

Spaces of Spectral Distributions and Their Natural Geometry

Reiner Lenz

Department of Science and Technology, Linköping University,
SE-60174 Norrköping, Sweden

Abstract

It is well-known that the knowledge of the natural geometry of a problem is often crucial in finding solutions. Problems involving functions on a circle are, for example, often solved using the theory of Fourier series. This is the mathematical explanation of the enormous success of the DFT, FFT and DCT-based methods. Another example is the relation between scaling properties and wavelet theory.

In this paper we show that spaces of spectral distributions, like color stimuli, have a natural cone-like structure. We use the framework of the Karhunen-Loève transform in a Hilbert space context to describe this cone-like structure and demonstrate how to compute natural coordinate systems from empirical data, like multi-spectral measurements and images. We will illustrate the theoretical findings with databases consisting of collections of multi-spectral measurements of color chips from color systems like Munsell, NCS and Pantone, multi-channel images of natural scenes, satellite data and daylight spectra.

We will also comment on the possible application of group theoretical methods in color science based on those findings.

1. PCA for spectral distributions

Color spectra are by definition non-negative functions $s(\lambda)$ where λ is the wavelength variable with values in the visible range of the spectrum. Typically λ takes values in the interval: $400nm \leq \lambda \leq 800nm$. The space of such spectra is an infinite-dimensional Hilbert space and therefore one can describe these spectra by the coefficients σ_n in a series expansion:

$$s(\lambda) = \sum_{n=0}^{\infty} \sigma_n b_n(\lambda) \quad (1)$$

When the basis functions b_n form an orthonormal system then the expansion coefficients σ_n are given by the scalar products $\langle s, b_n \rangle$ in the Hilbert space. In the following we will assume that the scalar product can be computed by an integral (which is not necessarily the ordinary integral)

$$\langle s_1(\lambda), s_2(\lambda) \rangle = \int s_1(\lambda) s_2(\lambda) d\mu(\lambda) \quad (2)$$

In many applications only a few coefficients are used to give a description of the spectrum s by the finite vector

$(\sigma_0, \sigma_1, \dots, \sigma_N)$. The mapping

$$s \mapsto s_N = \sum_{n=0}^N \sigma_n b_n = \sum_{n=0}^N \langle s, b_n \rangle b_n \quad (3)$$

is a projection from the Hilbert space to the $(N+1)$ -dimensional subspace spanned by the basis functions. When we consider the spectra as realizations of a stochastic process then we write $s = s(\lambda; \omega)$ where ω is the stochastic variable. In this case also the coefficients are random variables: $\sigma_n = \sigma_n(\omega)$. For each realization ω the projection $s(\omega) \mapsto s_N(\omega)$ gives an approximation error

$$\|\epsilon(\omega)\|^2 = \sum_{n=N+1}^{\infty} \sigma_n(\omega)^2 \quad (4)$$

that depends on the selected basis. The minimization of the mean squared approximation error is equivalent to maximization of the mean squared norm of the approximation vector: $\int \sum_{n=0}^N \sigma_n(\omega)^2 d\omega$. For the one-dimensional approximation with one basis function $b_0(\lambda)$ the mean of the squared norm is given by the double integral:

$$\begin{aligned} & \int \int \int s(\lambda_1, \omega) b_0(\lambda_1) s(\lambda_2, \omega) b_0(\lambda_2) d\mu(\lambda_1) d\mu(\lambda_2) d\omega \\ &= \int \int K(\lambda_1, \lambda_2) b_0(\lambda_1) d\mu(\lambda_1) b_0(\lambda_2) d\mu(\lambda_2) \\ &= \langle b_0, K b_0 \rangle \end{aligned} \quad (5)$$

Here K is the autocorrelation function

$$K(\lambda_1, \lambda_2) = \int s(\lambda_1, \omega) s(\lambda_2, \omega) d\omega \quad (6)$$

It defines an integral operator given by

$$f \mapsto (Kf)(\lambda_2) = \int K(\lambda_1, \lambda_2) f(\lambda_1) d\mu(\lambda_1) \quad (7)$$

The mean of the squared norm of the approximations is maximum if the basis vector b_0 is the eigenfunction of the integral operator $K : K b_0 = c_0 b_0$ with the largest eigenvalue c_0 . In the same way it can be shown that the other eigenfunctions of K are the basis functions that maximize the approximation length. This expansion is known as the Karhunen-Loève Transform (KLT) or Principal Component Analysis (PCA). A more detailed description of the KLT in a Hilbert Space setting can be found in [5, 8].

This selection of the basis functions as the eigenfunctions of the auto-correlation function is the usual description of Principal Component Analysis of color spectra (see

for example [4]). For stochastic processes of spectral distributions it was often observed that the first eigenvector is a positive function. The expansion coefficient for the first eigenfunction is therefore non-negative and the remaining eigenfunctions must have negative values since they are orthogonal to the first eigenfunction. The coefficients of all spectra that are elements of the finite-dimensional vector space spanned by the first N eigenfunctions satisfy therefore the inequality:

$$\sigma_0^2 - \gamma_1 \sigma_1^2 - \dots - \gamma_N \sigma_N^2 \geq 0 \quad (8)$$

with constants γ_n that depend only on the basis functions (independent of the spectra). For all databases investigated so far these constants were given by $\gamma_k = 1$. We will summarize this fact by saying that the spectra are located in a cone. This fact has been verified for a large set of different multispectral distributions. For more details the reader is referred to [6, 7].

From an arbitrary series expansion of the spectra (6) and (1) we find:

$$\begin{aligned} K(\lambda_1, \lambda_2) &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \int \sigma_n(\omega) \sigma_m(\omega) d\omega b_n(\lambda_1) b_m(\lambda_2) \\ &= \sum_{n,m} \gamma_{nm} b_n(\lambda_1) b_m(\lambda_2) \end{aligned} \quad (9)$$

When the basis functions b_k are eigenfunctions of the integral operator with kernel K we find that the expansion coefficients are uncorrelated: $\gamma_{nm} = 0$ for $n \neq m$.

An alternative to the PCA based on the auto-correlation function of the stochastic process $s(\lambda; \omega)$ is the PCA derived from the centered new process $t(\lambda; \omega) = s(\lambda; \omega) - m(\lambda)$ where $m(\lambda)$ is the mean of the original process s . If $t \mapsto t_N$ is the projection defined by the auto-correlation basis functions $b_n^{(t)}$ of the centered process t then we get the approximation of the original process s as:

$$\begin{aligned} s - m &= t \approx \sum_{n=0}^N \langle t, b_n^{(t)} \rangle b_n^{(t)} \\ &= \sum_{n=0}^N \left(\langle s, b_n^{(t)} \rangle - \langle m, b_n^{(t)} \rangle \right) b_n^{(t)} \\ &= \sum_{n=0}^N \langle s, b_n^{(t)} \rangle b_n^{(t)} - m_N \end{aligned} \quad (10)$$

$$\text{or } s \approx m - m_N + \sum_{n=0}^N \langle s, b_n^{(t)} \rangle b_n^{(t)}$$

In general nothing can be said about the differences between the two approximations (10) and (3). In the spectral databases we investigated we found however that the mean was identical to a constant multiple of the first eigenvector. In this case both approaches result in the same eigenfunctions. To see this write the spectra s as sum of the mean and a centered process t : $s = m + t = \hat{\sigma} b_0 + t$. For the expansion of t we get $t = (\sigma_0 - \hat{\sigma}) b_0 + \sum_{n=1}^{\infty} \sigma_n b_n$.

We denote the auto-correlation functions of s and t by K_s and K_t . These functions have an expansion of the type (9) with coefficients $\gamma_{nm}^{(s)}$ and $\gamma_{nm}^{(t)}$ respectively. From the expansion of the processes s and t it follows that $\gamma_{nm}^{(s)} = \gamma_{nm}^{(t)}$ for all pairs $(n, m) \neq (0, 0)$. The eigenfunctions of the operators given by K_s and K_t are thus identical.

A summary of the basic facts described so far is as follows

- The first eigenvector of the auto-correlation operator is positive everywhere.
- The mean is a multiple of the first eigenvector
- The coefficient vectors of the eigenvector expansion are all located in a cone.
- The auto-covariance and the auto-correlation function have identical eigenvectors.

2. Intensity and Chromaticity

The first eigenfunction is equal to the mean and positive everywhere. The first coefficient in the eigenfunction expansion of a spectral distribution is the scalar product of the spectral distribution and the first eigenfunction. It follows that the first expansion coefficient is a measure of the intensity of the spectral distribution. The remaining coefficients describe the chromaticity of the distribution and from the general theory follows that the intensity coefficient and the chromaticity coefficients are uncorrelated. Given an N -dimensional approximation $s \approx \sum_{n=0}^N \sigma_n b_n$ it is natural to describe s by the intensity coefficient σ_0 and the intensity-independent vector

$$\left(\frac{\sigma_1}{\sigma_0}, \dots, \frac{\sigma_N}{\sigma_0} \right) \quad (11)$$

The space formed by these projected vectors will be called the chromaticity plane or hyperspace. Since the mean was a multiple of the first eigenfunction it is also clear that the mean is projected to the origin in the chromaticity hyperspace. In most cases the mean is approximately a constant function of the wavelength. It is therefore natural to introduce polar coordinates in the chromaticity hyperspace and identify the radius with the saturation and the direction with the hue. For a two-dimensional chromaticity space the hue corresponds to an angle.

Geometrically it is also clear that the monochromatic distributions (ie. the distributions that are non-zero only for one wavelength) are at the boundary of the space of spectral distributions. Expanding the monochromatic distribution located at λ we get in (11) the chromaticity coefficient $b_j(\lambda)/b_0(\lambda)$. This describes for each value of j a function of λ and for $N = 2$ this defines a curve

$$\left(\frac{b_1}{b_0}(\lambda), \frac{b_2}{b_0}(\lambda) \right)$$

which corresponds to the spectral locus.

Next we characterize the coefficient vectors which lead to valid spectral distributions, ie. linear combinations with non-negative function values. We consider the case $N = 2$ and $\sigma_0 = 1$ for which we get the combination

$$b_0(\lambda) + \sigma_1 b_1(\lambda) + \sigma_2 b_2(\lambda)$$

which must be non-negative for all λ . This is equivalent to

$$\tilde{s}(\lambda) = 1 + \sigma_1 \frac{b_1(\lambda)}{b_0(\lambda)} + \sigma_2 \frac{b_2(\lambda)}{b_0(\lambda)} \geq 0 \quad (12)$$

In the coefficient space we introduce polar coordinates

$$(\sigma_1, \sigma_2) = \rho (\cos(\psi), \sin(\psi))$$

The curve where the reconstructed $\tilde{s}$ is zero for some λ is now constructed in two steps: First we compute for a given value of ψ the λ_0 for which

$$q(\lambda, \psi) = \cos(\psi) \frac{b_1(\lambda)}{b_0(\lambda)} + \sin(\psi) \frac{b_2(\lambda)}{b_0(\lambda)}$$

is minimal. From this result the value of ρ can be computed. A necessary condition for a minimum point is

$$\frac{\partial q(\lambda, \psi)}{\partial \lambda}(\lambda_0, \psi) = 0$$

which has to be solved for λ_0 . For the case where b_0 is constant this leads to the equation $\tan \psi = -b'_1(\lambda)/b'_2(\lambda)$ where $b'_j = \frac{db_j(\lambda)}{d\lambda}$. In our investigations we saw that values of ψ for which b'_1/b'_2 had significant singularities correspond to corner points of the spectral locus.

3. Fourier Series, Color Atlas and Multispectral Images

As an example where we can do all the necessary calculations analytically we consider the model where we assume that the spectra are functions defined on a circle thus gluing the ends of the wavelength interval together. The wavelength variable λ describes now an angle. A spectral distribution is then a realization of a stochastic process and if we assume that this stochastic process is invariant under shifts of the underlying circle then it can be shown that PCA leads to an expansion of the spectral distributions in Fourier series. In this case $b_0(\lambda)$ is a constant and $b_1(\lambda) = \cos(\lambda)$, $b_2(\lambda) = \sin(\lambda)$. Both, the spectral locus and the border of the parameter region leading to non-negative reconstructed distributions, are circles in this case. The space of 3-D coefficient vectors that result in non-negative reconstructed distributions is thus a cone.

Next we illustrate the results with experiments using a database consisting of 1750 spectra measured from the NCS color system in the range 400nm to 700nm in 10nm steps. In Figure 1 illustrate the similarity between the first

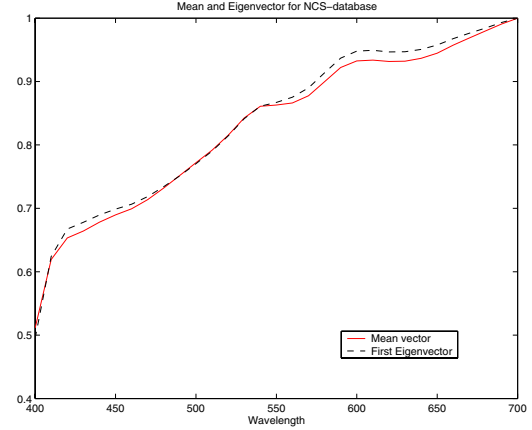


Figure 1: Scaled Mean Vector and First Eigenvector

eigenvector of the correlation matrix and the mean vector. Both were normalized so that their maximum value is equal to one. Similar relations between the mean vector and the first eigenvector were obtained from databases based on color systems like the Munsell and the Pantone system, from satellite and collections of daylight spectra.

Figure 2 illustrates some of the other properties of the spectra in the NCS database. Here we computed first the first three eigenvectors of the spectra and projected the spectra to the three-dimensional space spanned by these eigenvectors. Then we used the perspective projection in Equation (11) to compute two dimensional chromaticity vectors. The single points in the Figure are the coordinate vectors of the chips in the database. The inner, solid curve describes the boundary of the coefficient space which lead to non-negative reconstruction vectors. The outer, dashed curve is the spectral locus, ie. the location of the projected mono-chromatic spectra. From the Figure it can be seen that nearly all (except the red points in the upper, right corner) coordinate vectors lead to non-negative reconstructions and that all mono-chromatic spectra are outside the non-negativity region.

Also shown in the Figure are the projections of the three photopigment optical density curves (see [12, 11]). The location of these three vectors are shown as solid red (L-cone), green (M-cone) and blue (S-cone) dots. Next we converted the cartesian coordinates of these points to polar coordinates (where the origin was the projection of the first eigenvector) and adjusted their radial value such that the scaled points were located on the curve bounding the positivity region. These projected points are shown as the red, green and blue * in the Figure. When we reflected the projected green point on the origin we obtained the point marked as opp-green in the Figure. We see that the three points marked as "red", "opp-green" and "blue" are very near the corners of the triangle-shaped region.

Essentially the same results can be obtained from other color systems. We used another database measured from

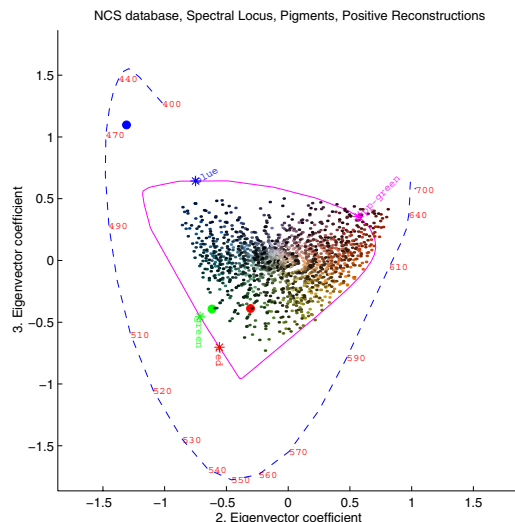


Figure 2: Properties of the NCS-database

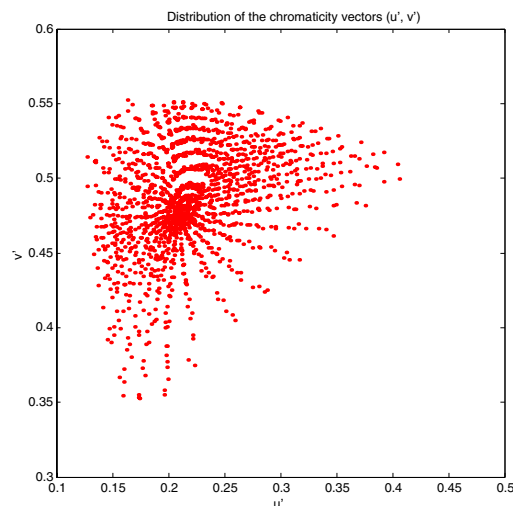
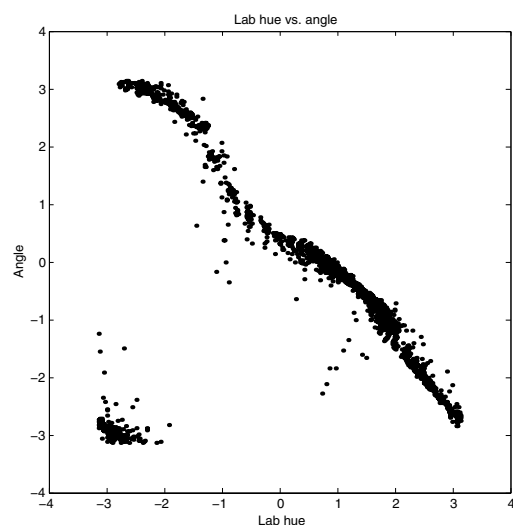
the NCS-system, a database measured from the Pantone system and two databases measured from the chips in the Munsell color atlas. All of them show essentially the same features.

We also investigated the properties of 22 multispectral images, 10 taken in coral reefs and 12 in forests (for a detailed description of the images see [2] and [3]). From these images we computed two eigenvector systems: one from the coral and one from the forest images. We then described the spectra from these images in both the Munsell-NCS coordinate system and in the system computed from the spectra in the same class.

For the forest images we find that only three spectra did not satisfy the inequality (8). For the underwater coral images the corresponding numbers are not as small as in the forest case but also here there are only around 200 exceptions (out of 10×128^2). Many of these outliers have a very low intensity.

Another dataset used data from the Global Ozone Monitoring Experiment (GOME). The spectrometer mounted on the satellite measures the solar radiation scattered by the atmosphere. In our experiments we extracted data from two detectors (with 1024 detector pixels each) and interpolated the measurements so that the final spectra represented the satellite data in the range 407nm to 794nm with 1nm sampling. We rejected spectra which were obviously flawed and collected finally 59975 spectra from 31 original files. We found that 619 spectra did not satisfy (8). An inspection of these outliers showed that almost all of them were measured at consecutive time steps. For all of them the data in the third band had very low values whereas the data vectors from the fourth band had normal values. We therefore concluded that these spectra were obtained while one of the sensors did not work properly. The number of points not fulfilling inequality (8) mentioned so far is com-

puted from a three term approximation of the spectra. We found however that the corresponding inequality also holds for higher order approximations. For the NCS database with 1750 spectra only the sixth- and seventh order approximation of one spectral distribution do not satisfy (8).

Figure 3: Chromaticity in u' , v' 

polar coordinate system introduced in the projection plane and the polar coordinates in the standard CIE-(a,b) chromaticity coordinates. The relation between the two angles (related to the hue of the color spectrum) is shown in Figure 4. Most of the spectra for which the linear relation between the angles do not hold are achromatic. For these colors the angular value is not very meaningful.

For the radial variable, roughly corresponding to the saturation of a spectral distribution, the correlation is not as strong as can be seen from Figure (5). Analyzing the data further shows however that the relation between these radial values are almost linear when only colors in a restricted hue sector are considered. The proportionality factor varies however with the hue sector.

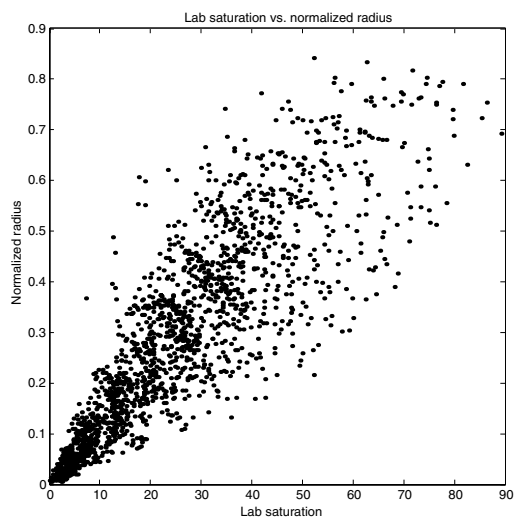


Figure 5: Comparison of saturation radii in (a,b) and projection

4. Group theoretical methods

A key observation of this paper is the relation between the non-negativity of the spectral distributions and the inequality in Equation (8) which implies that the space of spectral distributions has a conical structure. It has been mentioned before that the shift-invariance of stochastic processes lead to Fourier series and Fourier transformations as solutions to the corresponding PCA-approximations and also to solutions to invariant recognition strategies. Given the fact that the spectral distributions form a cone-like space it is therefore tempting to assume that the natural transformations operating on color spectra are the elements of the motion group of the cone. This leads to the hypothesis that the Lorentz-group is the natural transformation group of color space. The relevance of group theoretical methods in general, and the Lorentz group in particular, for color processing is not new. A few references are [13, 1] and [9]. All of these approaches are however based on facts on human color perception. In this paper we showed that the

positivity of the spectral distributions is enough to give the space of spectral distributions a cone-like structure. Before we sketch some potential consequences of these results it seems necessary to mention one important open problem that has not been investigated yet. Up to now we have shown that the spectral distributions are contained in a space that has a cone-like structure, that the boundary of this space is formed by the mono-chromatic spectra and that the main axis is given by the achromatic distributions which are scalar multiples of the mean of the databases which is equal to the first eigenvector of the correlation function. A basic idea in the theory of groups is however the definition of a group as a set of linear transformations that leave some distance measure invariant: Euclidean motions are those linear transformations that preserve the Euclidean distance in a vector space and the Lorentz-group consists of those linear transformations that preserve the hyperbolic distance in a cone. In our application we have only shown that the coordinate vectors of the spectral distributions are located within a cone-like part of ordinary Euclidean space. Nothing has been said about a natural distance measure between two such coordinate vectors.

From the inequality (8) we have:

$$\sigma_0^2 - \gamma_1 \sigma_1^2 - \dots - \gamma_N \sigma_N^2 \geq 0$$

and it is therefore possible to define $\sigma_0^2 - \gamma_1 \sigma_1^2 - \dots - \gamma_N \sigma_N^2$ as the norm of the coordinate vector $(\sigma_0, \sigma_1, \dots, \sigma_N)$ but we feel that such a definition requires further justification.

If we assume that the space of spectral distributions possesses a group of natural transformations then we can use the powerful tools from the theory of groups to solve problems in color science. As one example we mention briefly one application in which dynamical illuminations changes are described in the framework of Lie-groups (see [6] for details). The basic idea behind this approach is the observation that each illumination spectrum can be described by a coefficient vector that is an element in the cone defined by the spectral distributions. When the spectral characteristics of the illumination changes as a function of time these coordinate vectors describe a curve in the cone. Under some assumptions it can be shown that the transformations $T(t)$ that move the coordinate vector at a given time $t=0$ to the coordinate vector at time t form a subgroup of the group of natural transformations of the cone. Since the elements in this group depend on the time-variable t alone it is seen that these transformations $T(t)$ form a one-parameter subgroup of the transformation group.

The illumination spectra interact with the spectral reflection distributions of the objects in the scene and the spectral characteristics of the sensors to produce a measurement vector. A typical measurement vector is the RGB vector at a pixel in an image. Under an illumination change these measurement vectors change and in that way the one-parameter group $T(t)$ operates on the measurement vectors. A basic task in the interpretation of these images is now to estimate the characteristics of the illumination change

from the measured pixel vectors. Formulating the problem in this abstract setting has the advantage that the general theory of Lie-theory and differential equations can be applied to estimate the characteristics of the transformation group $T(t)$ from the measurements.

This is an application where the assumption of a group theoretical structure of the problem leads to well-known solution strategies. Other applications were the conical and group theoretical framework might lead to new results are the design of new color edge detection methods and the design of new scale-space methods for color image processing, for example in the framework of the Beltrami-framework investigated by Sochen et.al. (see [10]).

5. Conclusion

We found that the space of spectral distributions has a natural conical geometry. We showed that many geometrical properties of representative databases of spectral distributions have direct correlates in color spaces that were derived from properties of human color perception alone. This suggests that human color perception is taking advantage of these geometrical properties of the space of color stimuli.

In the second part of the paper we sketched some possible applications of group theoretical methods in color science. These include the study of Lorentz-transformations that have been used earlier in color science.

Acknowledgements

The databases with the Munsell-spectra and Pantone spectra were supplied by the Computer Science and Physics Departments at the University of Joensuu, Finland.

The NCS database was obtained from the Scandinavian Color Institute. The multi-spectral images are described in [2, 3].

The optical densities of the pigments were downloaded from the CVRL-database as file ssabance_1.txt from the site: <http://cvision.ucsd.edu>.

The GOME spectra were supplied by ESA under contract number 1098.

The daylight spectra were collected and investigated by Prof. Javier Hernández-Andrés, Departamento de Óptica, Universidad de Granada, Spain.

The financial support of CENIIT (Center for Industrial Information Technology, Linköping University) and TFR (Swedish Research Council for Technical Research) is gratefully acknowledged.

References

- [1] M. Brill and G. West. Group theory of chromatic adaptation. *Die Farbe*, 31(1-3):4–22, 1983/84.

- [2] C.C. Chiao, T.W. Cronin, and D. Osoria. Color signals in natural scenes: characteristics of reflectance spectra and effects of natural illuminants. *J. Opt. Soc. Am. A*, 17(2):218–224, Feb. 2000.
- [3] C.C. Chiao, D. Osoria, 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(10):1713–1721, Oct. 2000.
- [4] T. Jaaskelainen, J. Parkkinen, and S. Toyooka. Vector subspace model for color representation. *Journal of the Optical Society of America A*, 7(4):725–730, April 1990.
- [5] R. Lenz. *Group Theoretical Methods in Image Processing*. Lecture Notes in Computer Science (Vol. 413). Springer Verlag, Heidelberg, Berlin, New York, 1990.
- [6] Reiner Lenz. Estimation of illumination characteristics. *IEEE Transactions Image Processing*, 10(7):1031–1038, July 2001.
- [7] Reiner Lenz. A geometric foundation of colorimetry. In *Ninth Color Imaging Conference*, pages 46–51. IS&T/SID, 2001.
- [8] A. Levy and J. Rubinstein. Hilbert-space Karhunen-Loève transform with application to image analysis. *J. Opt. Soc. Am. A*, 16(1):28–35, January 1999.
- [9] H. L. Resnikoff. Differential geometry and color perception. *J. Mathematical Biology*, 1:97–131, 1974.
- [10] N. Sochen, R. Kimmel, and R. Malladi. A general framework for low level vision. *IEEE-Trans. IP*, 7(3):310–318, 1998.
- [11] A. Stockman and L. T. Sharpe. Spectral sensitivities of the middle- and long-wavelength sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40:1711–1737, 2000.
- [12] A. Stockman, L. T. Sharpe, and C. C. Fach. The spectral sensitivity of the human short-wavelength cones. *Vision Research*, 39:2901–2927, 1999.
- [13] H. Yilmaz. *On Color Perception*, volume XX of *Int. School Physics, Enrico Fermi*, pages 239–251. Academic Press, 1962.

Biography

Reiner Lenz is associate professor at the Department of Science and Technology, Campus Norrköping, Linköping University, Sweden. He received a diploma in mathematics from the Georg-August University in Göttingen, Germany and a PhD-degree in computer engineering from Linköping University. From 1979 to 1982 he was with the Department of Electrical Engineering, Stuttgart University, Germany. From 1983 to 1997 he was with the Image Processing Laboratory, Department of Electrical Engineering, Linköping University, Sweden. He held positions as invited researcher at the ZEISS Central Research and Development Laboratory, Oberkochen, Germany, the Advanced Telecommunication Research Institute (ATR), Kyoto, Japan, the Mechanical Engineering Laboratory, Tsukuba, Japan and at Rutgers University, USA. He received an honorable mention for the Pattern Recognition Society Award and the SAAB-Combitech Award for his work in image processing. He is interested in the application of group-theoretical methods in signal processing and color and multi-spectral image processing.