (e.g., one sampled every 5 nm) consists of abrupt emission and absorption features that are due to either the solar or the terrestrial atmosphere. Our aim here is to examine the reliability of a 5-nm-resolution spectral-daylight recovery algorithm by using a set containing only a few Gaussian sensors, carefully chosen for their close similarity to the spectral sensitivity of commercial camera channels. The algorithm is based on linear models frequently used in artificial-vision algorithms for recognizing and identifying colors.[9–12] We investigated systematically

1084-7529/2004/010013-11$15.00

the influence of the number of sensors, their spectral location, and their bandwidth. To evaluate the accuracy of our recoveries, we used both spectral and colorimetric error measurements and an extensive set of experimental daylight spectra. We completed our study with a commercial CCD to compare its performance with a daylight-estimation algorithm of this sort.

2. DAYLIGHT LINEAR MODELS

Daylight-measurement campaigns[13–17] were conducted in several countries during the 1960s and 1970s to determine representative daylight spectra. Despite the variety of techniques and equipment used by the various researchers in these campaigns, their studies consistently arrived at two basic conclusions. First, the chromaticities of different phases of daylight lie near the Planckian locus of the 1931 chromaticity diagram adopted by the Commission Internationale de l'Eclairage (CIE); second, different daylight power spectra are closely correlated with one another, and this underlying similarity has many practical uses. In fact, a low-dimensional, linear efficient representation of a daylight SPD benefits from daylight spectral correlation.

We denote a daylight SPD by $E(\lambda)$, where λ is the wavelength variable with values in the visible range of the spectrum. Daylight $E(\lambda)$ SPDs are nonnegative functions. We can describe these spectra by a linear model[9–12]:

$$E(\lambda) = \sum_{i=1}^{p} \epsilon_i V_i(\lambda), \tag{1}$$

where $V_i(\lambda)$ are fixed and known basis functions and ϵ_i are weighting coefficients (expansion coefficients) or coordinates. But how well does a linear model capture the range of the spectral variation of daylight? To put it another way, how many basis functions are required to approximate accurately a large daylight data set $E(\lambda)$ that corresponds to different atmospheric conditions? If a daylight SPD is sampled over a wavelength range of 380 to 780 nm at 5-nm intervals ($N = 81$ samples), we need in principle 81 basis functions in Eq. (1) in order to represent $E(\lambda)$ exactly. In practice, however, the correlation among daylight spectra means that we may set $p < N$ without losing any meaningful spectral information.

There are infinite possible choices of basis functions $V_i(\lambda)$. If the basis functions in Eq. (1) are calculated to minimize the mean squared error[9,10,12] of a set of empirical spectra, then they are all orthogonal eigenvectors and may be obtained from a principal component analysis (PCA). Mathematically stated,

$$\langle V_j | V_k \rangle = 0 = \sum_{i=1}^{N} V_j(\lambda_i) V_k(\lambda_i), \qquad j \neq k, \tag{2}$$

where $\langle | \rangle$ is the inner product. In this case the weighting coefficients ϵ_i can be obtained by

$$\epsilon_i = \langle E(\lambda) | V_i(\lambda) \rangle. \tag{3}$$

Therefore the mathematical reconstruction $E_R(\lambda)$ of the original SPD, $E(\lambda)$ from the eigenvectors $V_i(\lambda)$ is given by

$$E_R(\lambda) = \sum_{i=1}^{p} \langle E(\lambda) | V_i(\lambda) \rangle V_i(\lambda), \tag{4}$$

where p is the number of eigenvectors with which we wish to recover the spectral distribution.

In a previous study[18] we made a PCA using 2600 daylight spectra (global spectral irradiances on a horizontal surface from direct sunlight—when present—and the entire sky) recorded over a period of two years in the city of Granada (Spain) from sunrise to sunset, under nearly all weather conditions. We found that the first five eigenvectors (basis functions) accounted for 99.991% of the observed variance between 380 and 780 nm and that up to five eigenvectors were needed to recover daylight SPDs accurately in the visible region of the spectrum (380–780 nm). We found in fact that, according to CIE recommendations,[19] it was possible with only three eigenvectors to achieve daylight recoveries that are colorimetrically indistinguishable from the corresponding original daylight SPD.

3. SPECTRAL-DAYLIGHT RECOVERY

Previous results indicate that low-dimensional linear models suffice for the representation of both natural[18,20–22] and artificial[23] illuminants. More recently, DiCarlo and Wandell[24] described a possible improvement in the quality of linear models if the coefficients ϵ_i are highly structured (i.e., if some of these coefficients are a function of others), it being feasible to estimate values for several coefficients from a knowledge of others. They asserted that a knowledge of the structure of the coefficients ϵ_i may increase the quality of the estimation with fewer sensors and lower cost and thus help in sensor design.

No attempts have been made, however, to find the optimum sensor design when what is of interest is the practicality of linear models for the estimation of natural illuminant spectra. It is impossible, for example, to have sensors with the same spectral transmittance $R_k(\lambda)$ as the eigenvectors $V_i(\lambda)$ obtained from a PCA, because eigenvectors, being orthogonal, must have negative values, as shown in Fig. 1. Therefore Eq. (4) is only a mathematical approach to the problem, not a realistic one. Consequently, the practicality of such a daylight-recovery algorithm relies heavily on the suitable choice of an optimum real set of sensors.

Let us assume that an array of q sensors, each with a different spectral sensitivity $R_k(\lambda)$, are directly detecting the illuminant. For daylight, $E(\lambda)$, a sensor with a spectral sensitivity $R_k(\lambda)$ will give the following response to the incident light:

$$\rho_k = \sum_{n=1}^{N} E(\lambda_n) R_k(\lambda_n). \tag{5}$$

Let $\bar{\rho}$ be the vector of the q sensor responses $\bar{\rho} = (\rho_1, \rho_2, \ldots, \rho_q)^{\mathrm{T}}$, where the T superscript denotes transpose. If we use a linear representation of $E(\lambda)$, as in Eq. (4), then Eq. (5) becomes a simple matrix equation,

$$\bar{\rho} = \hat{\Lambda} \bar{\epsilon}, \tag{6}$$

where

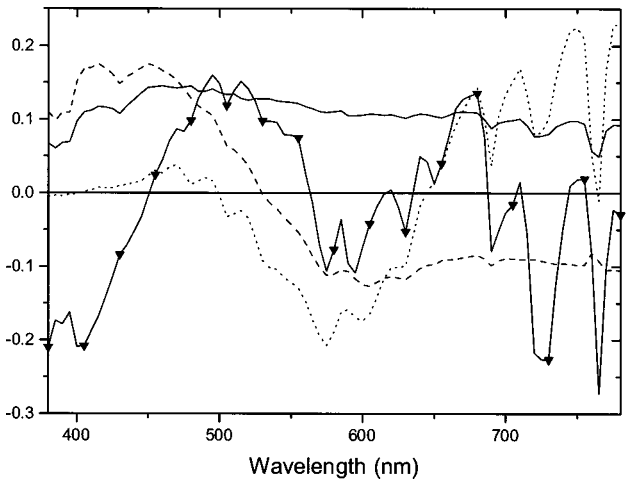

Fig. 1. Spectral distribution of eigenvectors $V_1(\lambda)$, $V_2(\lambda)$, $V_3(\lambda)$, and $V_4(\lambda)$ for our 2600 daylight measurements (see Ref. 18). Plain solid curve, $V_1(\lambda)$; dashed curve, $V_2(\lambda)$; dotted curve, $V_3(\lambda)$; solid curve with triangles, $V_4(\lambda)$.

$$(\hat{\Lambda})_{ki} = \sum_{n=1}^{N} V_i(\lambda_n) R_k(\lambda_n), \tag{7}$$

$$(\bar{\epsilon})_i = \langle E(\lambda) | V_i(\lambda) \rangle. \tag{8}$$

We have a set of simultaneous matrix equations [Eq. (6)], all sharing the same $\hat{\Lambda}$ matrix. If the number of sensors is the same as the number of basis functions (i.e., $p = q$), then we can solve the equation $\bar{\rho} = \hat{\Lambda}\bar{\epsilon}$ for $\bar{\epsilon}$ by inverting the matrix $\hat{\Lambda}$, thus obtaining the reconstructed daylight spectrum as

$$E_R(\lambda) = \sum_{i=1}^{p} (\bar{\epsilon})_i V_i(\lambda). \tag{9}$$

If this daylight-recovery algorithm must be implemented through the use of a particular set of sensors, then, as discussed above, it is essential to choose the most adequate sensors—obviously with the constraint of positive $R_k(\lambda)$—as we now do below.

4. SEARCH FOR OPTIMUM SENSORS

Among different methods for selecting the optimal filters (see Refs. 25 and 26 for an overview of different approaches for filter selection and design), we have made an exhaustive search to find the best set of sensors for providing the least error in recovering our 2600 daylight spectral curves, using our daylight eigenvectors $V_i(\lambda)$ (Fig. 1) in the range from 380 to 780 nm with a spectral resolution of 5 nm. To do this we imposed certain restrictions and made several assumptions *a priori*. First, we limited our study to sensors that are Gaussian functions of wavelength. Although some researchers have proposed different mathematical techniques to obtain nonnegative reflectance basis functions that maintain the optimum compression properties of the PCA scheme while providing accurate reconstructions of reflectance spectra (see, for example, Refs. 27 and 28), we chose Gaussian sensors because the spectral sensitivity of the sensors of many commercial cameras is close to that of Gaussian functions.[29] Second, we limited the number of sensors to between three and five because this is the optimum number of eigenvectors needed to represent daylight spectra, depending upon the degree of accuracy required. Third, we assumed that the number of sensors is the same as the number of basis functions ($p = q$) in the recovery process, which leads to a squared matrix $\hat{\Lambda}$. And fourth, we took it that the spectral sensitivity of the sensors does not depend on the level of daylight irradiance.

We exhaustively and systematically investigated the influence of the number of Gaussian sensors, their spectral location, and their bandwidths. We allowed the peak sensitivity of each sensor to be at any wavelength from 380 to 780 nm in 5-nm steps, and the FWHM to vary from 10 to 400 nm in 5-nm steps. In this way we included in our search both narrowband and wideband sensors. Allowing the full parameter space (number of sensors, sensor peak sensitivity, and sensor bandwidth) to be searched involves high computational cost: The search for the three optimum Gaussian sensors, for instance, required the evaluation of 2.6×10^{11} filter combinations ($2.6 \times 10^{11} = 81^3 \times 79^3$). Note also that the optimum sensors will depend on the experimental data used to carry out the search and also on the eigenvectors used in the linear recovery model [Eqs. (5)–(9)].

While carrying out this exhaustive and systematic search, by using Eq. (9) we compared the reconstructed function, $E_R(\lambda)$, with the original, $E(\lambda)$, making a triple evaluation in accordance with the recommendations of Imai *et al.*,[30] who indicated that there is no single parameter that permits us to assess the validity of a spectral reconstruction. Therefore we used a triple cost function: Our evaluation consists of a spectral index (goodness-of-fit coefficient GFC,[31] based on Schwartz's inequality, previously used in Refs. 18, 20, 22, 23, and 30), a colorimetric index (CIELUV color difference[32,33]), and the usual, essential parameter in solar research, the relative error between integrated irradiances (both original and reconstructed) throughout the visible spectrum (380–780 nm). Our set of optimum sensors therefore must simultaneously maximize the mean GFC value, minimize the mean CIELUV color difference, and minimize the mean integrated irradiance error.

We have found in previous studies[18,22] that colorimetrically accurate daylight recoveries require GFC > 0.995, whereas what we might call a *good* spectral fit requires GFC > 0.999, and an almost-exact fit GFC > 0.9999. Note also that three CIELUV color-difference units are often taken to be one just-noticeable difference in technical and industrial applications. Other authors have studied and proposed different quality metrics for the optimal design of camera spectral-sensitivity functions.[34,35]

Thus by using these three cost functions to evaluate the quality of the reconstruction and to find the optimum sensors, we avoid focusing our results on just one aspect.[30] Therefore we avoid an exhaustive search that produces as optimum sensors both those that generate metameric daylight reconstructions and those that allow us to recover daylight with similar integrated irradiance but with low colorimetric and spectral quality (another form of metamerism).

5. RESULTS

In this section we present an evaluation of the accuracy of our optimization search when it is tested against a set of 2600 daylight spectra[18] and our daylight basis functions (Fig. 1). Table 1 shows the characteristics of the optimum sensors (optimized peak sensitivity and FWHM), which are also represented in the graph in Fig. 2. Connah *et al.*[7] found that the "nature of the cost function affects the spectral properties of the filters": When they used a root mean square (RMS) cost function, their three sensors were evenly spread throughout the spectrum, and when they used the CIELAB ΔE error as the cost func-

Table 1. Optimum Spectral Sensitivities of the Gaussian Sensors for Daylight Recovery

Number of Sensors	Gaussian Sensor Characteristics	Spectral Sensitivity (nm)				
		First Sensor	Second Sensor	Third Sensor	Fourth Sensor	Fifth Sensor
3	Peak sensitivity	425	440	740		
	FWHM	250	120	300		
4	Peak sensitivity	430	445	520	750	
	FWHM	290	100	40	330	
5	Peak sensitivity	450	485	505	590	750
	FWHM	230	110	110	15	370

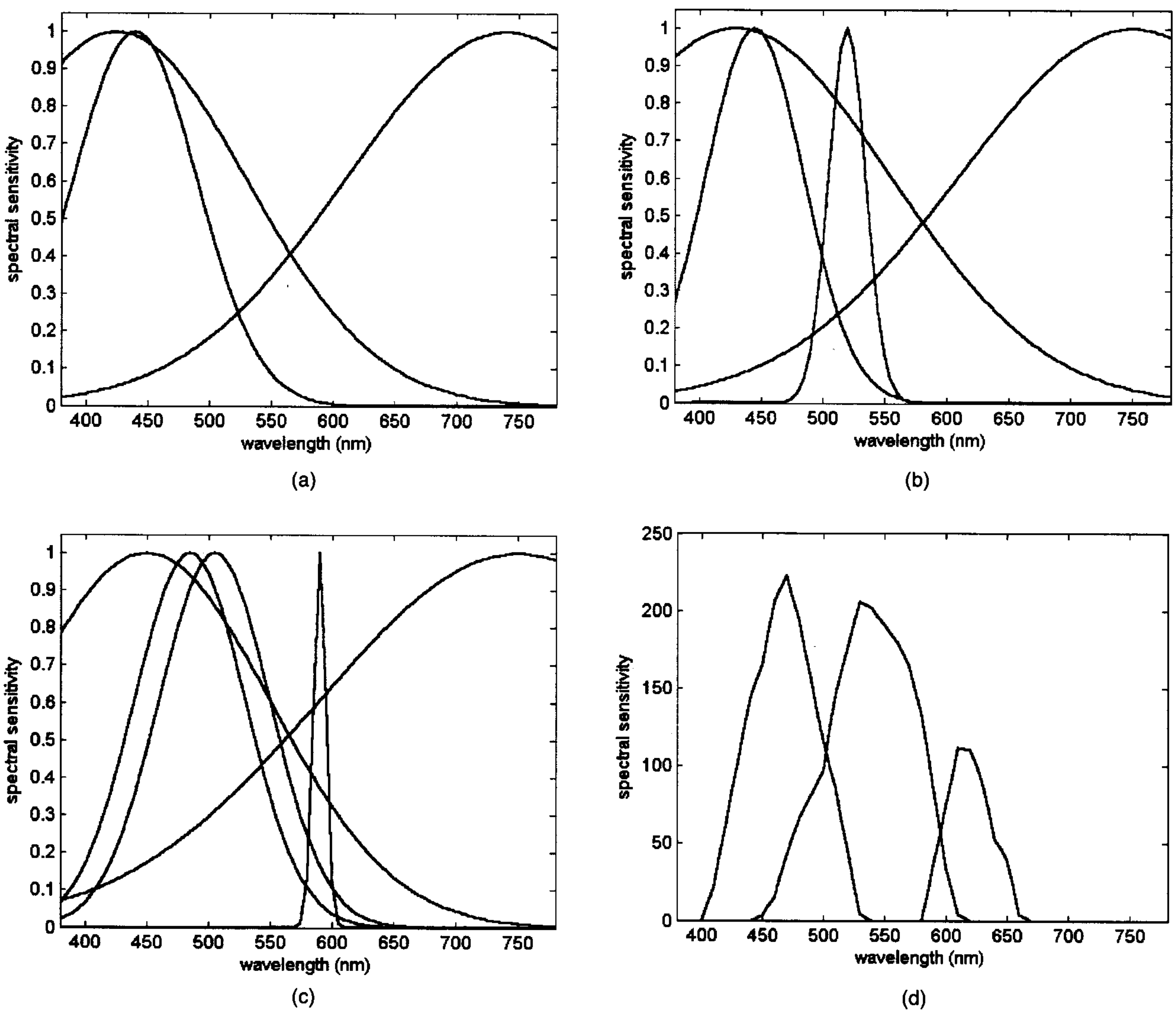

Fig. 2. Optimized spectral sensitivities of Gaussian sensors to recover daylight: (a) three optimum Gaussian sensors, (b) four optimum Gaussian sensors, (c) five optimum Gaussian sensors, (d) spectral response function of the sensors of the JVC TK-1270E CCD color camera.

Table 2. Mean, Standard Deviations, and 99th Percentile Results Obtained with the Optimum Sensors (Table 1) and Our Daylight Eigenvectors (Fig. 1) Tested against a Set of 2600 Daylight Spectra[a]

	GFC		ΔE_{uv}		Integrated Irradiance Error		Fractional Deviation		RMS Error	
Number of Sensors	Mean (SD)	99th Percentile	Mean (SD)	99th Percentile	Mean (SD)	99th Percentile	Mean (SD)	99th Percentile	Mean (SD)	99th Percentile
3	0.9997 (0.0004)	0.9983	0.3360 (0.2983)	1.1855	0.0133% (0.0116%)	0.0572%	0.0173 (0.0092)	0.0497	0.3715 (0.4078)	1.6072
CCD camera	0.9991 (0.0015)	0.9921	0.1967 (0.1740)	0.6825	1.1874% (1.0116%)	5.1053%	0.0298 (0.0198)	0.1083	0.6295 (0.7686)	2.6421
4	0.9998 (0.0003)	0.9990	0.1186 (0.1208)	0.4354	0.0056% (0.0051%)	0.0209%	0.0112 (0.0061)	0.0328	0.2620 (0.3001)	1.086
5	0.9999 (0.0002)	0.9995	0.1072 (0.1180)	0.3972	0.0058% (0.0052%)	0.0222%	0.0076 (0.0044)	0.0226	0.1771 (0.2234)	0.7225

[a] Also included are the results obtained with a JVC TK-1270E CCD color camera. Standard deviations (SD) are given in parentheses.

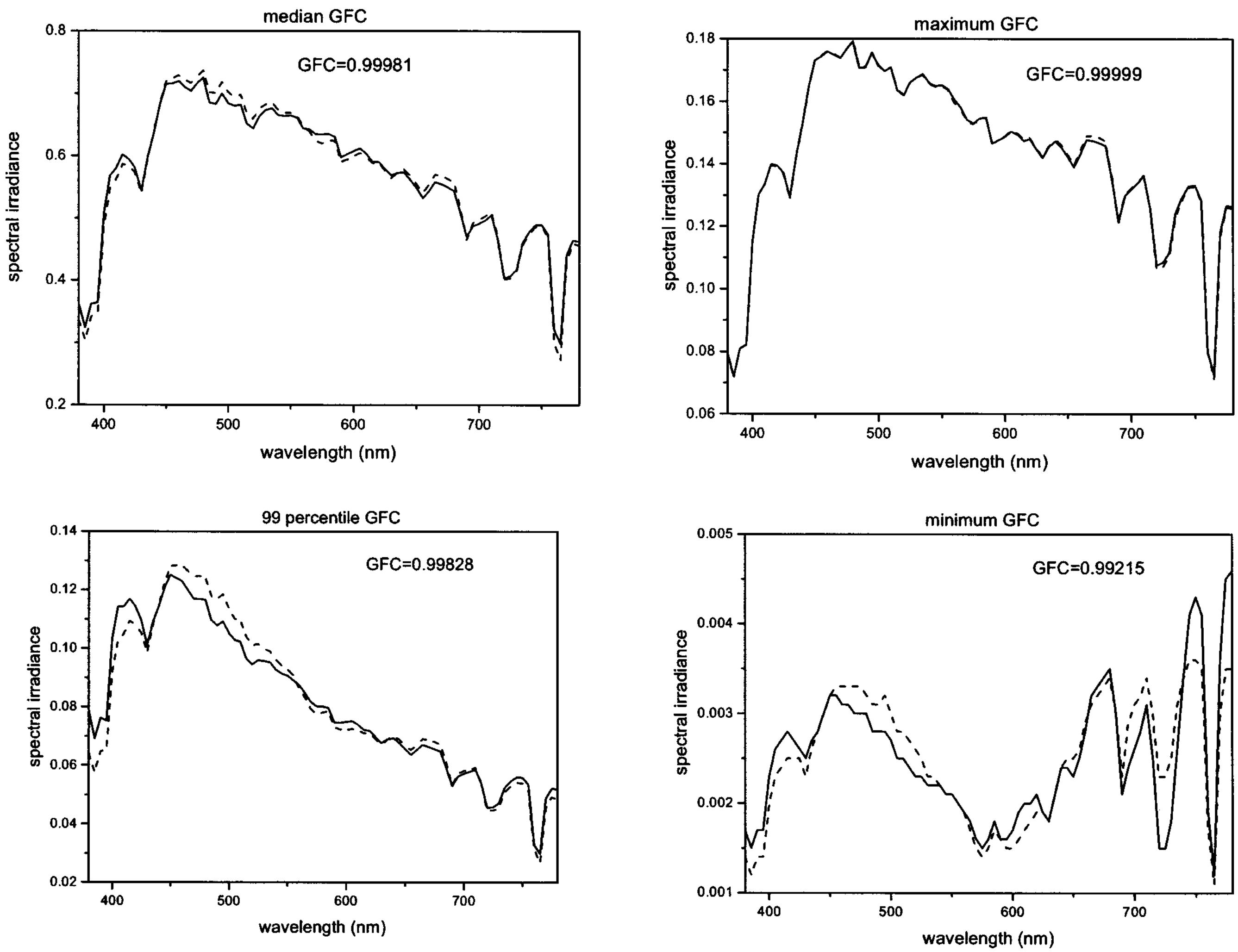

Fig. 3. GFC results obtained with the three optimum Gaussian sensors in Fig. 2(a) versus our set of 2600 daylight spectra. Solid curves, original; dashed curves, recovery.

tion, the spectral sensitivities of the sensors were strikingly similar to those of the human cone fundamentals.[7] Unlike Connah *et al.*,[7] we focus our work on the recovery of daylight, not on spectral reflectances, and, more important, we have used a triple cost function that takes into account a spectral measure (the GFC), the properties of the human visual system (the CIELUV color difference), and the relative error between integrated irradiances. Therefore the nature of these three cost functions biases the spectral profile of our optimum sensors.

The mean, standard deviations, and 99th percentile results of the recoveries are summarized in Table 2, where the fractional deviations are added to help compare our results with those of Michalsky,[2] who found an average fractional deviation of 0.052 when using seven sensors. We have also included the RMS errors and the values for the quality of the estimation when the three nonoptimum color channels (not obtained after an optimization search) of a typical CCD camera [shown in Fig. 2(d)] are considered in the recovery algorithm.

From the results shown in Table 2 we may conclude that even with a set of only three optimum sensors, daylight spectra can be recovered to a very high degree of spectral and colorimetric accuracy: The 99th percentile for the GFC is 0.99828, and the 99th percentile for ΔE_{uv} is 1.1855. We found that the performance of the optimum sensors was in general comparable to that of the eigenvectors[22] and that we got very low values in comparison with any kind of spectral and color tolerances. Moreover, the average fractional deviation with only three optimum sensors (0.0173) was three times lower than that pertaining to Michalsky's results.[2]

Table 2 also indicates that an increase in the number of optimum sensors enhances both the spectral and the colorimetric performance of the recovery algorithm; with five optimum sensors the accuracy is almost perfect, producing an average GFC of 0.9999, an average ΔE_{uv} of 0.1072, and an average fractional deviation of 0.0076. Table 2 shows an unexpected result however: The mean integrated error with five sensors is slightly worse than with four sensors (0.0056% versus 0.0058%). This result can be put down to the cost function, which in our work is a triple cost function, rather than to the optimization search design. So in fact the RMS error (which is directly related to the eigenvector expansion because the eigenvector expansion is the best mean-squared-error approximation) with five sensors is no worse than with four sensors.

When a commercial CCD camera is used, the results are worse, particularly with regard to the integrated irradiance error (99th percentile ~5%), but for colorimetric purposes this CCD camera could be quite suitable as a daylight-estimation device, as found by Chiao *et al.*[36] Figures 3–6 show daylight recoveries with the median,

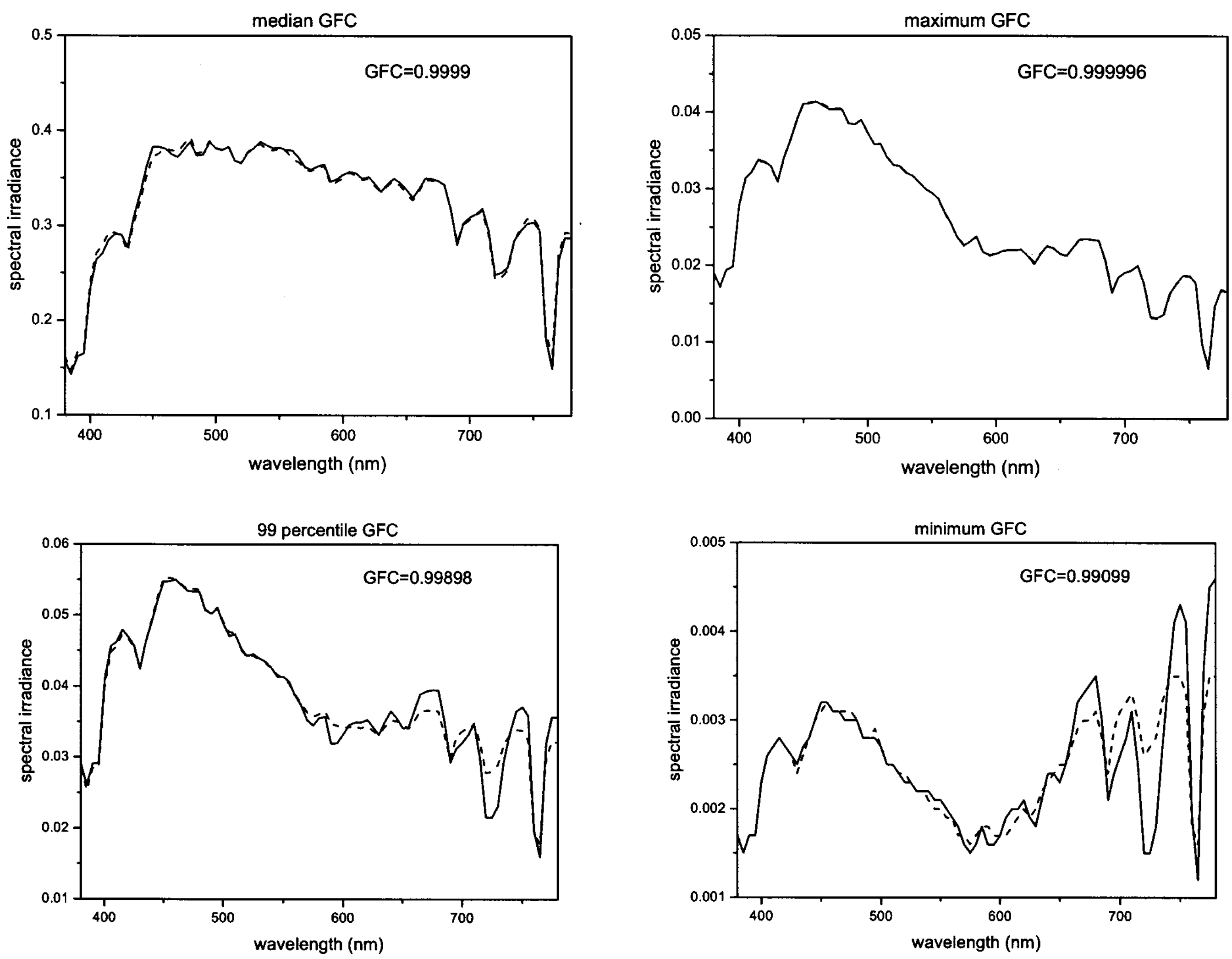

Fig. 4. GFC results obtained with the four optimum Gaussian sensors in Fig. 2(b) versus our set of 2600 daylight spectra. Solid curves, original; dashed curves, recovery.

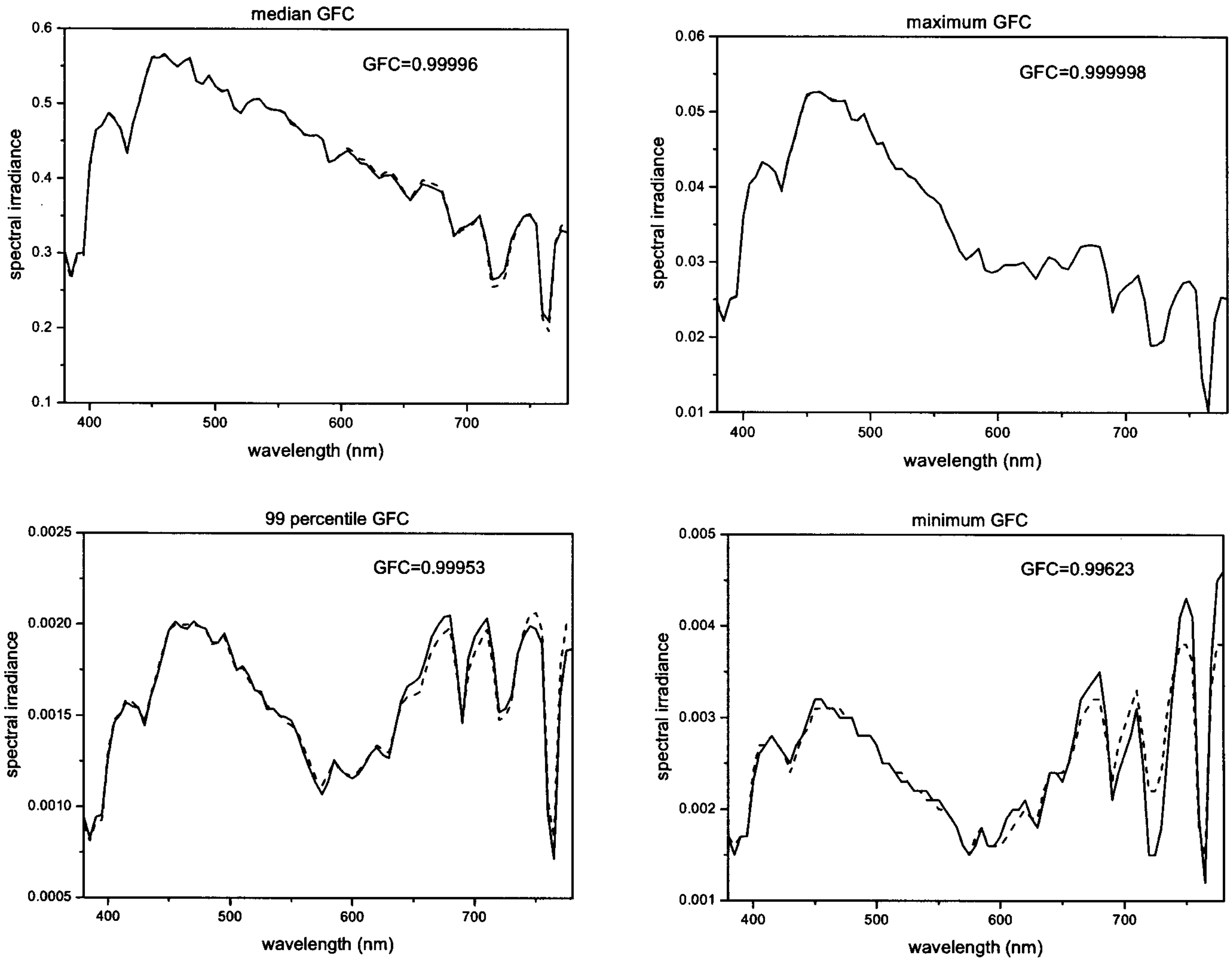

Fig. 5. GFC results obtained with the five optimum Gaussian sensors in Fig. 2(c) versus our set of 2600 daylight spectra. Solid curves, original; dashed curves, recovery.

maximum, minimum, and 99 percentile GFC values obtained with the optimum sensors and the CCD camera.

In Table 3 the results obtained with a theoretical eigenvector expansion [Eq. (4)] can be compared with those obtained with the optimum sensors, and it can be seen how much is lost in using positive sensors instead of the eigenvectors. Although these optimum Gaussian sensors are neither orthogonal nor negative, as are the eigenvectors, the accuracy obtained is comparable.

Our approach is open to the criticism that our conclusions cannot be generalized beyond the collection of spectra analyzed because the matrix $\hat{\Lambda}$, which had to be inverted, depends on both the sensors $R_k(\lambda)$ and the eigenvectors $V_i(\lambda)$. As Connah *et al.* pointed out,[6] the reconstruction errors might be high when the linear system [Eq. (6)] is ill-conditioned, which can be quantified by the condition number of the matrix $\hat{\Lambda}$ calculated by the product of the norms of $\hat{\Lambda}$ and its inverse, $\hat{\Lambda}^{-1}$. Note that a high condition number implies that the coefficients of $\hat{\Lambda}$ are becoming increasingly correlated. Consequently, we extended our analysis to three different collections of natural light to test the influence of the choice of eigenvectors both on the reconstruction errors and on the ill-conditionness or otherwise of the matrix $\hat{\Lambda}$. The three additional sets of natural light were (1) nine typical CIE daylight curves for different color-correlated temperatures (CCTs) from 2000 to 20,000 K at intervals of 2000K,[19] (2) 12 real daylight spectra, measured in the US by R. L. Lee, Jr., and not included in the PCA described in Section 2, and (3) a set of 1567 skylight spectra.[22] This third set was chosen because daylight and skylight spectra usually differ at any given place, as do their chromaticities, but traditionally they have not been distinguished in the confusing terminology used for daylight.

Table 4 shows the recovery errors for the set of nine typical CIE daylight spectra with use of the optimum sensors and our daylight eigenvectors, where it can be seen that both the colorimetric and the spectral results have deteriorated. In a previous paper[20] we analyzed the influence of the daylight basis functions (there are different sets of basis functions published in the literature) on reconstruction quality and found that a spectral resolution of 5 nm, as opposed to 10 nm, improved the quality of the reconstructions. Bearing in mind first that the original

CIE daylight spectra[13] were measured at a resolution of 10 nm (afterward interpolated to a sampling rate of 5 nm), second that typical CIE daylight recoveries do not correspond to real measurements but to a method for calculating the relative SPD of typical daylight phases from CCT values alone, and last that daylight colorimetric and spectral characteristics depend heavily on the particular conditions under which the spectra were acquired, we are hardly surprised by the slightly worse results shown in Table 4.

When using the 12 US real daylight spectra (not included in the PCA) (Table 5) and our daylight basis functions, unsurprisingly we got results very similar to the Granada daylight-recovery results. These very good results indicate that the eigenvectors obtained from our set of 2600 daylight spectra (which provided the broadest range of daylight chromaticities and CCTs to date[18]) are very representative of daylight and highlight the portability of the optimum sensors and our eigenvectors to recover daylight at different sites and for varying atmospheric conditions.

The third additional set of natural-light samples tested consisted of 1567 skylight spectra measured with a field of view of 3° along four sky meridians.[22] The spectral and colorimetric characteristics of skylight differ greatly from those of daylight (which includes sunlight), but unfortunately the tendency in the past has always been to combine heterogeneous collections of curves (see, for example, Judd *et al.*[13]), thus complicating still further the already confusing nature of daylight terminology.[37] This is evident when we recover skylight spectra from the response of our sensors optimized for daylight recovery and the use of our daylight eigenvectors, as shown in Table 6. The results are quite similar to those shown in Table 4 with the CIE daylight spectra, supporting our suspicion that the work of Judd *et al.*[13] included predominantly skylight rather than daylight spectra.

If we use skylight eigenvectors instead of daylight ones to recover skylight with the optimum sensors set out in Table 1, the results (Table 7) improve considerably, becoming almost identical to those shown in Table 2. This happens because the change in the eigenvectors scarcely alters the condition number of the matrix $\hat{\Lambda}$ (e.g., with three optimum sensors it changes from 8.81 to 18.53). If, however, we try to use the CCD camera as a skylight- (6), estimation device with the skylight eigenvectors in Eq.

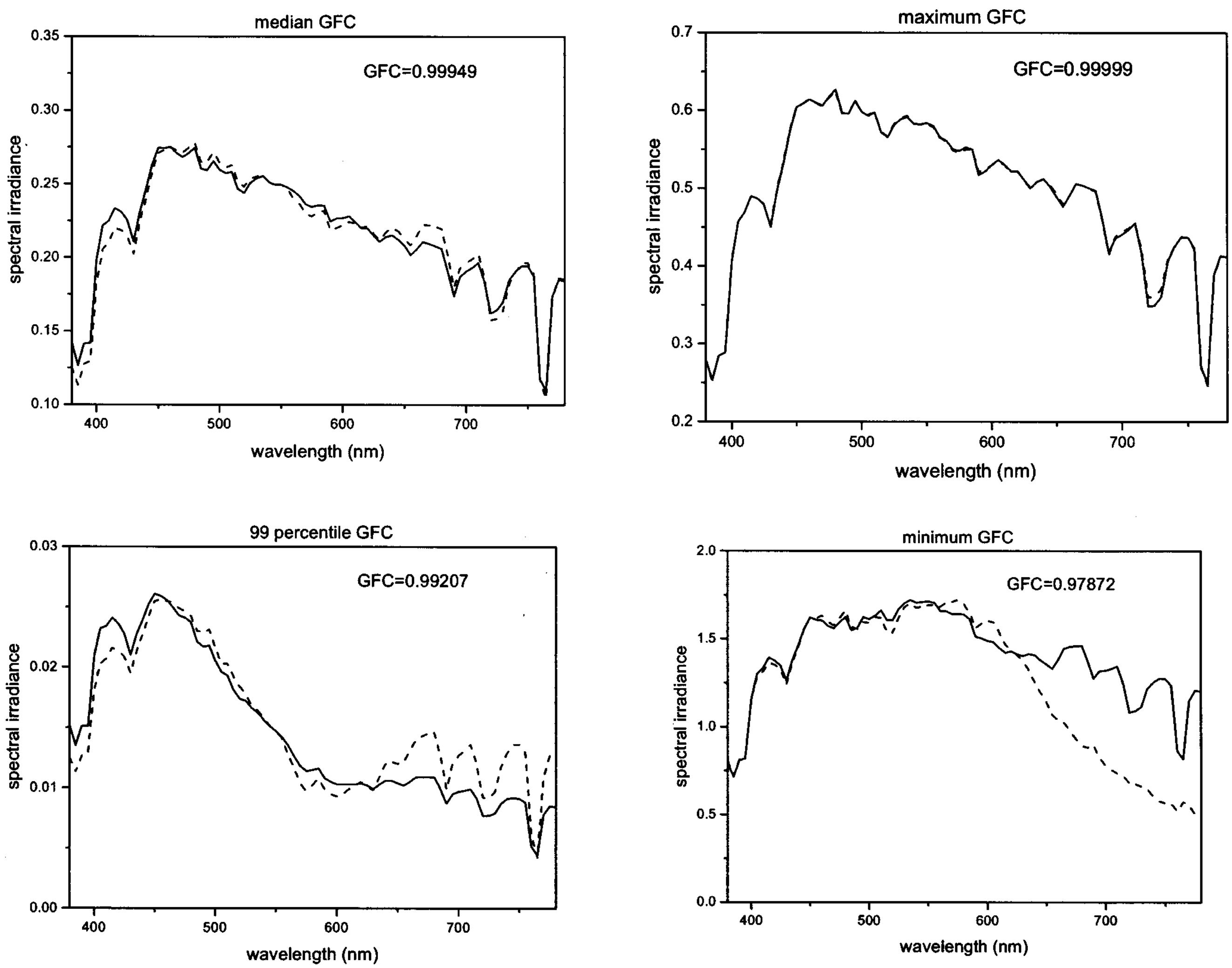

Fig. 6. GFC results obtained with the sensors of a CCD color camera (JVC TK-1270E), shown in Fig. 2(d) versus our set of 2600 daylight spectra. Solid curves, original; dashed curves, recovery.

Table 3. Means and Standard Deviations Obtained with the Optimum Sensors (Table 1) and Our Daylight Eigenvectors (Fig. 1) Tested against a Set of 2600 Daylight Spectra, Compared (in Italics) with the Results Obtained with a Theoretical Eigenvector Expansion [(Eq. 4)][a]

	Mean GFC (SD)	Mean RMS Error (SD)
3 sensors	0.99971 (0.00039)	0.3715 (0.4078)
3 eigenvectors	*0.99973 (0.00038)*	*0.3442 (0.3775)*
4 sensors	0.99983 (0.00030)	0.2620 (0.3001)
4 eigenvectors	*0.99986 (0.00025)*	*0.2503 (0.2868)*
5 sensors	0.99992 (0.00017)	0.1771 (0.2234)
5 eigenvectors	*0.99995 (0.00008)*	*0.1516 (0.1881)*

[a] Standard deviations are given in parentheses.

Table 4. Means and Standard Deviations Obtained with Our Optimum Sensors (Table 1) and Our Daylight Eigenvectors (Fig. 1) Tested against Nine Typical CIE Daylight Spectra of Different CCTs[a]

Number of Sensors	Mean GFC (SD)	Mean ΔE_{uv} (SD)	Mean Integrated Irradiance Error (SD)	Mean Fractional Deviation (SD)
3	0.9946 (0.0033)	24.8328 (16.1775)	0.0416% (0.0074%)	0.0420 (0.0121)
CCD camera	0.9831 (0.0143)	14.8362 (9.0366)	3.8200% (2.1130%)	0.1420 (0.0820)
4	0.9979 (0.0011)	8.3701 (4.1103)	0.0078% (0.0048%)	0.0298 (0.0062)
5	0.9980 (0.0011)	4.9978 (1.9412)	0.0328% (0.0177%)	0.0322 (0.0022)

[a] 4000, 6000, 8000, 10,000, 12,000, 14,000, 16,000, 18,000, and 20,000 K.

the condition number worsens dramatically from 18.16 to 267.88, and this is reflected in results of quite poor quality because most of the skylight recoveries have negative values.

6. CONCLUSIONS

Researchers into artificial-vision algorithms for recognizing and identifying colors have generally assumed that linear models are accurate enough to recover not only spectral reflectances but also spectral illumination. This paper represents an initial step in determining whether the physical implementation of linear models is suitable for estimating daylight spectra from the response of a few carefully chosen Gaussian sensors.

Although both daylight and skylight have complex, uneven spectral profiles with different absorption bands caused by such factors as water vapor, oxygen, ozone, and aerosols, and the strength of these bands depends on the day or even the time of the day, we have found that a linear recovery algorithm with a set of a few optimum Gaussian sensors returns very-high-quality reconstructions of both daylight and skylight. We have described here a procedure to find the best set of Gaussian sensors with optimum spectral position and bandwidth.

Although these optimum Gaussian sensors are not orthogonal, as are the eigenvectors, daylight spectra can be recovered from the response of these sensors with a very high degree of spectral and colorimetric accuracy. In fact, the performance of the optimal sensors was in general comparable to that of the eigenvectors. Increasing the number of optimum sensors enhances both the spectral and the colorimetric performance of the recovery algorithm. Our results show that the combination of linear models with optimal Gaussian sensors is an accurate

Table 5. Means and Standard Deviations Obtained with Our Optimum Sensors (Table 1) and Our Daylight Eigenvectors (Fig. 1) versus 12 Daylight Spectra Measured in the US by Lee[37] [a]

Number of Sensors	Mean GFC (SD)	Mean ΔE_{uv} (SD)	Mean Integrated Irradiance Error (SD)	Mean Fractional Deviation (SD)
3	0.9992 (0.0006)	0.2847 (0.2141)	0.0323% (0.0319%)	0.0296 (0.0145)
CCD camera	0.9967 (0.0054)	0.1776 (0.1512)	2.2306% (2.9605%)	0.0501 (0.0472)
4	0.9995 (0.0003)	0.1136 (0.0818)	0.0110% (0.0060%)	0.0208 (0.0066)
5	0.9996 (0.0002)	0.1531 (0.1364)	0.0060% (0.0061%)	0.0181 (0.0061)

[a] These 12 spectral daylight measurements have different CCTs from 5750 to 16,780 K and a sampling rate of 5 nm. Standard deviations are given in parentheses.

Table 6. Means and Standard Deviations Obtained with Our Optimum Sensors (Table 1) and Daylight Eigenvectors (Fig. 1) versus Our Set of 1567 Skylight Spectra with a Sampling Rate of 5 nm[a]

Number of Sensors	Mean GFC (SD)	Mean ΔE_{uv} (SD)	Mean Integrated Irradiance Error (SD)	Mean Fractional Deviation (SD)
3	0.9959 (0.0037)	0.6224 (0.4208)	0.0669% (0.0562%)	0.0696 (0.0332)
CCD camera	0.9854 (0.0131)	0.3459 (0.2409)	3.9113% (3.3562%)	0.1375 (0.0753)
4	0.9981 (0.0017)	0.2192 (0.1510)	0.0134% (0.0111%)	0.0474 (0.0249)
5	0.9985 (0.0016)	0.1501 (0.1278)	0.0212% (0.0173%)	0.0413 (0.0235)

[a] For more details see Ref. 22.

Table 7. Means and Standard Deviations Obtained with Our Optimum Sensors (Table 1) and Our Skylight Eigenvectors (see Ref. 22) versus Our Set of 1567 Skylight Spectra with a Sampling Rate of 5 nm

Number of Sensors	Mean GFC (SD)	Mean ΔE_{uv} (SD)	Mean Integrated Irradiance Error (SD)	Mean Fractional Deviation (SD)
3	0.9979 (0.0049)	0.3898 (0.8592)	0.0701% (0.0640%)	0.0403 (0.0303)
CCD camera	0.7800 (0.2323)	2.0551 (8.5778)	14.1358% (12.4656%)	0.5711 (0.4939)
4	0.9994 (0.0008)	0.1812 (0.2821)	0.0096% (0.0078%)	0.0240 (0.0157)
5	0.9999 (0.0001)	0.0513 (0.0996)	0.0061% (0.0049%)	0.0099 (0.0052)

method to recover spectral daylight because of its precision, efficiency, and portability.

Our intention for the future is to extend this work to different kinds of sensor, to analyze the effects of reducing the constraints assumed in Section 4 (i.e., to relax the constraint of an equal number of sensors as basis functions), and to consider the presence of random noise and quantization noise.

ACKNOWLEDGMENTS

The authors thank Raymond L. Lee, Jr., who kindly provided us with some US daylight spectral measurements. Thanks also go to A. L. Tate for revising our English text and to our Ph.D. student Miguel Ángel López for his insightful comments. This work was supported by the Comisión Interministerial de Ciencia y Tecnología (CICYT), Ministerio de Educación y Ciencia, Spain, under research grant BMF2000-1473.

Correspondence should be addressed to Javier Hernández-Andrés by e-mail, javierha@ugr.es.

REFERENCES AND NOTES

1. J. J. Michalsky and E. W. Kleckner, "Estimation of continuous solar spectral distributions from discrete filter measurements," Sol. Energy **33**, 57–64 (1984).
2. J. J. Michalsky, "Estimation of continuous solar spectral distributions from discrete filter measurements: II. A demonstration of practicability," Sol. Energy **34**, 439–445 (1985).
3. S. K. Park and F. O. Huck, "Estimation of spectral reflectance curves from multispectral image data," Appl. Opt. **16**, 3107–3114 (1977).
4. D. H. Brainard and W. T. Freeman, "Bayesian color constancy," J. Opt. Soc. Am. A **14**, 1393–1411 (1997).
5. A. García-Beltrán, J. L. Nieves, J. Hernández-Andrés, and J. Romero, "Spectral reflectance recovering through linear model and narrow-band filters," in *Digest of 1995 OSA Annual Meeting* (Optical Society of America, Washington, D.C., 1995), p. 132.
6. D. Connah, S. Westland, and M. G. A. Thomson, "Recovering spectral information using digital camera systems," Color Technol. **117**, 309–312 (2001).
7. D. Connah, S. Westland, and M. G. A. Thomson, "Optimization of a multispectral imaging system," in *Proceedings of the First European Conference on Colour in Graphics, Imaging and Vision* (Society for Imaging Science and Technology, Springfield, Va., 2002), pp. 619–622.
8. D. Connah, S. Westland, and M. G. A. Thomson, "Parametric investigation of Multispectral imaging," in *9th Congress of the International Color Association*, R. Chung and A. Rodrigues, eds., Proc. SPIE **4421**, 943–946 (2001).
9. L. T. Maloney and B. A. Wandell, "Color constancy: a method for recovering surface spectral reflectance," J. Opt. Soc. Am. A **3**, 29–33 (1986).
10. L. T. Maloney, "Physics-based approaches to modeling surface color perception," in *Color Vision: From Genes to Perception*, K. R. Gegenfurtner and L. T. Sharpe, eds. (Cambridge U. Press, New York, 1999), pp. 393–398.
11. D. H. Marimont and B. A. Wandell, "Linear models of surface and illuminant spectra," J. Opt. Soc. Am. A **9**, 1905–1913 (1992).
12. L. T. Maloney, "Surface color perception and environmental constraints," in *Colour Vision: From Light to Object*, R. Mausfeld and D. Heyer, eds. (Oxford U. Press, Oxford, UK, to be published).
13. D. B. Judd, D. L. MacAdam, and G. Wyszecki, "Spectral distribution of typical daylight as a function of correlated colour temperature," J. Opt. Soc. Am. **54**, 1031–1041 (1964).
14. G. T. Winch, M. C. Boshoff, C. J. Kok, and A. G. du Toit, "Spectroradiometric and colorimetric characteristics of daylight in the southern hemisphere: Pretoria, South Africa," J. Opt. Soc. Am. **56**, 456–464 (1966).
15. V. D. P. Sastri and S. R. Das, "Typical spectral distributions and colour for tropical daylight," J. Opt. Soc. Am. **58**, 391–398 (1968).
16. A. W. S. Tarrant, "The spectral power distribution of daylight," Trans. Illum. Eng. Soc. **33**, 75–82 (1968).
17. E. R. Dixon, "Spectral distribution of Australian daylight," J. Opt. Soc. Am. **68**, 437–450 (1978).
18. 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).
19. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae* (Wiley, New York, 1982), pp. 144–146.
20. J. Hernández-Andrés, J. Romero, A. García-Beltrán, and J. L. Nieves, "Testing linear models on spectral daylight measurements," Appl. Opt. **37**, 971–977 (1998).
21. D. Slater and G. Healey, "Analyzing the spectral dimensionality of outdoor visible and near-infrared illumination functions," J. Opt. Soc. Am. A **15**, 2913–2920 (1998).
22. J. Hernández-Andrés, J. Romero, and R. L. Lee, Jr., "Colorimetric and spectroradiometric characteristics of narrow-field-of-view clear skylight in Granada, Spain," J. Opt. Soc. Am. A **18**, 412–420 (2001).
23. J. Romero, A. García-Beltrán, and J. Hernández-Andrés, "Linear bases for representation of natural and artificial illuminants," J. Opt. Soc. Am. A **14**, 1007–1014 (1997).
24. J. M. DiCarlo and B. A. Wandell, "Illuminant estimation: beyond the bases," in *Proceedings of Eighth Color Imaging Conference: Color Science, Systems, and Applications* (Society for Imaging Science and Technology, Springfield, Va., 2000), pp. 91–96.
25. J. Y. Hardeberg, "Acquisition and reproduction of color images: colorimetric and multispectral approaches" (Dissertation.com, Parkland, Fla., 2001). (Revised second edition of Ph.D. dissertation, Ecole Nationale Supérieure des Télécommunications, Paris, 1999), pp. 157–174.
26. J. Y. Hardeberg, "Filter selection for multispectral color image adquisition," in *Proceedings of PICS 2003* (Society for Imaging Science and Technology, Springfield, Va., 2003), pp. 177–182.
27. R. Piché, "Nonnegative color spectrum analysis filters from principal component analysis characteristic spectra," J. Opt. Soc. Am. A **19**, 1946–1951 (2002).
28. R. Lenz, M. Österberg, J. Hiltunen, T. Jaaskelainen, and J. Parkinnen, "Unsupervised filtering of color spectra," J. Opt. Soc. Am. A **13**, 1315–1324 (1996).

29. M. G. A. Thomson and S. Westland, “Color-imager calibration by parametric fitting of sensor responses,” Color Res. Appl. **26**, 442–449 (2001).
30. F. H. Imai, M. R. Rosen, and R. S. Berns, “Comparative study of metrics for spectral match quality,” in *Proceedings of the First European Conference on Colour in Graphics, Imaging and Vision* (Society for Imaging Science and Technology, Springfield, Va., 2002), pp. 492–496.
31. Note that the GFC is the multiple correlation coefficient R and the square root of the variance-accounted-for coefficient. Note that GFC ranges from 0 to 1, where 1 indicates a perfect reconstruction.
32. See Ref. 19, pp. 306–310 and 828–829.
33. ASTM Committee E-12, “Standard practice for computing the colors of objects by using the CIE system (E 308-95), in ASTM standards on color and appearance measurements” (American Society for Testing and Materials, Philadelphia, Pa., 1996), pp. 262–263.
34. S. Quan and N. Ohta, “Evaluating hypothetical spectral sensitivities with quality factors,” J. Imaging Sci. Technol. **46**, 8–14 (2002).
35. S. Quan, N. Ohta, R. S. Berns, and X. Jiang, “Unified measure of goodness and optimal design of spectral sensitivity functions,” J. Imaging Sci. Technol. **46**, 485–497 (2002).
36. C.-C. Chiao, D. Osorio, M. Vorobyev, and T. W. Cronin, “Characterization of natural illuminants in forests and the use of digital video data to reconstruct illuminant spectra,” J. Opt. Soc. Am. A **17**, 1713–1721 (2000).
37. R. L. Lee, Jr., “Twilight and daytime colors of the clear sky,” Appl. Opt. **33**, 4629–4638, 4959 (1994).