Optimal sensors for recovering skylight in the presence of noise

M.A. López-Álvarez[1], J. Hernández-Andrés[1], R.L. Lee Jr[2]. and J. Romero[1]
[1]*Departamento de Óptica, Facultad de Ciencias, Universidad de Granada, Granada (SPAIN)*
[2]*Physics Department, U. S. Naval Academy, Annapolis, USA*
Corresponding author: M.A. López-Álvarez (migangel@ugr.es)

ABSTRACT

In a previous work [J. Opt. Soc. Am. A **21**, 13-23 (2004)] it was shown that a combination of linear models with optimum Gaussian sensors obtained by an exhaustive search provided a theoretical algorithm to recover daylight spectra accurately. Here we improve the simulation of such a multispectral device by considering the different kinds of noise that are always present in electronic devices such as CCDs, and extend our research to a different kind of natural illumination, skylight. We opted for a simulated annealing algorithm to obtain the optimum sensors because of its fastness, which requires the minimization of a single cost function, so we develop a new one that calculates both the spectral and colorimetric similarity of any pair of skylight spectra. Our technique lets us study the properties of optimal sensors in the presence of noise, one side effect of which is that adding more sensors may not improve the spectral recovery.

1. INTRODUCTION

The profit of low-cost multispectral imaging systems in estimating spectral power distributions has been widely studied[1-4]. Because researchers could benefit from high-resolution angular maps of skylight's spectral power distribution (SPD), we need to measure many skylight spectra simultaneously across the sky dome. Multispectral imaging systems let us do so.

There are various mathematical methods[5-7] (like PCA, Wiener estimation method, spline interpolation, MDST, non-negative matrix factorization or direct transformations) which permit reconstruct accurately a spectrum from the response of a small set of sensors. This is a classic inverse problem that requires a mathematical estimation algorithm. We opt here for a linear model based on a principal components analysis (PCA).

An important issue in this task is the influence of noise[1,7,8], its propagation through the mathematical transformations and how the selection of the sensors of the multispectral system may affect the accuracy of skylight spectral recovery. To include all these factors in an exhaustive search is a highly demanding computational task. Our alternative approach greatly reduces computing time with a simulated annealing algorithm[9] that minimizes a single cost function. To this end, we propose a cost function that evaluates the quality of our recovered skylight spectra from the colorimetric and spectral points of view: the colorimetric and spectral combined metric, or CSCM function.

Finally, we present the results of spectral recoveries for two different sets of skylight spectra. One dataset was measured in Granada, Spain and was used as a training set both in the PCA and in the simulated annealing algorithm. The second dataset was measured in Owings, Maryland, and was used to test the accuracy of our simulated spectral imaging system.

2. METHOD

Our aim here is to find the spectral properties (peak sensitivities from 380-780 nm in 5nm steps, and full width at half-maximum, FWHM, from 10-300 nm in 10-nm steps) of a small number of Gaussian sensors (three to five), similar to commercial ones, which give the optimum spectral recovery of skylight in the presence of thermal noise (measured via the signal to noise ratio, SNR, in three cases: 40, 30 and 26dB) and quantization noise (simulating the analog to digital conversion with 8, 10 and 12 bits) by using a linear reconstruction based on PCA eigenvectors (also three to five) obtained from a set of 1567 skylight spectral measurements in Granada (Spain)[10].

No consensus exists about what "optimum" means in such a system. For us, one set of sensors is clearly better than another if its reconstructed skylight spectra are more accurate when tested by some metric. The key question is what metric to use in order to assess color reconstruction from both colorimetric and spectral standpoints. Day[8] used thresholds for RMSE and CIEDE2000 metrics when searching for optimum sensors; Hernández-Andrés *et al.*[11] used GFC, CIELUV, and IIE(%) (integrated irradiance error, a metric widely used in solar radiation measurements) in a similar way.

We must use a single cost function when developing a simulated annealing algorithm, so we actually construct a simple single-cost function or metric that combines several metrics at once. We use GFC as a spectral metric, CIELAB as a colorimetric cost function, and IIE(%) as a metric for comparing the spectral curves of natural illuminants. In principle, our new metric should approach zero for near-perfect matches (in contrast with GFC, which tends to unity for perfect matches) and give approximately the same weight to the GFC, CIELAB, and IIE(%) metrics. Our combined CSCM metric is calculated by:

$$CSCM = Ln(1 + 1000(1 - GFC)) + \Delta E_{ab} + IIE(\%) \tag{1}$$

Equation (1) is a combined metric that is zero for perfect matches, and thus is a good candidate for developing an annealing search algorithm. Its chief advantage is that it quantifies spectral mismatches among metamers, perceptual differences in color matches, and differences in such integrated radiometric quantities as irradiance or radiance. This metric clearly combines the properties of various metrics relevant to skylight spectra.

Simulated annealing algorithms have been widely used as search algorithms in physics[9] and in the design of multispectral imaging systems[2], but typically they evaluate only a single metric such as CIELAB ΔE_{ab} or spectral RMSE. Such algorithms are based on simulating the process of annealing (slow cooling after heating) of a thermodynamic system (e.g., a gas) that is always in thermal equilibrium. The algorithm searches for the minimum energy state when temperature decreases with time. We must construct a rule for changing the existing state, calculate its energy, and accept it as the system state with a probability given by Bolztmann's factor. In this case, energy is replaced with our CSCM cost function, which makes clear why the CSCM must be a single function that equals zero for perfect matches. The state is represented by a given set of sensors (the peak sensitivities and FWHM of which determine the energy of the state), and the rule of state-changing will statistically favor those states whose energy is close to that of the current state.

3. RESULTS

We found (see fig. 1) that, for a given number of sensors, the peak sensitivities and FWHM are almost the same in every noise situation and for any number of eigenvectors (from three to five) used in the linear reconstruction, which is desirable for constructing a practical multispectral system from this theoretical study. As other authors have noted[1,3], sensors tend to sharpen when noise is high (i.e., low SNR). This occurs because sharper sensors make the matrix to be inverted more robust to noise by decreasing its condition number. Not surprisingly, the curves also sharpen as we increase their number (i.e., as we approach a narrow-band hyperspectral imaging system).

The optimal number of eigenvectors used in the linear reconstruction depends on the noise situation and the number of sensors. As we can see in table 1, for low noise (high SNR) it is better to equal the number of eigenvectors to the number of sensors and there is an improvement in increasing the number of sensors. For high noise (low SNR) it is better to use always five eigenvectors and there is no improvement in using more sensors, because adding more sensors simply increases system's noise. This can be appreciated by noting that a system with more sensors likely has more connections, more transistors in the CCD, and more memory cells. All these elements individually contribute noise, and so each raises the total noise level.

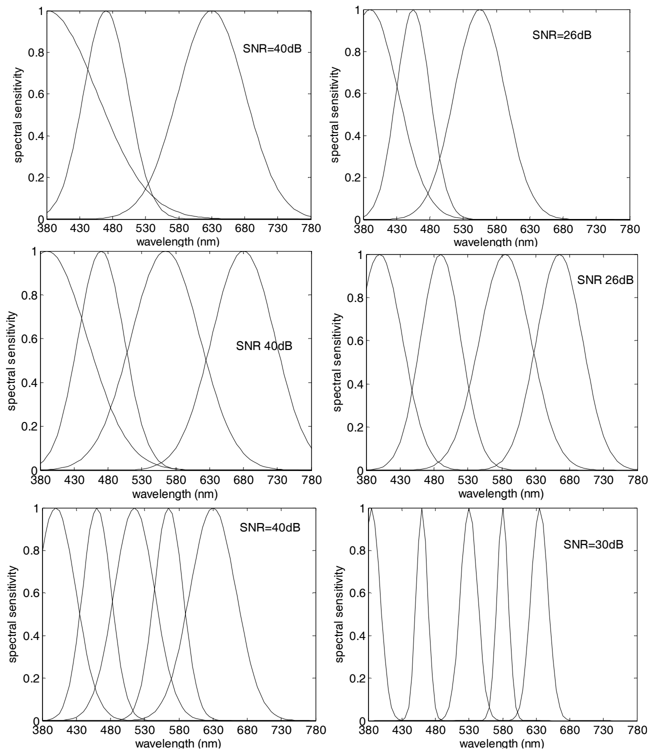

Fig.1 Best 3, 4 and 5 sensors (and equal number of eigenvectors in each case) for various SNR.

Table 1. Means for the 1567 skylight spectra used in PCA with 3 to 5 sensors for various SNR levels and 12-bit quantization.

SNR	sensors	eigenvectors	GFC±SD	CIELAB ΔE_{ab}±SD	IIE(%)±SD	CSCM±SD
40dB	3	3	0.9993±0.0012	0.7±0.5	1.3±0.7	2.4±1.1
	3	5	0.9988±0.0023	1.0±0.6	0.9±0.6	2.5±1.2
	4	3	0.9994±0.0011	0.8±0.6	1.2±0.7	2.4±1.2
	4	4	0.9997±0.0003	0.6±0.3	1.2±0.5	2.0±0.7
	5	3	0.9993±0.0012	0.8±0.5	1.1±0.7	2.3±1.2
	5	5	0.9998±0.0002	0.3±0.2	1.1±0.8	1.6±0.8
30dB	3	3	0.9987±0.0016	0.9±0.5	3.2±1.9	4.8±2.2
	3	5	0.9940±0.0070	1.6±0.7	1.7±1.4	4.9±2.1
	4	3	0.9990±0.0017	0.9±0.4	3.5±2.1	4.9±2.4
	4	4	0.9991±0.0008	0.9±0.5	3.5±1.4	4.9±1.8
	5	3	0.9985±0.0017	1.1±0.6	2.8±1.6	4.7±2.0
	5	5	0.9991±0.0009	0.8±0.4	3.3±2.1	4.6±2.3
26dB	3	3	0.9981±0.0023	1.3±0.7	5.0±3.5	7.3±4.1
	3	5	0.9870±0.0130	2.0±0.7	2.4±2.1	6.9±2.8
	4	3	0.9986±0.0017	1.2±0.6	5.9±2.9	7.9±3.2
	4	5	0.9902±0.0014	2.0±0.8	2.7±2.2	6.9±3.1
	5	3	0.9976±0.0025	1.4±0.7	5.7±3.1	8.2±3.5
	5	5	0.9988±0.0012	1.1±0.5	6.1±2.6	7.9±3.1

Naturally our system must work for spectra other than those used to calculate the original dataset's eigenvector matrix V. Thus we extend our analysis to skylight spectra measured at another site in Owings, Maryland. Using 12-bit quantization and the best sensors for each case (see Table 1), we analyze metrics for these new spectra in Table 2. It too shows small errors for the recovered spectra, demonstrating that both our spectral recovery method and optimum sensors can be used to develop a reliable system for imaging skylight spectra.

Table 2. Means for 242 skylight spectra measured in Owings recovered with the best set of sensors at 12-bit quantization.

SNR	Number of sensors	Number of eigenvectors	GFC±SD	CIELAB ΔE_{ab}±SD	IIE(%)±SD	CSCM±SD
40dB	5	5	0.9990±0.0002	0.2±0.1	0.5±0.4	1.5±0.5
30dB	5	5	0.9984±0.0009	0.7±0.3	1.7±1.3	3.3±1.6
26dB	3	5	0.9902±0.0021	2.1±0.5	1.5±1.1	5.9±1.4

4. CONCLUSIONS

We have shown that linear methods based on PCA allow accurate recovery of skylight spectra from broadband camera sensors. We have presented a simulated annealing algorithm, using CSCM as a single cost function, as a fast method for searching a limited number of Gaussian sensors to construct an optimum multispectral imaging system. We have simulated some common noise sources present in any digital imaging system in order to mimic noise in real images. We have shown that increasing the number of sensors does not necessarily improve the accuracy of recovered spectra if sensor noise is high because each sensor's individual noise contributions degrade the overall quality of the spectral reconstruction. Thus the optimum number of sensors depends on noise levels inherent to the given multispectral system.

Acknowledgments
López-Álvarez thanks Consejería de Innovación y Desarrollo Tecnológico de la Junta de Andalucía for the grant that supports his investigation.

References

1. S. Noriyuki, "Optimal spectral sensitivities of a color image acquisition device in the presence of noise," in *Proceedings of the 2nd European Conference on Colour Graphics, Imaging and Vision* (Society for Imaging Science and Technology, Springfield, Va., 2004), pp. 379-383.
2. D. Connah, S. Westland, and M. G. A. Thomson, "Optimization of a multispectral imaging system," in *Proceedings of the 1st European Conference on Colour Graphics, Image and Vision* (Society for Imaging Science and Technology, Springfield, Va., 2002), pp. 619-622.
3. D. Connah, S. Westland, and M. G. A. Thomson, "Recovering spectral information using digital camera systems," Color Technol. **117**, 309-312 (2001).
4. J. Y. Hardeberg (2001), "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.
5. 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).
6. G. Buchsbaum and O. Bloch, "Color categories revealed by non-negative matrix factorization of Munsell color spectra," Vis. Res. **42**, 559-563 (2002).
7. P. D. Burns, "Analysis of image noise in multispectral color acquisition," Ph. D. Dissertation, Rochester Institute of Technology (1997).
8. D. C. Day, "Filter selection for spectral estimation using a trichromatic camera," Ph. D. Dissertation, Rochester Institute of Technology (2003).
9. N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equations of state calculations by fast computing machines," J. Chem. Phys. **21**, 1087-1092 (1953).
10. 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).
11. J. Hernández-Andrés, J. L. Nieves, E. M. Valero, and J. Romero, "Spectral-daylight recovery by use of only a few sensors," J. Opt. Soc. Am. A 21, 13-23 (2004).