Spectral recovery of artificial illuminants using a CCD colour camera with Non-negative Matrix Factorization and Independent Component Analysis

Juan L. Nieves, Eva M. Valero, Javier Romero, Javier Henández-Andrés; Department of Optics. University of Granada; Granada (SPAIN)

Abstract

We investigated the quality of the spectral estimation of incandescent and fluorescent illuminants using Non-negative Matrix Factorization (NMF), Independent Component Analysis (ICA) and a direct pseudo-inverse approach. We simulated the response of a commercial digital CCD camera coupled or not with coloured filters to a set of natural and artificial illuminants. None of the recovery algorithms used here needed information about spectral sensitivities of the camera sensors or eigenvectors to estimate the spectral power distributions of illuminants. Although nonnegative algorithms can reduce the computational cost of spectral devices, experiments show that ICA and direct pseudo-inverse methods consistently outperforms the NMF approaches, even for fluorescent lights and even using a reduced training set of illuminants.

1. INTRODUCTION

Multispectral analysis and synthesis of the spectral power distribution (SPD) of illuminants and spectral reflectances has been explored intensively during the last few years.[1]-[6] Different computational approaches have been introduced to improve both the spectral and colorimetric quality of spectral recovery and to reduce the number of components which are needed to recover the computed spectra.[4]-[5] In a previous work it was found that it is possible to recover daylight spectra with high spectral and colorimetric accuracy with a reduced number of 3 to 9 spectral bands with very promising results when a direct pseudo-inverse transformation or *direct-mapping* is established between spectra and digital counts.[6] The great advantage of this method by comparison with other algorithms using CCD cameras [7] is that it does not need information about spectral sensitivities of the camera sensors nor eigenvector analysis. In that work the training set of illuminants was clearly dominated by daylight spectra and thus spiky SPDs was avoided in the analysis. These spiky profiles are evident in fluorescent-type illuminants and cannot be represented adequately with linear models of reduced number of parameters. In addition, different approaches have been proposed for classifying fluorescent scene illumination instead of trying to recover their spectral profile.[8]-[9]

The aim of the present work was to investigate the quality of the spectral estimation of fluorescent illuminants using a CCD colour camera. Recent approaches based on Independent Component Analysis (ICA) and non-negative linear modelling [10] have proven of particular interest to design physically realizable sensors for recovering spectral functions. These algorithms are used to derive appropriate basis for identifying the pseudo-inverse of the basis vectors with the optical sensors to be used by a spectral device. We simulated the response of a commercial digital CCD camera coupled or not with coloured filters to a set of natural and artificial illuminants, and used different Non-negative Matrix Factorization (NMF) algorithms, ICA and the direct pseudo-inverse method mentioned above to recover the SPD of the illuminants from the camera responses.

2. METHODS

We simulated the responses of a digital CCD colour camera from QImaging (model Retiga 1300, QImaging Corp., Canada) with 12 bits intensity resolution per channel coupled with different coloured filters. The coloured filters were the colour glass filter OG550 and RG630 from OWIS GMBH once we had tested other filter combinations. Different sets of illuminants were used for the training and testing of the different algorithms to obtain the transformation between camera responses and the SPD of the illuminants: the training set was formed from a reduced number of t= 82 SPDs of daylights and fluorescent illuminants; and the test sets comprised a set of m=20 commercial fluorescent-type illuminants.[11] The camera was assumed to point to a uniform white reference surface with spectral reflectance function $r_W(\lambda)$. The response of the kth sensor ρ_k is,

$$\rho_k = \sum_{\lambda=400}^{700} E(\lambda) r_W(\lambda) Q_k(\lambda) \qquad (1)$$

where $Q_k(\lambda)$ is the spectral sensitivity of the kth sensor and $E(\lambda)$ is the SPD of the illuminant impinging on the white reference surface (chip number 19 from the GretagMacBeth ColorChecker). The camera responses were obtained with a single set of k=3 sensors, the three native RGB channels of the camera, a set of k=6 sensors, with three filtered RGB channels, and a set of k=9 sensors with six filtered RGB.

2.1. NMF and ICA algorithms

Given a non-negative data matrix **E** (an $n \times m$ matrix), non-negative matrix factorization (NMF) finds an approximate factorization into two non-negative matrix factors **W** (an $n \times p$ matrix of basis vectors) and **H** (a $p \times m$ matrix of p coefficient vectors), where p is a smaller number compared to n and m.[12] In this work, the data matrix is a set of unknown illuminants **E** (an $n \times m$ matrix of m illuminant spectra sampled at n wavelengths) which will be derived using a relationship involving the coefficients vector within each column of **H**,

$$\mathbf{E} = \mathbf{W}\,\mathbf{H} \qquad (2)$$

Since the intent of the spectral recovery is to estimate illuminant spectra from the responses of a CCD colour camera, we first computed a set of sensors outputs ρ_o (a $k \times t$ matrix of t training spectra which are captured by k sensors) and coefficient matrix $\mathbf{H_o}$ from a training set of illuminants. Thus, given a data matrix of unknown illuminants $\mathbf{E}$, whose sensor outputs are ρ(a $k \times m$ matrix), the corresponding coefficient matrix $\mathbf{H}$ is computed from the training set as,

$$\mathbf{H} = \left(\mathbf{H}_o\ \boldsymbol{\rho}_o^+\right)\boldsymbol{\rho} \tag{3}$$

where ρ_o+ is the pseudo-inverse matrix of ρ_o; next, the equation (2) is applied to recover the estimated spectra. The rank of factorization p can be adjusted depending on the input data matrix in order to reduce the computational cost of spectral estimation.

The ICA algorithm approximates data using a similar decomposition as (2) and finds basis vectors that are uncorrelated and also independent but not necessarily orthogonal. In this paper, we have used the ICA by Hyvarinen [13], and two NMF algorithms which use two different error functions for the optimal choice of $\mathbf{W}$ and $\mathbf{H}$; the NMF algorithms were the Euclidean and the divergence updates by Lee and Seung.[14]

2.2. Direct pseudo-inverse transformation

This method is also based on a direct transformation between the estimated illuminant spectra $\mathbf{E}$ (an $n \times m$ matrix) and sensor responses ρ(an $k \times m$ matrix) expressed by,

$$\mathbf{E} = \mathbf{F}\ \boldsymbol{\rho} \tag{4}$$

In this expression, the matrix $\mathbf{F}$ is derived following a Wiener-based method by,

$$\mathbf{F} = \mathbf{E}_o\ \boldsymbol{\rho}_o^+ \tag{5}$$

where $\mathbf{E_o}$ (an $n \times m$ matrix) and ρ_o (an $k \times t$ matrix) are the SPDs and their corresponding sensor outputs for the training set of illuminants.[6]

2.3. Metrics for quality evaluation

To quantify the quality of the reconstructions we have used four different metrics [15]: the Goodness-of-Fit-Coefficient (GFC), the Root-Mean-Square-Error (RMSE), the CIELAB colour difference ΔE^*_{ab}, and the integrated irradiance (IIE). The GFC is based on the Schwartz's inequality and it is defined as,

$$GFC = \frac{\left|\sum_j f(\lambda_j) f_r(\lambda_j)\right|}{\sqrt{\left|\sum_j \left[f(\lambda_j)\right]^2\right| \left|\sum_j \left[f_r(\lambda_j)\right]^2\right|}} \tag{6}$$

where $f(\lambda)$ and $f_r(\lambda)$ are the original and the estimated spectral functions, respectively. Colorimetrically accurate illuminant estimations require GFC>0.995, GFC>0.999 indicates quite good spectral fit, and values GFC>0.9999 signifies an almost-exact fit. The CIELab colour difference formula was used to evaluate the colorimetric quality and differences of less than 3 CIELab units between the original and the estimated spectra were considered acceptable.

3. RESULTS

Table 1 shows some of the results using only a reduced number of coefficients (p=3) and all of them (p=29). The rank of factorisation p=29 has been selected according to the convergence rate derived from ICA algorithm.[13] Where n.a. appears, it means that the algorithm does not converge appropriately, e.g. ICA for p< 29 coefficients, or it makes not sense, i.e. direct-pseudo-inverse algorithm as it does not use coefficients.[5] First, the results show that the spectral recovery qualities are far from being good. In all cases, GFC values are clearly below 0.995 on average, which is the limit value for a spectrally accurate estimation. This is an expected result because of the spiky spectral profiles of the SPDs, which corresponds to the SPD of fluorescent lights that we have used here. On the contrary the colour differences are always around or below 3 CIELab units which indicates a quite good colorimetric estimation. These subtle differences between the colorimetric (ΔE^*_{ab}) and the spectral (GFC) metrics are originated in the presence of peaks of different heights in the SPDs of fluorescent illuminants.

Figure 1 we show examples of recoveries from some SPDs of fluorescent lights and the different estimation algorithms with k=3, k=6 and k=9 sensors. The integrated irradiance values (IIE) reveals the differences among the different methods and signal the importance of avoiding mononumerosis to evaluate the recovered spectra.

Second, the advantage of the NMF algorithm is that we can adjust the size of the coefficient matrix $\mathbf{H}$ to minimize the computational cost of SPD recovery. But by working with a reduced rank of p=3 both the spectral and the colorimetric

Table 1: Spectral and colorimetric quality of spectral recovery of test illuminants for the algorithms tested. The results are mean values for a set of 20 fluorescent SPDs.

Rank of parametrization		*p= 3*		
		GFC	Colour dif.	IIE
NMF Euclidean	*3 sensors*	0.9416	3.32	0.2931
	6 sensors	0.9654	2.88	0.2194
	9 sensors	0.9676	3.72	0.2037
NMF Divergence	*3 sensors*	0.9464	2.8	0.2697
	6 sensors	0.9678	2.46	0.1929
	9 sensors	0.9684	3.09	0.1811
ICA	*3 sensors*			
	6 sensors		n.a.	
	9 sensors			
Direct pseudo-inv	*3 sensors*			
	6 sensors		n.a.	
	9 sensors			

Rank of parametrization		*All components*		
		GFC	Colour dif.	IIE
NMF Euclidean	*3 sensors*	0.9506	1.95	0.254
	6 sensors	0.9786	0.96	0.1466
	9 sensors	0.9785	0.98	0.1347
NMF Divergence	*3 sensors*	0.9507	1.96	0.2537
	6 sensors	0.9789	0.73	0.1464
	9 sensors	0.9793	0.45	0.1305
ICA	*3 sensors*	0.9506	1.95	0.2537
	6 sensors	0.9787	0.74	0.1469
	9 sensors	0.9798	0.36	0.1299
Direct pseudo-inv	*3 sensors*	0.9507	1.95	0.2537
	6 sensors	0.9787	0.75	0.1464
	9 sensors	0.9798	0.36	0.1294

quality of recoveries are reduced in a similar way. In addition, results suggest that ICA and direct pseudo-inverse approaches lead to very similar results, and NMF is a little worse. In a previous work we found that daylight spectra could be recovered with a reduced number of sensors based on a priori analysis of an RGB set of signals from a white surface captured by a digital CCD camera.[6] The results shown here from the ICA and the direct pseudo-inverse methods confirm the CCD's potential as illuminant-estimation device even for a reduced training set of illuminants and the spiky profiles of some of these SPDs.

As we can see in the left column of figure 1, the values of the colour differences ΔE^*_{ab} suggest that acceptable illuminant colour identification is possible using a reduced number of sensors. The examples in the right column of the figure also show that colorimetric and spectral quality increase as the number of sensors increases. Nevertheless the recoveries worsen for "daylight fluorescents", which are characterised by different peaks and a background continuum (e.g. upper right SPD in figure 1).

4. CONCLUSSIONS

We have used three different algorithms, which are based on different mathematical background, for recovering SPD of fluorescent lights. The spectral profile of these lights are characterised by the presence of prominent peaks along the visible spectrum, and these peaks are difficult to recover using linear models with reduced number of parameters. The results shown here suggest that fluorescent lights can be recovered with both high spectral and colorimetric accuracy using no information about spectral sensitivities of the camera sensors or eigenvectors. Although nonnegative algorithms can reduce the computational cost of spectral devices, experiments show that ICA and direct pseudo-inverse methods consistently outperforms the NMF approaches, even for fluorescent lights and even using a reduced training set of illuminants.

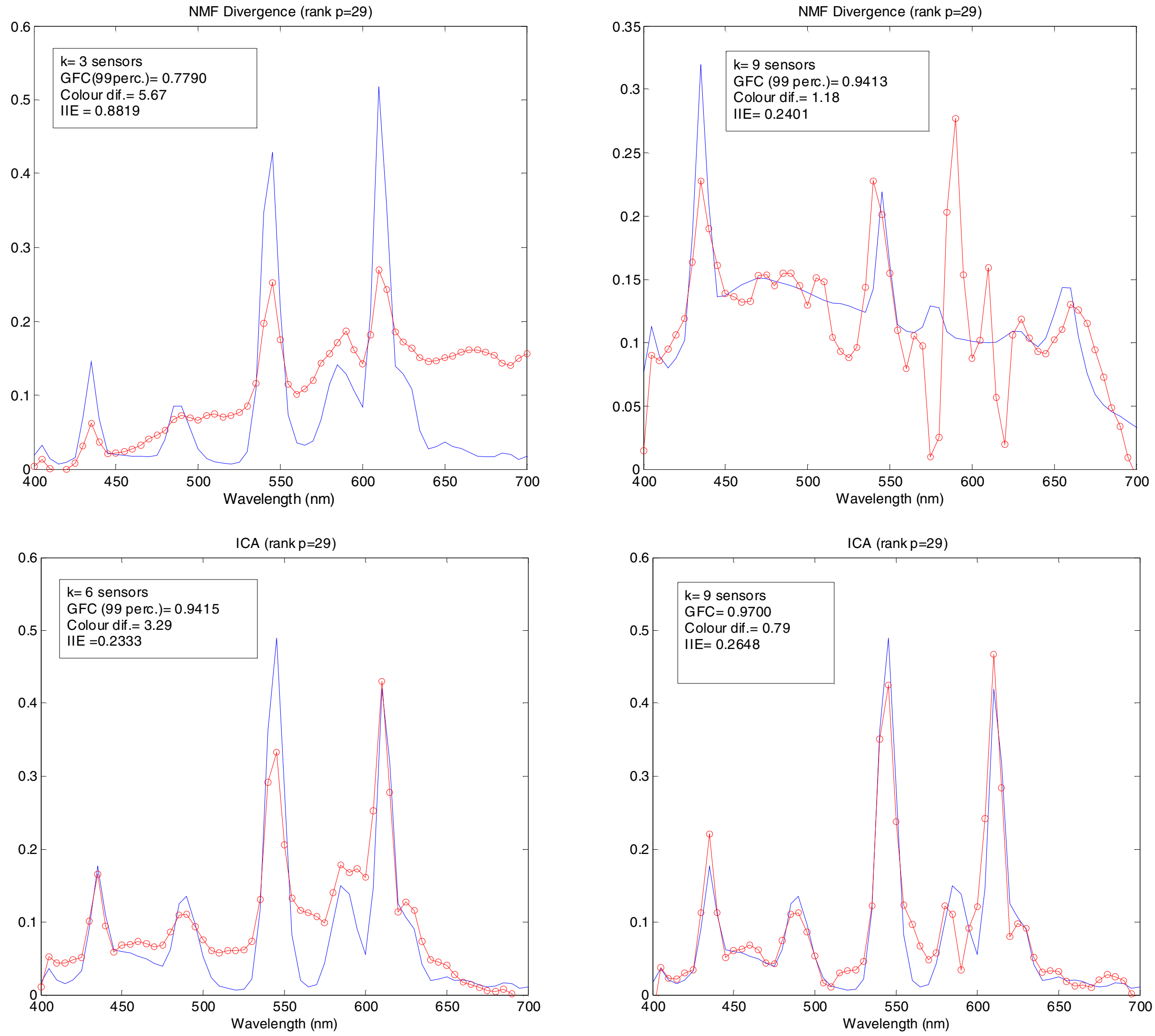

Figure 1: *Original (—) and recovered (o) spectrum using the NMF divergente update and ICA algorithm with simulated digital counts and different number of sensors.*

Acknowledgements

This work has been supported by the Comisión Interministerial de Ciencia y Tecnología (CICYT), Spain, under research grant DPI 2004-03734.

References

[1]. F.H. Imai, R. Berns and Di-Y. Tzeng, "A comparative analysis of spectral reflectance estimated in various spaces using a trichromatic camera system", J.Imaging Sci. Technol. 44, 280-371 (2000).

[2]. J. Hardeberg, F. Schmitt and H. Brettel, "Multispectral color image capture using a liquid crystal tunable filter", Opt. Eng. 41, 2532-2548 (2002).

[3]. F. H. Imai and R. S. Berns, Spectral estimation of oil paints using multi-filter trichromatic imaging, in Proc. of the 9th Congress of the International Colour Association, Rochester (NY, 2001), pp. 504-507.

[4]. J. Hernández-Andrés, J.L. Nieves, E. Valero and J. Romero, "Spectral-daylight recovery by use of only a few sensors", J.Opt.Soc.Am. A 21, 13-23 (2004).

[5]. M.A. Lopez-Alvarez, J. Romero, R. L. Lee, Jr. and J.Hernández-Andrés, "Designing a practical system for spectral imaging of skylight", Appl. Opt., 44, 5688-5695 (2005).

[6]. J.L. Nieves, E. Valero, S.M.C. Nascimento, J. Hernández-Andrés and J.Romero, "Multispectral synthesis of daylight using a commercial digital CCD camera", Appl. Opt. 44, 5696-5703 (2005).

[7]. 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, 1713-11721 (2000).

[8]. S. Tominaga, S. Ebisui and B.A. Wandell, "Scene illuminant classification: brighter is better", J. Opt. Soc. Am. A 1, 55-64 (2001).

[9]. S Tominaga and H. Haraguchi, "A spectral imaging method for classifying fluorescent scene illuminant", in Proceedings of 10th Congress of the International Colour Association, (Granada, 2005), pp. 193-196.

[10]. W. Xiong and B. Funt, "Independent component analysis and nonnegative linear model analysis of illuminant and reflectance spectra", in Proceedings of 10th Congress of the International Colour Association, Granada (Granada, 2005), 503-506.

[11]. K. Barnard, L. Martin, B. Funt and A. Coath, "A data set for color research", Color Res.Appl., 27, 147-151 (2002).

[12]. P. Paatero and U. Tapper, "Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values", Environmetrics, 5, 111-126 (1994).

[13]. A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis",. IEEE Transactions on Neural Networks, 10, 626–634 (1999).

[14]. D. D. Lee and H. S. Seung. "Algorithms for non-negative matrix factorization", Adv. Neural Info. Proc. Syst., 13, 556-562 (2001).

[15]. 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).

Author Biography

Juan L. Nieves received his BS in physics (1991) and his PhD in Sciences (1996) from University of Granada, Spain. Since then he has worked in the Department of Optics at the Science Faculty in the same University. His work has focused on colour vision, colour constancy and the development of multispectral techniques for recovering reflectances and illuminants.

