Reconstructing Spectral Reflectances with Mixture Density Networks

Alejandro Ribés and Francis Schmit
Ecole Nationale Supérieure des Télécommunications
46 Rue Barrault 75634 Paris Cedex 13, France
{ribes,schmitt}@tsi.enst.fr

Abstract

We consider the problem of spectral reconstruction from multispectral images by using non-linear methods. In the search for a neural network able to provide noise resistance and good generalization we apply Mixture Density Networks. This approach has been tested and compared with a linear method already used for spectral reconstruction of fine art paintings. This has been done using simulated and real data. Mixture Density Network based methods provide very good results in both cases. In particular, for real data acquisition we have scanned a Gretag-Macbeth™ color chart using a Minolta CS-100 spectroradiometer and a PCO SensiCam 370 KL monochrome camera with an electronically tunable liquid crystal spectral filter VariSpec VIS2. The results obtained using the data from this experiment clearly show the superiority of the Mixture Density Network based approach over the linear one used as a reference.

Introduction

We consider the problem of spectral reconstruction from multispectral images. The pixel value of a channel in a multispectral image is the result 1) of the spectral interaction of the light radiant distribution with the reflectance of an object surface and 2) of the spectral sensitivity of the camera combined with the transmittance of the optical path including the filter corresponding to this channel. Retrieving the spectral reflectance function of the object surface at each pixel is highly desirable. It allows a more general representation which is independent from light spectral distribution and from the camera used for the multispectral image acquisition. This representation can be used for many different purposes. Our interest is in high fidelity color reproduction of fine art paintings. As an example, knowing the spectral reflectances in each pixel allows us to simulate the appearance of a painting under any virtual illuminant.

In the particular case of color images the number N of channels is limited to three. Efforts have been made in order to characterize spectral reflectances using just three color channels. Some authors have proposed linear methods, as in Refs. 10 or 8. Others have proposed non-linear approaches using neural networks, see for instance Refs. 1 and 14 where spectral characterization is performed from RGB and YMC tristimulus values. On the other hand, neural networks have also been used for other purposes in colorimetry, see for instance Tominaga.[15]

In our case, we consider multispectral images with a higher number of channels (N>3) and we aim for a more precise spectral reconstruction than a raw estimation just satisfactory for subjective color reproduction purposes. Various linear and non-linear methods like splines, modified discrete sine transform (MDST), MDST with aperture correction, pseudoinverse, smoothing inverse or Wiener inverse have already been proposed as indicated by König and Praefcke in Ref. 9, see also Burns and Berns,[3] Herzog et al.[7] In particular, in the field of digital archives for fine art paintings, reconstruction of pigment spectral reflectance curves has mainly been obtained using linear methods, see for instance Refs. 12, 4 or 6. The first attempt using neural networks was proposed in a previous paper of the authors[13] where we studied the resistance to quantization noise of the spectral reconstruction obtained with different conventional neural networks and compared them with a linear method already used for spectral reconstruction of fine art paintings.[5]

In this paper we consider another approach based on a Mixture Density Network, in the search for a neural network able to provide noise resistance and good generalization, allowing good reconstruction for patterns not included in the training set.

Density Mixture Networks

A Mixture Density Network (MDN) is a method for solving regression or classification problems that consists in building a conditional probability density function between outputs and inputs of a given problem [[2]]. In the following $\mathbf{C}$ represents an input vector of dimension c, and $\mathbf{S}$ represents an output vector of dimension s.

The desired conditional probability density is modelled by a mixture of basis functions, usually chosen as Gaussians. The parameters of this mixture model are estimated from a set of known data (pairs of C and S vectors) using a neural network which can be any conventional neural network with universal approximation capabilities. In our case, the neural network used has a classical feedforward structure. The mixture model that represent the conditional probability density is of the form,

$$p(S|C) = \sum_{i=1}^{m} \alpha_i(C) g_i(S|C)$$

where m is the number of Gaussians used, $\alpha_i(\boldsymbol{C})$ are mixing coefficients, and every $g_i(\boldsymbol{S}|\boldsymbol{C})$ are the following a multidimensional Gaussian function:

$$g_i(S|C) = \frac{1}{(2\pi)^{2s}\sigma_i(C)^s} \exp\left\{ \frac{\|C - \mu_i(C)\|^2}{2\sigma_i(C)^2} \right\}$$

parameterized by scalars σ_i for the variances (all dimensions having the same variance) and vectors μ_i of dimension s representing their centres. Consequently, the vector $\boldsymbol{V}$ which parameterises the mixture model contains $m(1+1+s)$ elements.

A Mixture Density Network being based on a neural network, it needs a training phase. In this phase the neural network learns the mapping between each input vectors $\boldsymbol{C}$ and its associated parameter vector $\boldsymbol{V}$ defining a conditional probability density function. The learning process is driven by the minimization of the negative logarithm of the likelihood, formally:

$$E = \sum_{P} -\ln\left\{ \sum_{i=1}^{m} \alpha_i(C_p) g_i(S_p|C_p) \right\},$$

where p represents an index of a pattern and P is the number of patterns. Consequently we are training the system over a set of P pairs $(\boldsymbol{C}_p, \boldsymbol{S}_p)$. In the following section, for simplicity we will avoid the use of P and p.

Estimating Reflectances

Our aim is to estimate the spectral reflectance of pigments from multiespectral images. We are interested in the reconstruction of spectral curves between 400 and 760 nm which is the visible part of the spectrum. We reconstruct a spectral reflectance curve as a sequence of s regularly sampled values taken from 400 to 760 nm at constant d nm intervals.

Our problem consist in the construction of a system that maps a vector $\boldsymbol{C}$ containing camera values to a vector $\boldsymbol{S}$ representing a sampled spectral curve. As long as pairs $(\boldsymbol{C},\boldsymbol{S})$ are known this problem can be solved by the construction of a MDN system from this data.

In this context the probability $P(\boldsymbol{S}|\boldsymbol{C})$ becomes the conditional probability of a spectral curve $\boldsymbol{S}$ being obtained from a particular camera response vector $\boldsymbol{C}$. That means, we are building a function that assigns probabilities to all possible vectors $\boldsymbol{S}$ in a s dimensional space. Every point of this space represents the probability of a particular vector $\boldsymbol{S}$ being the counterpart of the given input $\boldsymbol{C}$.

Minimizing the negative logarithm of the likelihood over a database of pairs $(\boldsymbol{C},\boldsymbol{S})$ we fix the weights of the neural network of the MDN. Once the neural network trained, the MDN provides a mapping between a camera response vector $\boldsymbol{C}$ and a parameters vector $\boldsymbol{V}$. Of course, we are interested in finding a single sampled spectral curve $\boldsymbol{S}$ that provides the best estimation given a vector $\boldsymbol{C}$. For that purpose we need to chose a way to extract this vector $\boldsymbol{S}$ from the mixture model represented by the parameter vector $\boldsymbol{V}$.

In general, maximizing the obtained conditional density will give us the $\boldsymbol{S}$ with highest probability, that is indeed what we are looking for. But maximizing the mixture model is a problem not solved in closed form and implies the application of an iterative optimization procedure that is CPU consuming. We use a much quicker and simpler strategy by keeping as solution the vector $\boldsymbol{S}$ associated to the Gaussian with bigger mixing coefficient:

$$\max_i\{\alpha_i(C)\}.$$

This strategy is justify as long as in our problem we systematically obtain mixture models in which one Gaussian has a much bigger mixing coefficient than the others. In fact, we have compared results coming from different strategies and the one used (max) and the actual optimization of the function obtain mostly the same results. In our case, the maximum of the mixture model is well approximated by the means of the biggest Gaussian. In Figure 1 a graphical summary of the method can be seen.

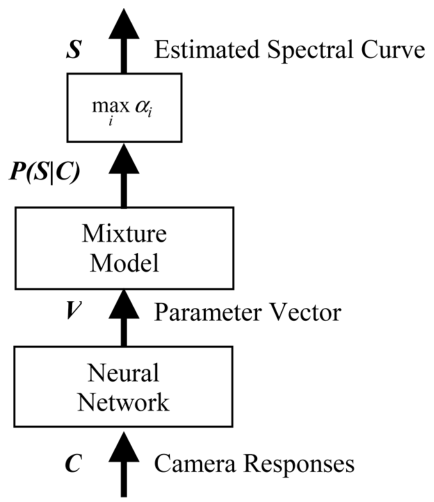

Figure 1. MDN Spectral Estimation.

Experimental Results

We have tested the proposed reconstruction approach by using both simulated and real data. We compare the results obtained using this mixture density network method with those obtained using the pseudo-inverse based reconstruction method described in Ref. 5. This method takes into account a database of spectral reflectances in order to constraint the solutions of the pseudo-inverse.

Simulated Data

Comparisons are performed over the four following spectral reflectance databases of pigments, the first three of them kindly provided by D. Saunders from The National Gallery, London:

- the "Kremer" database contains 184 spectral samples of pigments produce by Kremer Pigmente, Germany.

We use this database for training the MDN and to determine the linear transformation in the pseudo-inverse based method.

- the "Selected Artists" database contains 67 pigments chosen among a collection of artist's paintings.
- the "Restoration" database contains a selection of 64 pigments used in oil painting restoration.
- the "Munsell" database is not issue from the same canvas painting environment. It contains spectral curves corresponding to 1269 matte Munsell colour chart samples.

These databases having been sampled at different rates and with different limits, we resampled them in order to represent each spectral reflectance curve as a sequence of regularly sampled values from 400 to 760 nm at 10 nm intervals, which corresponds to s=37 values. To obtain the multispectral camera responses we use a simulated seven channel camera with equidistributed Gaussian filters over the range 400 to 760 nm, with 50 nm half-bandwidth. We choose as spectral sensitivity of the camera sensors a typical response of CCD arrays. If no noise is introduced in this simulation process, we remark that the theoretical camera model remains a perfect linear process. This is the reason that justifies the use of a linear based method as a reference method for spectral reconstruction.

In order to study the robustness of these methods in the presence of noise, we simulate acquisitions with quantization noise by using different numbers of bits for representing the camera channels. We present simulation results that shows the resistance of a Mixture Density Network for camera responses being quantized at 12, 10 and 8 bits. The choice of these three levels corresponds to the actual quantization levels observed on digital cameras currently available. The large signal to noise ratio (SNR) corresponding to 8 bits quantization is representative of most common digital images. The much lower SNR corresponding to 12 bits is available at the present time only on high-end digital cameras. Simulations performed with 12 bit quantization are indeed close to simulations without noise, and they provide results very similar to a perfect linear theoretical model. On the other hand, for 8 bit quantization the linear relationship is strongly corrupted by noise and the robustness of a reconstruction method against noise becomes predominant, which is not in favor of linear reconstruction methods.

We have tested different numbers of Gaussians in the mixture model and we have chosen the simplest MDN containing just one Gaussian (m=1, $\mathbf{V}$-dimension = 39) that indeed performs well in these simulations. The associated neural network hidden layer contains 28 neurons which correspond to a network with 1288 weights. The choice of the MDN architecture parameters is behind the scope of this paper and will not be discussed here.

In Table 1 we present our results as spectral reconstruction errors. For a given database they are calculated as the average of the L^1 distance (mean value of the absolute differences) between each real spectral curve and its reconstructed counterpart. We can see that at 8 bits this error is decreased about 40% for all databases tested. This result confirms that the MDN based method used is more robust in presence of noise than the linear reference one. It is also remarkable that the MDN response on 12 bits continues to be slightly better than the reference method, even if at this signal to noise ratio the reconstruction problem is nearly linear. Furthermore, we note that the MDN based method generalizes well over the three databases not used as training set, specially over the Munsell database since this database is not based on oil pigments as it is the case for the training set and the two others.

In order to compared the colorimetric behaviour of the reconstructed curves with the original ones, Table 2 shows the CIELAB errors corresponding to the same experiments as Table 1. For each database the CIELAB error is the average of the CIE 1976 CIELAB colour-difference between each real spectral reflectance curve and its reconstruction, D50 being used as reference illuminant. We observe the same general behaviour as in Table 1: the CIELAB error for the MDN method is always better in presence of strong noise than for the reference method and remains comparable when noise is low (12 bits quantization), although not clearly stated as it is for the spectral error.

Table 1. Spectral Error over different databases.

8 bits quantization	**pinv**	**MDN**
Kremer	0.0248	0.0138
Selected Artists	0.0230	0.0154
Restoration	0.0219	0.0136
Munsell	0.0202	0.0144
10 bits quantization	**pinv**	**MDN**
Kremer	0.0126	0.0094
Selected Artists	0.0119	0.0110
Restoration	0.0113	0.0186
Munsell	0.0114	0.0098
12 bits quantization	**pinv**	**MDN**
Kremer	0.0109	0.0089
Selected Artists	0.0105	0.0107
Restoration	0.0093	0.0081
Munsell	0.0103	0.0094

Table 2. CIELAB Error Over Different Databases.

8 bits quantization	**pinv**	**MDN**
Kremer	4.6996	2.9995
Selected Artists	4.2582	3.9300
Restoration	3.8773	2.7178
Munsell	2.8551	2.6556
10 bits quantization	**pinv**	**MDN**
Kremer	1.6944	1.4398
Selected Artists	1.7265	1.5712
Restoration	1.4521	1.1781
Munsell	1.3179	1.4599
12 bits quantization	**pinv**	**MDN**
Kremer	1.3351	1.2227
Selected Artists	1.1909	1.4603
Restoration	1.0956	1.0041
Munsell	1.0944	1.3353

We have scanned a GretagMacbeth™ color chart using a Minolta CS-100 spectroradiometer and a PCO SensiCam 370 KL monochrome camera with an electronically tunable liquid crystal spectral filter VariSpec VIS2. From this experiment we obtained 200 spectral curves from 380 to 780 nm sampled at 1 nm intervals, each curve corresponding to a patch of the chart. We also acquired 12 images of the Gretag-Macbeth™ chart using the PCO digital camera and 12 band-pass Gaussian-shaped filters using the tuneable filter, their centres being equally distributed from 400 to 740 nm with a mean half-bandwidth of 30 nm.

Table 3. Spectral Error Over *GretagMacbeth*™ Chart.

	pinv	MDN
Training Set	0.0267	0.0162
Test Set	0.0239	0.0134

In Table 3 we compare the spectral reconstruction errors (L^1 distance) obtained by a pseudo-inverse based method with a Mixture Density Network using 8 Gaussians in its mixture model and 40 neurons in the hidden layer of its feed-forward neural network. This comparison is performed over two complementary sets of measured patches belonging to the GretagMacbeth™ chart. Set 1 contains 150 patches and is used for training, set 2 contains 50 patches not included in the training set. We can see that the MDN based method globally decreases the errors about 40% on the training set and about 44% on the test set.

Table 4. CIELAB Error Over *GretagMacbeth*™ Chart

	pinv	MDN
Training Set	3.9707	2.6730
Test Set	4.1533	2.3248

Table 4 shows the same information as table 1 but for CIELAB errors. We observe that the MDN based method globally decreases CIELAB errors about 33% on the training set and about 44% on the test set.

In order to better compare the reconstruction behaviour of both methods we show in Figure 3 the spectral error histograms for the pseudo-inverse based and the MDN based method. The error has been linearly quantized into ten bands represented by bars. Each bar indicates the number of spectral curves belonging to its error band. We clearly see that the error distribution is much better for the MDN method, most spectral curve reconstruction errors remaining in the first three bands.

In figure 3 we include some examples of spectral curves in order to visually compare both reconstruction methods. Although we have observed that for same samples the linear method performs comparably or even better than the MDN method, in general we clearly see that MDN reconstructed curves match better the real reflectance curves. This is sensible as MDN spectral errors are statistically 40% better than the errors obtained by the linear reference method.

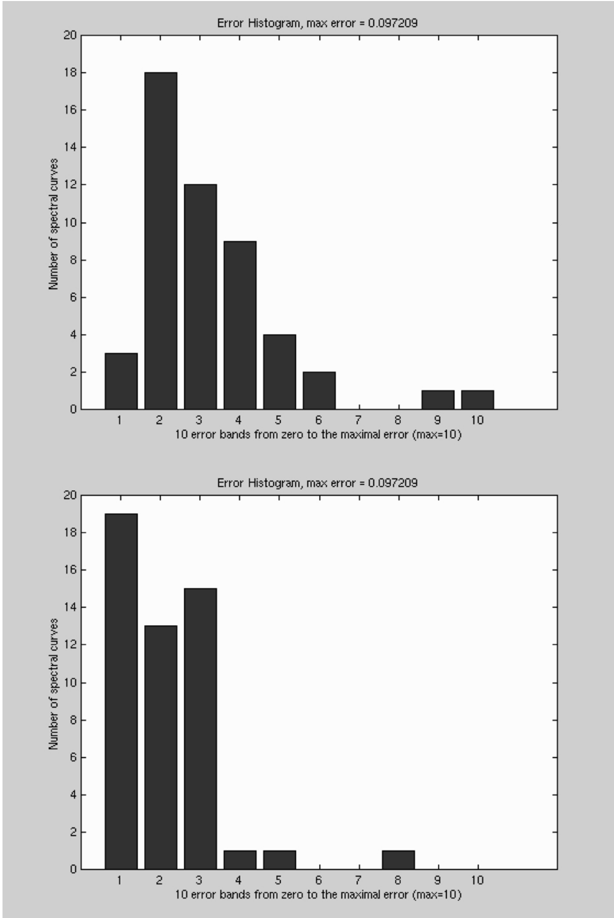

Figure 2. Histograms of the error for the pseudo-inverse based method (up panel) and the Mixture Density Network (MDN) based method (bottom panel).

Conclusion

We have developed a new spectral reconstruction method based in a mixture density network. We have compared this new method with a pseudo-inverse based one described in Ref. 5. For this comparison we have used simulated data in order to show the reaction of both methods in presence of noise. The new method performs better in mostly all cases.

We have used also real data as an end test to the new method. This real data was acquired in our laboratory using a spectroradiometer and a multispectral camera in controlled conditions. Afterwards both methods were applied to this real data and the mixture density network based one shows clearly superior results.

References

1. Y. Arai, S. Nakauchi and S. Usui. Color Correction Method Based on the Spectral Reflectance Estimation using a Neural Network. In proceedings of Fourth Color Imaging Conference Color Science, Systems, and Applications. November 1996. pp. 5-9.

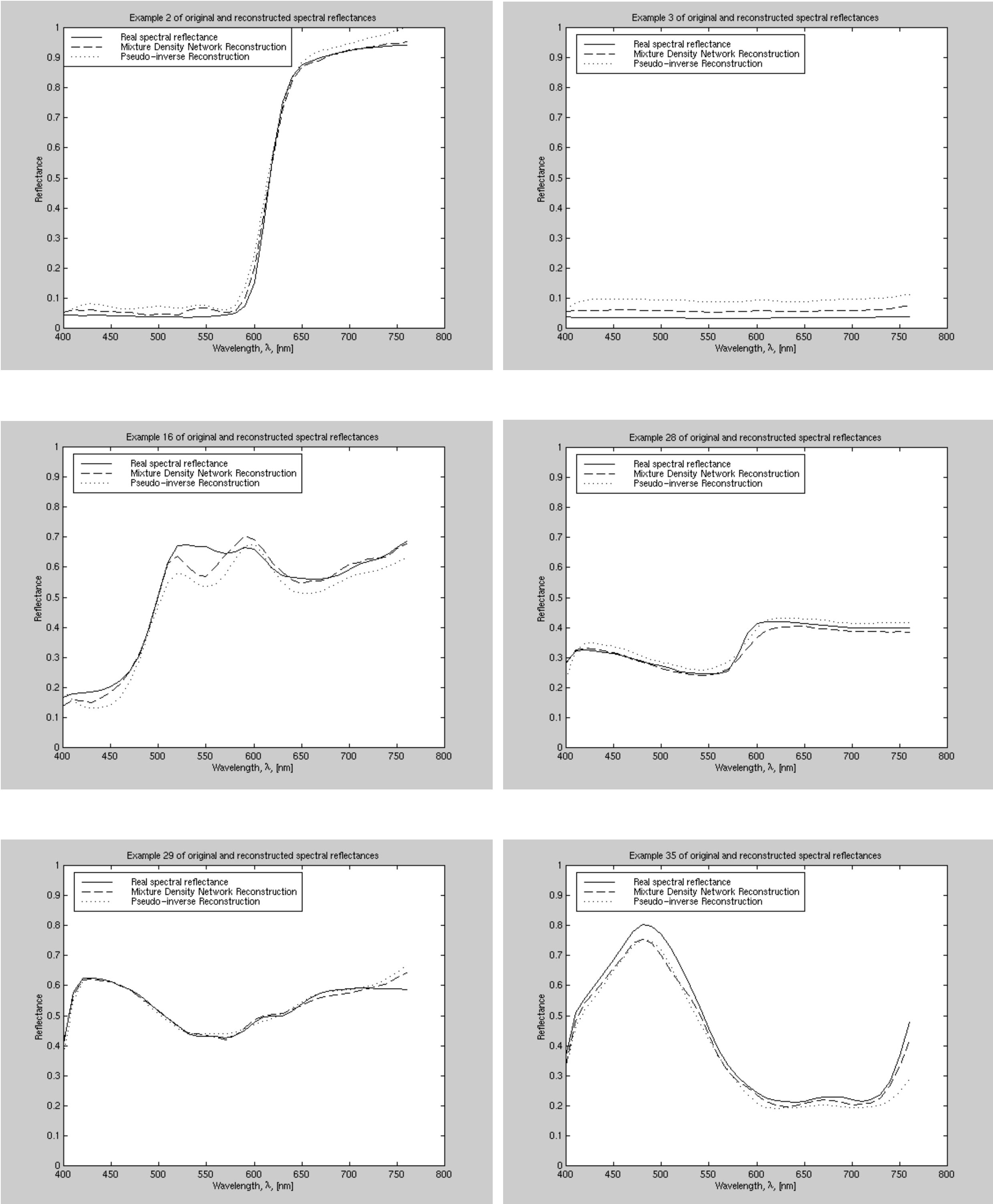

Figure 3. Six samples of reconstructed real curves taken from the GretagMacbeth^{TM} colour chart not belonging to the training set. Black continuous curves have been obtained by using a Minolta CS-100 spectroradiometer, dotted curves are reconstructed by the linear reference method and half-dotted curves are reconstructed by the MDN based method.

2. Chris M. Bishop. Mixture Density Networks. Neural Computing Research Croup Report NCRG/4288. Aston University, February 1994.
3. P. D. Burns and R. S. Berns. Analysis Multispectral Image Capture. In proceedings of Fourth Color Imaging Conference Color Science, Systems, and Applications. November 1996. pp. 19-22.
4. J. E. Farrell, J. Cupitt, D. Saunders and B. A. Wandell. Estimating Spectral Reflectances of Digital Images of Art. In proceedings of International Symposium on Multi-spectral Imaging and Color Reproduction for Digital Archives. October 1999. pp. 58-64.
5. J.Y. Hardeberg, F. Schmitt, H. Brettel, J. Crettez, H. Maître. Multispectral image acquisition and simulation of illuminant changes. Chapter in Ref. 11 pp. 145-164.
6. H. Haneishi, T. Hasegawa, N. Tsumura, and Y. Miyake. Design of Color Filters for recording Art Works. In proceedings of IS&T 50th Annual Conference, 1997. pp. 369-372.
7. P. G. Herzog, D. Knipp, H. Stiebig and F. König. Characterization of novel three and six channel color moiré free sensors. In proceedings of SPIE, Color Imaging: Device Independent Color, Color Hardcopy, and Graphic Arts IV. January 1999. pp. 48-59.
8. F. H. Imai and R. S. Berns. Spectral Estimation Using Trichromatic Digital Cameras. In proceedings of International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives. October 1999. pp. 42-49.
9. F. König and W. Praefcke. A multispectral scanner. Chapter in [[11]] pp. 129-144.
10. H. Kotera, H. Motomura and T. Fumoto. Recovery of Fundamental Spectrum from Color Signals. In proceedings of International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives. October 1999. pp. 141-144.
11. Color Imaging. Vision and Technology. Edited by L. W. MacDonald and M. R. Luo. John Wiley & Sons, Ltd. 1999.
12. H. Maître, F. Schmitt, J. Crettez, Y. Wu, and J.Y. Hardeberg. Spectrophotometric image analysis of fine art paintings. In Proc. IS&T and SID Third Imaging Conference, 1996. pp. 50-53.
13. A. Ribés, F. Schmitt, and H. Brettel. Reconstructing Spectral Reflectances with Neural Networks. In proceedings of Third International Conference on Multispectral Color Science. June 2001. pp. 9-12.
14. T. Sato, Y. Nakano, T. Iga, S. Nakauchi and S. Usui. Color Reproduction Based on Low Dimensional Spectral Reflectance Using the Principal Component Analysis. In proceedings of International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives. October 1999. pp. 185-188.
15. Shoji Tominaga. Color coordinate conversion via neural networks. Chapter in [[11]] pp. 165-178.

Biography

Alejandro Ribés received a computer science engineering degree from the *Universitat Jaume I*, Spain, and a DEA (one-year French postgraduate degree in research) from the *Université de Nice-Sophia Antipolis* specialized in image processing and artificial vision. He is currently preparing a Ph.D. in multispectral imaging at the *Ecole Nationale Superieure des Telecommunications* in Paris, since September 2000.

Francis Schmitt received an engineering degree from the *Ecole Centrale de Lyon*, France, in 1973 and a PhD degree in physics at the *Université Paris VI* in 1979. He has been a member of *Ecole Nationale Superieure des Telecommunications* (ENST) since 1973. He is Full Professor at the Image and Signal Processing departement. His main interests are in computer vision, 3D modelling, computational geometry, picture analysis and synthesis, colorimetry, multispectral imagery. He is author or co-author of about 100 papers in these fields.