COMPARATIVE ANALYSIS OF GAUSSIAN AND LINEAR SPECTRAL MODELS FOR COLOUR CONSTANCY

Dmitry P. Nikolaev and Petr P. Nikolayev
Institute for Information Transmission Problems, RAS
Bol'shoj Karetnyj lane, 19
101447, Moscow, Russia
E-mail: nikol@iitp.ru

KEYWORDS
Colour constancy, spectral models, Gaussian spectral model.

ABSTRACT

The present work considers the features and advantages of a Gaussian model of spectral functions approximation for solving the colour constancy problem. The peculiarities and advantages of the Gaussian model, as compared with traditional linear approximation models, are examined. The possibility of stable numerical estimation of the reflectance features of object surfaces on the basis of the spectral stimulus captured by a camera at different conditions of scene illumination is demonstrated. The efficiency of colour constancy algorithms that use the Gaussian model is estimated.

INTRODUCTION

In machine vision, it is often required not only to compute the object segmentation map of an image (Nikolaev and Nikolayev 2004), but also to estimate the coloration of object surfaces. Such a problem arises, for example, when processing videos or indexing image databases in the case when a target object should be successively detected, regardless of the illumination and observation conditions, which aren't known *a priori*.

The ability of the vision system of the man and animals to estimate the reflective properties of surfaces in the case when the illumination chromaticity changes (Helmholtz 1910; Nyuberg et al. 1971a), that is, the phenomenon of colour constancy (CC), has been discussed quite explicitly. CC algorithms suitable for machine vision systems are known too (Forsyth, 1990; Finlayson et al., 2001; Finlayson and Schaefer, 2001; Barnard et al., 2002). But as a rule they don't involve the spectral models of scenes and so they aren't suitable for processing of the images captured by some different sensors.

The elaboration of a solution method for such CC tasks requires introducing some *a priori* restrictions on the optical properties of the medium in which the sensor system operates. Due to considerable irreversible information loss during the transformation of a radiation signal into a sensor system response, it is necessary to match the language of spectral descriptions with the colour discrimination capabilities of the system. For this purpose, various models are introduced for approximating the spectral characteristics of the whole vision process, from the emission properties of the illumination sources to the photosensitivities of the colour channels of the system. Such restrictions and approximations turn the CC problem into a solvable one, making it possible to build a mathematical model of the CC (Nikolayev 1985).

The CC problem allows various definitions that differ in the level of the requirements set for the final result. If it is possible to assign an etalon vector stimulus (e. g., 3-stimulus from "white" illumination) to each uniformly coloured object contained in the scene and if this stimulus is invariant to the illumination and observation conditions, then the problem of the constant estimation of the object coloration can be considered solved for this particular type of a sensor. As an etalon stimulus, it is convenient to use the stimulus that would be produced by the object observed in a hypothetic situation when it is illuminated by a diffuse equal-energy source of a given brightness. In such a definition, it is possible to describe the CC mechanism as a "correction of illumination chromaticity" (Helmholtz 1910). Such a solution of the CC problem does not allow one to predict what vector stimulus will be obtained at the output of a different sensor observing the same object. In this sense, the above problem definition can be called "weak". In the framework of this definition, the constant estimation of the source chromaticity is reduced to finding the locus of the vector stimuli obtained when directly observing diffuse sources of the same chromaticity and an arbitrary brightness. As for the problem of the constant estimation of the brightness of the scene illumination sources, it is in principle insoluble if there is no information on the scene geometry (on the arrangement of its objects and their shapes).

The estimates obtained in the framework of the "weak" CC definition are described by the magnitudes of the vectors of sensor response to incident radiation, that is, they are the projections of some spectral stimuli onto the colour space (CSp) of the particular sensor. To obtain constant estimates of the object coloration and the illumination chromaticity, which are also invariant to the choice of a sensor, it is necessary to retrieve a spectral stimulus from its CSp projection (from the

Proceedings 19th European Conference on Modelling and Simulation
Yuri Merkuryev, Richard Zobel, Eugène Kerckhoffs © ECMS, 2005
ISBN 1-84233-112-4 (Set) / ISBN 1-84233-113-2 (CD)

vector of response to it). Obviously, this problem is, in general, insoluble and becomes soluble only when a unique response corresponds to each of the possible etalon spectral stimuli. The restriction on the whole variety of colorations and illuminations to a set of the spectra meeting the above condition is usually performed by introducing a spectral model with a limited number of parameters. We will refer to the retrieval (in the framework of the spectral model accepted) of the spectral composition of illumination and the reflection spectra of objects as a "strong" definition of the CC problem.

Most researchers dealing with the CC problem use so-called linear spectral models (LSM). In a LSM, the space of spectral functions is confined to a 3D linear subspace of the function space. Interesting particular cases of LSM are models in which the bases are step-wise functions (so-called "banded spectral model" (Stiles and Wyszecki 1962; Land and McCann, 1971; Nyuberg et al. 1971b)) or the functions of spectral sensitivity of the sensor (Lee et al. 1995). However, all these models have the same drawback: they cannot adequately describe stimuli of a high saturation. This fact is quite obvious if taking into account that spectrally separated stimuli are linearly independent.

One of few nonlinear spectral models suggested so far is a Geusebroek model (Geusebroek et al. 2001), in which the spectral functions are approximated with second-order polynomials, and the functions of spectral sensitivity of the sensor are considered decomposable on some basis. As it will be demonstrated below, such decomposition is quite far from the ideal, both for human receptors and for an RGB-camera.

These drawbacks are no longer valid for the Gaussian spectral model (GSM) we suggested earlier (Nikolaev 1985). In this model, the spectral stimuli, the reflective properties of scene surfaces, and the sensor sensitivity functions are approximated with a 3-parameter set of various Gaussians. Let us consider some features of the GSM, trying to demonstrate its advantages.

SENSOR MODELING IN LSM AND GSM

The Geusebroek model is accurate only for those sensors whose sensitivity spectra are decomposable on a basis that consist of a Gaussian function and its first and second derivatives. Figure 1 shows the results of approximating the sensitivity spectra of the human eye (a, c, e (Dartnall et al. 1983, Table 2)) and an RGB camera (b, d, f – a FillFactory camera with an infrared filter, $\lambda_0 \approx 630nm$). The solid marks correspond to the real experimental data, while the curves correspond to best data fit. Figures 1a,b show the best fit between the Geusebroek model and the real data for the human eye and the RGB camera (the root-mean-square errors are $s_a = 0.05$ and $s_b = 0.01$, respectively). As it can be seen from the plots, such an approximation cannot be considered good.

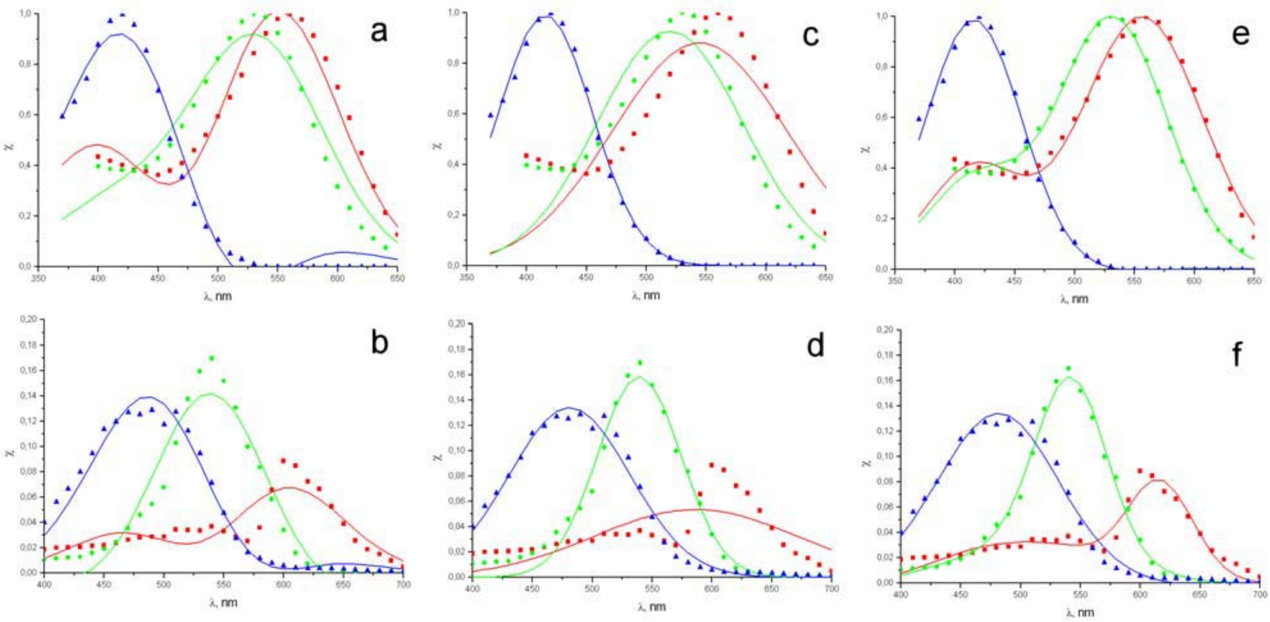

Figure 1: Results of Fitting the Sensitivity Spectra of the Human Eye and an RGB Camera

For the GSM, it is known (Nikolayev 1985) that the model is accurate for a sensor whose each sensitivity spectrum has a Gaussian profile. Unfortunately, for real sensors, such an approximation is even less adequate than the Geusebroek one (Figures 1c,d, the root-mean-square errors are $s_c = 0.09$ and $s_d = 0.01$, respectively). However this result can be considerably improved if taking into account that the response vectors of sensors with linearly dependent sensitivities are also linearly dependent. For example, if the sensitivity spectra of the sensor have the following form:

$$\chi_{1i}(\lambda)=\sum_{j=1}^{3} l_{ij}\cdot G(\lambda,h_j,s_j),\ 1\le i\le 3, \qquad (1)$$

where $G(\lambda,h,s)=\exp\left(-s\cdot(\lambda-h)^2\right)$, then there exists a linear transform of vectors from the CSp of this sensor to the Csp of a sensor with the sensitivity $\chi_{2i}(\lambda)=G(\lambda,h_i,s_i)$, for which the GSM is adequate. Figures 1e,f show the best fit between the extended GSM (EGSM) and the real curves for the human eye and the RGB camera (the root-mean-square errors are $s_e=0.02$ and $s_f=0.005$, respectively). As it can be seen from the plots, among all the models considered, the EGSM describes both sensor types most adequately. For the Lee model, evidently, the approximation error for the sensor sensitivity spectra is definitely zero.

To obtain the parameters of the sensor model for the EGSM, we used a shooting method that includes a few stages, in order to decrease the probability of finding a local optimum instead of the global one. At the first stage, shooting was performed for the function $\chi_{11}(\lambda)$ and the parameters l_{11},s_1,h_1. At the second stage – for $\chi_{11}(\lambda),\chi_{12}(\lambda)$ and $l_{11},l_{12},l_{21},l_{22},s_1,h_1,s_2,h_2$, using the results of the first stage as the initial estimate. At the third stage, shooting concerns all 15 parameters (9 weight coefficients l_{ij} and 3 pairs of parameters h_j and s_j, which define the three sensitivity functions in the form (1)), starting from the values obtained at the second stage.

RETRIEVAL OF A SPECTRAL STIMULUS FROM RESPONSES IN LSM AND GSM

After the spectral model is settled, it becomes possible to link the CSp of the sensor to the spectral stimuli function space and to solve the CC problem in the «strong» definition if it is already solved in the «weak» one. Thus, as applied to an LSM, the following expression can be obtained for the vector of the sensor response to a spectral stimulus:

$$\vec{a}=\int_0^\infty \vec{\chi}(\lambda)\cdot F(\lambda)\cdot d\lambda=\int_0^\infty \vec{\chi}(\lambda)\cdot\left(\vec{\alpha}\cdot\vec{b}(\lambda)\right)d\lambda=B\vec{\alpha}, \ (2)$$

where $F(\lambda)$ is the spectral stimulus, $\vec{\chi}(\lambda)$ is the spectral sensitivity vector of the sensor, $\vec{b}(\lambda)$ is the vector of basis functions of the model, and B is the matrix of responses to the basis functions: $B_{ij}=\int_0^\infty \chi_i(\lambda)\cdot b_j(\lambda)\cdot d\lambda$. It is evident that in an LSM a spectral stimulus is retrievable for any sensor with $|B|\ne 0$ and can be expressed as follows:

$$F(\lambda)=\vec{b}(\lambda)\cdot B^{-1}\vec{a}. \qquad (3)$$

For a trichromatic sensor that satisfies the GSM, we earlier described an analytic procedure of calculating the parameters of the spectral stimulus $F(\lambda)=L\cdot G(\lambda,H,S)$ from the corresponding response vector $\vec{a}$, provided that the dispersions of the sensor sensitivity spectra are equal ($s_1=s_2=s_3$) (Nikolayev 1985). Let us show how a Gaussian stimulus $F(\lambda)$ can be retrieved from the responses to it, produced by an arbitrary GSM sensor.

In the general case of a sensor with sensitivities $\chi_i(\lambda)=G(\lambda,h_i,s_i)$, it is not difficult to find the H and L components if the stimulus saturation, S, is already found:

$$\begin{cases} H=\dfrac{\sum_{i=1}^{3}\left(S\cdot s_i\cdot\left(h_i^2-1\right)+\left(S+s_i\right)\cdot\ln\dfrac{a_i\cdot\sqrt{S+s_i}}{\sqrt{\pi}}\right)\cdot r_i}{-2\cdot S\cdot\sum_{i=1}^{3} h_i\cdot s_i\cdot r_i} \\ L=\sum_{i=1}^{3} e_i\cdot a_i \Big/ \sum_{i=1}^{3} e_i^2 \end{cases},(4)$$

where $r_i=s_{(i+1)\bmod 3}-s_{(i+2)\bmod 3}$, and $e_i=\exp\left(-S\cdot s_j\cdot\left(H-h_j\right)^2/\left(S+s_j\right)\right)\cdot\sqrt{\pi/\left(S+s_j\right)}$. The equation for finding S is, in general, transcendental, has a complex structure, and, to all appearance, has no analytical solution. Nevertheless, one can solve this equation numerically, using, for example, a shooting method. In this case, it is reasonable to take as the initial approximation of S the solution obtained for the equal dispersions of the sensitivity spectra:

$$\frac{1}{S}=\frac{\sum_{i=1}^{3} h_i\cdot h_{(i+1)\bmod 3}\cdot\left(h_i-h_{(i+1)\bmod 3}\right)}{\left(h_1-h_2\right)\cdot\ln\dfrac{a_3}{a_2}+\left(h_2-h_3\right)\cdot\ln\dfrac{a_1}{a_2}}-\frac{1}{s_2}. \qquad (5)$$

Each iteration of the shooting procedure implies calculating the parameters $\{H,L\}$ form the current value of S, using Eq.(4), and subsequently finding the response vector for the Gaussian stimulus with the parameters $\{L,S,H\}$. Shooting is performed with respect to the parameter S, the goal being to minimize the discrepancy between the computed and the actual response.

ESTIMATION OF CC EFFICIENCY FOR LSM AND (E)GSM

Let us build an algorithm for numerically solving the CC problem for an LSM and (E)GSM under equal conditions (identical sets of coloured samples to test, the same set of illumination sources, and a single RGB camera) and compare the solution accuracies. Assume that the scene illumination conditions allow us to use the simplest mechanism of the color constancy – «illumination correction by a white sample» (Land 1977).

Thus, in each experiment (observation of the n-th sample illuminated with the k-th source) the algorithm receives two input vector stimuli: $\vec{a}_{n,k}=\int_0^\infty \vec{\chi}(\lambda)\cdot\Phi_n(\lambda)\cdot S_k(\lambda)\cdot d\lambda$ and

$\vec{a}_{0,k} = \int_0^\infty \vec{\chi}(\lambda)\cdot S_k(\lambda)\cdot d\lambda$ (from a white sample). The goal of the algorithm is to estimate $\Phi_n(\lambda)$, that is, to solve the CC problem in its «strong» definition. This task can be solved in two steps:

1) Using the methods described in the previous section, retrieve the spectral stimuli $F_{n,k}(\lambda) = \Phi_n(\lambda)\cdot S_k(\lambda)$ and $F_{0,k}(\lambda) = S_k(\lambda)$ from the response vectors $\vec{a}_{n,k}$.

2) Estimate the reflectance curve of the n-th sample as the ratio of the stimuli found at the first step: $\tilde{\Phi}_{n,k}(\lambda) = \tilde{F}_{n,k}(\lambda)/\tilde{F}_{0,k}(\lambda)$.

This is, in fact, the solution of the CC problem.

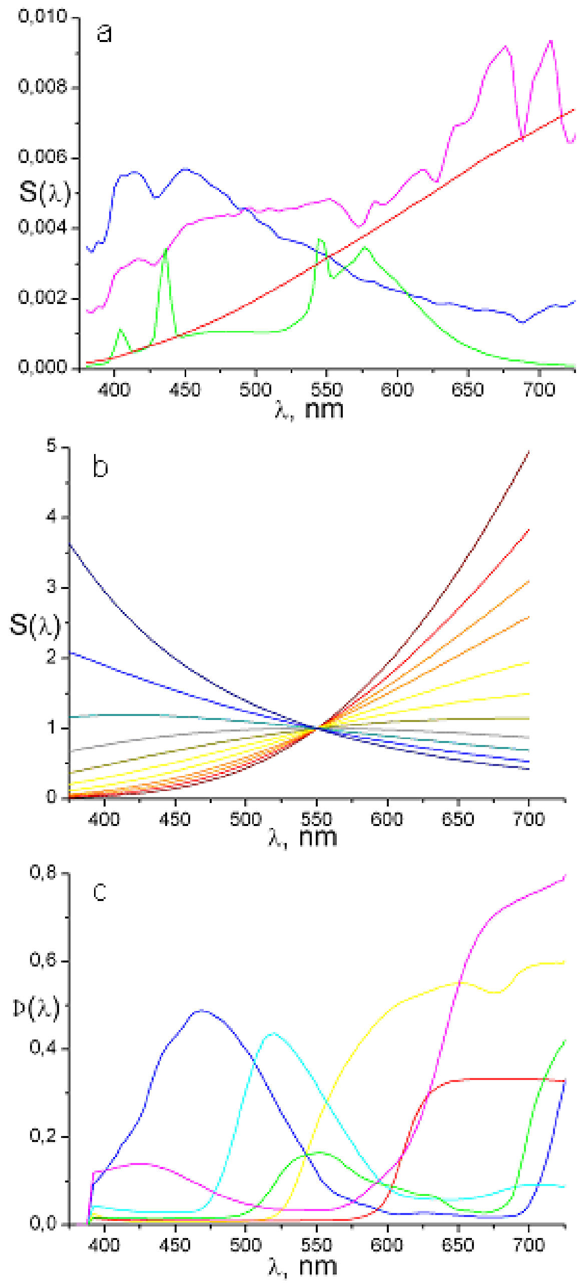

Figure 2: Emission and Reflectance Curves Used in the Experiment

To estimate the efficiency of a spectral model, let us calculate a «virtual vector stimulus», $\tilde{\vec{f}}_{n,k} = \int_0^\infty \vec{\chi}(\lambda)\cdot \tilde{\Phi}_{n,k}(\lambda)\cdot d\lambda$, for the same n-th sample being illuminated with an etalon equal-energy source and compare it with the true stimulus $\vec{f}_{n,k} = \int_0^\infty \vec{\chi}(\lambda)\cdot \Phi_{n,k}(\lambda)\cdot d\lambda$.

To compare LSM and EGSM, we chose among the LSM family the model suggested by Lee (Lee et al. 1995) and banded spectral model. As already mentioned above, Lee model uses the vector of spectral sensitivity of the sensor, $\vec{\chi}(\lambda)$, as the linear basis. The two models were compared along the following scheme: for a set of N samples and a selected sensor, we calculated the estimates $\tilde{\Phi}_{n,k}(\lambda)$ and fixed on the chromaticity space of the sensor the Euclidean distance, $r_{n,k}$, between the projections of the estimation vector $\vec{f}_{n,k}$ and the «etalon» $\vec{f}_n$. As a measure of the efficiency of each model for the given illumination conditions, we use the root-mean-square (over all N samples) deviation, r_k, of the actual estimate from the «absolutely» constant one. 170 reflectance functions of «natural colorants» and 4 emission spectra of «natural non-Planck sources» tabbed with a step of *4nm* in the range from *375nm* to *750nm* were taken from the database of the Computational Vision Lab of S. Fraser University, Canada (http://www.cs.sfu.ca/~colour). To better compare our results with results of other groups worldwide, we added to this set 11 Planck sources with colour temperatures ranging from 2000^0K to 30000^0K.

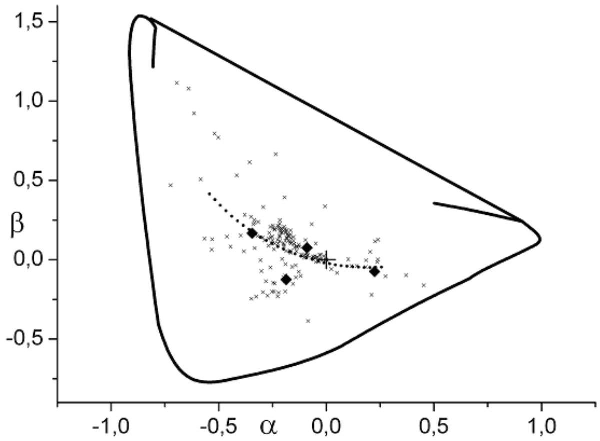

Figure 3: Positions of Stimuli on the Chromaticity Plane

Figure 2 shows normalized emission curves for 4 non-Planck sources (a) and 11 Planck sources (b); and 6 reflectance curves for the most saturated samples (c) (from those 170), which were used in the experiment with CC models. Figure 3 displays the positions of the vector stimuli from 15 sources and 170 samples on the chromaticity plane of the sensor (the latter were calculated for the case of an equal-energy illumination source).

Figure 4 maps onto the sensor chromaticity plane the composite diagrams for the LSM (left column – Lee model, right column – banded model) and EGSM (central column) under ideal Planck (bottom row) and «natural» (top row) illumination. The hollow circles in the central parts of the diagrams correspond to shifts of the *constant estimates* of sample chromaticities with respect to the «ideal CC» positions, while the solid dots correspond to shifts of the initial aconstant values.

The more compact the hollow circle cluster is relative to the solid dot cluster, the better the CC algorithm operates. The overall statistics of this experiment is the following. Under «natural» illumination (K=4), the efficiency is q =7.9 for the EGSM, while it is q =3.6 for the Lee LSM and q=6.2 for banded LSM. Planck illumination (K=11) increases q to 8.0 for the EGSM, while keeping it at 3.6 for the Lee LSM and increasing q to 7.0 for banded LSM.

The efficiency q was calculated by averaging (over the K sources) the ratio $r_k / r_{0,k}$, where $r_{0,k}$ is the root-mean-square deviation from the «ideal» estimate for the *aconstant* (i.e., input) stimulus, and r_k is the same deviation for the *constant* estimate of each sample in the CC model chosen. Thus, over the set of 170 types of functions $\Phi_n(\lambda)$ and for two sets of illumination sources, the q-criterion judges the EGSM to be more advantageous in the sense of the accuracy of estimating these functions.

Furthermore, the experiment allowed us to qualitatively conclude that the EGSM estimate accuracy increases with the proximity of $\Phi_n(\lambda)$ to a symmetric single-extremum function within the sensor sensitivity range.

In conclusion, we would like to note that the triad of parameters, h, s, and L, of the Gaussian curve, which fits in the GSM all the spectral curves of the vision process, substantially correlates with the three components of the colour perception of a trichromate: hue, saturation, and brightness (or lightness, which is exactly the same), respectively.

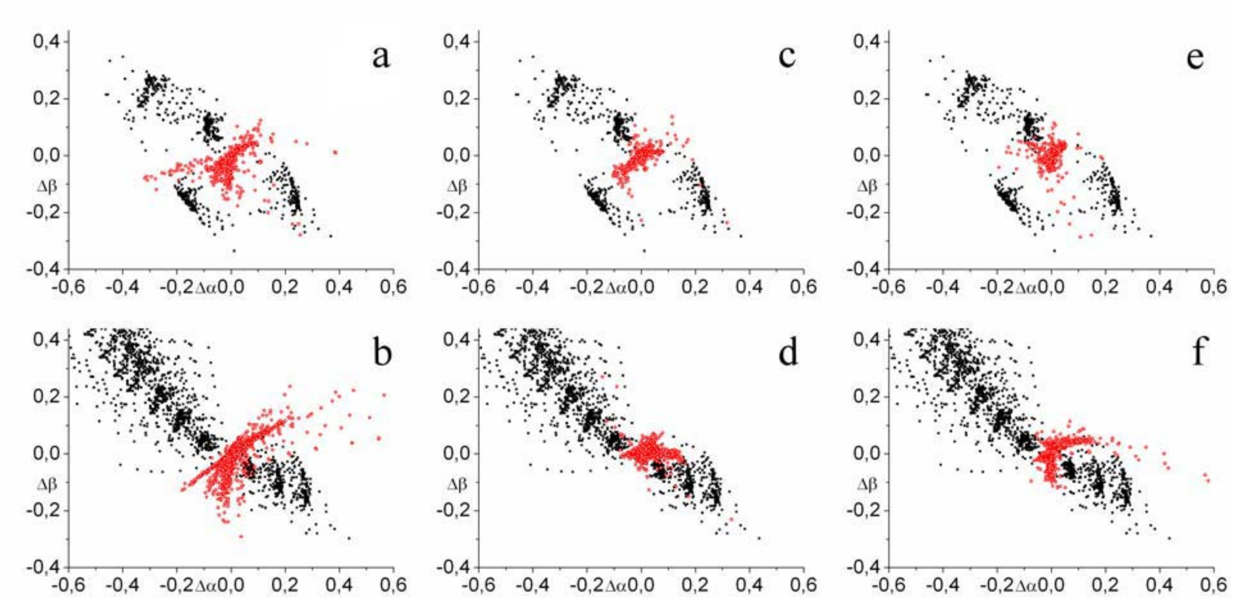

Figure 4: Colour Constancy Efficiency of EGSM and LSM

CONCLUSION

The present work suggests to use a Gaussian model of spectral functions approximation for solving the CC problem. This model demonstrates good performance in numeric modelling with the extensive set of natural pigments illuminated both ideal Planck and real light sources. Comparative analysis indicates that a Gaussian model is potentially more effective at estimating the chromaticity of the scene illuminant and the scene objects than the linear spectral models while using the same CC algorithm. This was demonstrated for the "banded spectral model" (that underlies scale-by-max Retinex algorithm) and Lee's linear model that uses functions of spectral sensitivity of the sensor as a basis. In the following we plan to proceed the comparison of the performance capabilities of our spectral model for CC with other existing linear and non-linear spectral models.

REFERENCES

Barnard K.; L. Martin, A. Coath and B. Funt. 2002. "A comparison of color constancy algorithms. Part two. Experiments with image data". *IEEE Trans. in Image Processing. 11*. No. 9, 985-996.

Dartnall, H.J.; J.K. Bowmaker and J.D. Mollon. 1983. "Human visual pigments: microspectrophotometric results from the eyes of seven persons". *Proc. of Royal Soc. of London. B 220*, 115-130.

Finlayson, G.D.; S. Hordley and P. Hubel. 2001. "Color by correlation: a simple, unifying framework for color constancy". *IEEE Trans. Pattern Anal. Mash. Intel. 23*, No. 12, 1209-1221.

Finlayson, G.D. and G. Schaefer. 2001. "Solving for colour constancy using a constrained dichromatic reflection model". *Int. J. Computer Vision.. 42,* No. 3, 127-144.

Forsyth D.A. 1990. "A novel algorithm for colour constancy". *Int. J. Computer Vision. 5,* No. 1, 5-36.

Geusebroek, J.-M.; R. van den Boomgaard and A. Smeulders. "Color invariance". 2001. *IEEE Trans. Pattern Anal. Mash. Intel. 23*, No. 12, 1338-1350.

Helmholtz, H.L.F. 1910. *Theatise on Physiological Optics.* English translation in 1925. Dover, N.J.

Land, E.H. and J. J. McCann. 1971. "Lightness and retinex theory". *J. Opt. Soc. Amer. 61*, 1-11.

Land, E.H. 1977. "The retinex theory of color vision". *Sci. Am. 237*, No. 6, 108-128.

Lee, S.D; C.Y. Kim and Y.S. Seo. 1995. "Linear model of surface and scanner characterization method". *Proceedings of SPIE. 2414*, 84-93, Device-Independent Color Imaging II, Eric Walowit; Ed.

Nikolayev, P.P. 1985. "Model of colour constancy for the case of continuous spectral functions". *Biophysics* 30, No. 1, 112-117 (in Russian and English translation).

Nikolaev, D.P. and P.P. Nikolayev. 2004. "Linear colour segmentation and its implementation". *Comp. Vision and Image Understanding. 94*, 115-139.

Nyuberg, N.D.; M.M. Bongard and P.P. Nikolayev. 1971a. "I. On the constancy of perception of coloration". *Biophysics* 16, No. 2, 285-293 (in Russian).

Nyuberg, N.D.; P.P. Nikolayev and M.M. Bongard. 1971b. "II. On the constancy of perception of coloration". *Biophysics* 16, No. 6, 1052-1063 (in Russian).

Stiles, W.S. and G.W. Wyszecki 1962. "Counting metameric object colors". *J. Opt. Soc. Amer 52,* No. 2, 313-328.

AUTHOR BIOGRAPHIES

Dmitry P. Nikolaev was born in Moscow, Russia. He studied physics, obtained master degree in 2000 and PhD degree in 2004 from Moscow State University. Since 2000, he has been a research scientist at the Institute for Information Transmission Problems, RAS. His research activities are in the areas of computer vision with primary application to colour image understanding. His e-mail address is dimonstr@iitp.ru and his Web-page can be found at http://www.ddt.ru/chaddt/Lace/pub.html.

Petr P. Nikolaev was born in Moscow, Russia. He studied physics and obtained master degree in 1966 from Moscow State University. He obtained PhD degree in biology in 1975 from Institute for Information Transmission Problems, RAS and Dr. of Science degree in physics in 1993 from Institute for Biophysics, RAS. Since 1966, he has been a research scientist at the Institute for Information Transmission Problems, RAS where he is now leading a research group in the field of computer colour vision. His e-mail address is nikol@iitp.ru.