

The Hyperbolic Geometry of Illumination-Induced Chromaticity Changes

Reiner Lenz

Dept. Science and Technology, Linköping University, Norrköping, Sweden, reile@itn.liu.se

Pedro Latorre Carmona

Comp. Languages and Systems, Jaume I University, Castellón, Spain, latorre@lsi.uji.es

Peter Meer

Electrical and Computer Eng., Rutgers University, Piscataway, USA, meer@caip.rutgers.edu

Abstract

The non-negativity of color signals implies that they span a conical space with a hyperbolic geometry. We use perspective projections to separate intensity from chromaticity, and for 3-D color descriptors the chromatic properties are represented by points on the unit disk. Descriptors derived from the same object point but under different imaging conditions can be joined by a hyperbolic geodesic. The properties of this model are investigated using multichannel images of natural scenes and black body illuminants of different temperatures. We show, over a series of static scenes with different illuminants, how illumination changes influence the hyperbolic distances and the geodesics. Descriptors derived from conventional RGB images are also addressed.

1. Introduction

Processing of color images is difficult and has received less interest in the past than gray-value image processing. A fundamental problem of color signal processing is the fact that the output signals generated by the sensors (the pixel vectors) are the result of a complex interaction between the properties of the illumination source, the object reflection and the sensors sensitivity properties. Understanding these processes is, however, useful in a great number of applications, ranging from better color correction in cameras, over improvements in automatic inspection tasks to a better understanding of the human visual system.

In this paper we develop a geometrical theory of color signal processing (see also [1, 3, 4, 14] and [5] for a comprehensive historical overview over geometrical models of color spaces). We observe that all spectral distributions involved in color imaging describe energies, probabilities or sensitivities. They can therefore only assume non-negative

function values. This implies that these functions are always located in the positive part of the corresponding function space. These regions in the function spaces have a special geometry and admissible transformations of these functions can never move the resulting functions outside these regions. Based on this observation we argue that the function spaces of interest in color signal processing can be described by direct products of the positive-axis (representing intensity) and the unit ball (describing chromaticity-related properties). The properties related to intensity are investigated by gray-value-based image processing and here we concentrate on the study of the chromaticity part.

The natural geometry for two-dimensional chromaticity descriptors is the hyperbolic geometry of the unit disk with the pseudo-unitary group $SU(1,1)$ as symmetry group (an early application of group theory in color is [1]). It was earlier shown [8, 9, 10] that illumination induced chromaticity changes can be described by curves generated by one-parameter subgroups of $SU(1,1)$. The methods used to estimate these curves from measurements require however that at least three points on these curves are known (in practice more than three since they are based on regression methods). In the following we will describe descriptors, derived from the straight lines in hyperbolic geometry, that only require two measurements. The straight lines in this geometry are the Euclidean circles meeting the bounding unit circle at perpendicular angles and they can therefore be parameterized by the two points where they cross the unit circle. The space of all hyperbolic lines can thus be identified with the torus.

We will show that the chromaticity descriptors of a scene point generated by illumination changes define a curve on the unit disk and that these curves can be approximated by hyperbolic lines. In Section 2 we introduce the basic conical model. Section 3 describes the application of hyperbolic geometry to color changes and in Section 4 we show in our experiments that the descriptors derived from purely geo-

metrical properties are useful for separating the influence of the illumination changes from the reflection properties of the scene points. Most of our results are derived from images captured with multispectral cameras but we will also show that RGB images captured by conventional cameras contain sufficient information to estimate the same results.

2. The conical model of color signal processing

An illumination spectrum or a spectral power distribution $l(\lambda)$ is a non-negative function (or vector) defined on the wavelength interval of interest (usually in the range 400nm to 700nm with 10nm sampling). The value $l(\lambda)$ describes the emission of photons of wavelength λ from the light source and can only assume non-negative values. The optical properties of materials are characterized by their reflection spectra r where $r(x, \lambda)$ is the probability that a photon of wavelength λ is reflected from point x of the object. We have $0 \leq r \leq 1$. The product $s(x, \lambda) = r(x, \lambda) \cdot l(\lambda)$ defines the color signal and describes the number of photons of wavelength λ reflected from point x . The spectral sensitivity functions c of the camera or sensors describe the contribution of an incoming photon of wavelength λ to the final output signal of the sensor. We use the simplest model to describe color imaging. In this model the output $p(x)$ of the sensor pointed at position x in the scene is given by

$$p(x) = \int l(\lambda)r(x, \lambda)c(\lambda) d\lambda \quad (1)$$

where integration is over the relevant wavelength interval and effects like fluorescence and all dependencies on the geometrical relations between the components involved are ignored.

The key components of the conical model are now explained, first in a mathematical Hilbert or vector space setting, and then they are translated into concepts from color image processing and color science. All functions $l(\lambda), r(x, \lambda), c(\lambda)$ used in the model can only assume non-negative values. The set of all such functions can therefore never span the full function (or vector) space. Instead these functions define a conical subspace, ie. a space closed under addition and scalar multiplication with non-negative constants. Some properties of these conical spaces are crucial:

In a space consisting of non-negative functions it is possible to select an orthogonal basis $b_0(\lambda), b_1(\lambda), b_2(\lambda), \dots$ with the following properties.

1. A non-negative function s has an expansion

$$s(\lambda) = \alpha_0 b_0(\lambda) + \alpha_1 b_1(\lambda) + \alpha_2 b_2(\lambda) + \dots \quad (2)$$

2. $b_0(\lambda)$ has strictly positive values and therefore $\alpha_0 > 0$ for non-zero functions s ;

3. If we break the series expansion after $K + 1$ terms then we can find a constant C_K such that

$$\alpha_1^2 + \alpha_2^2 + \dots + \alpha_K^2 < C_K \alpha_0^2 \quad (3)$$

After applying a perspective projection $\xi_k = \frac{\alpha_k}{\alpha_0}$ we get

$$\xi_1^2 + \xi_2^2 + \dots + \xi_K^2 < C_K \quad (4)$$

For color image processing the basis is often derived by Principal Component Analysis (PCA) from representative collections of spectral data such as the Munsell or NCS color atlas. In these cases it was shown [8] that the first eigenvector b_0 is proportional to the mean vector. The coefficient α_0 is computed as the scalar product of the spectral distribution s to be analyzed and b_0 , it is thus a weighted average of s . Since both, s and b_0 are non-negative, we get $\alpha_0 > 0$. The ratios $\xi_k = \frac{\alpha_k}{\alpha_0}$ are well-defined and independent of scalings of the original signal s . The ξ_k are constant under intensity changes of the original color signal and thus descriptors of its chromatic properties.

3. Hyperbolic Geometry and Color Changes

From the inequality in Eq. (4) follows that the vectors $(\xi_1, \xi_2, \dots, \xi_K)$ are points inside a K -dimensional sphere. In the rest of the paper we will restrict us to the case $K = 2$, write $x = \xi_1, y = \xi_2$ and refer to the values α_0, x, y as color descriptors.

Restriction to three descriptors is motivated by mathematical simplicity, the trichromatic structure of normal human color vision and, as will be demonstrated by our experiments, by its close relation to the three-dimensional RGB-vectors.

Mathematically the simplest conical structures of interest are three-dimensional, and the 3-D case illustrates many of the characteristic properties of non-euclidean geometry that will be of interest in the rest of the paper. Most color systems related to human color vision are three-dimensional mainly because the human eye contains three different types of sensors. In [7] it was shown that there is a close relation between the geometry of the color signal space and the CIELAB coordinate system, derived from color matching experiments of human observers. There it was demonstrated that if the basis in the color signal space is derived by PCA from a color atlas, then the angle $\arctan(x, y)$ correlates to the hue-angle in the CIELAB coordinate system, and the radius $\sqrt{x^2 + y^2}$ is related to the saturation. In the experiments described below, we will also show that the RGB-vectors of a conventional camera can be used to estimate the values of these descriptors.

In the following we assume that the axes are scaled such that $x^2 + y^2 < 1$ for all points. We also identify (x, y) with $z = x + iy$. In the complex plane the unit circle

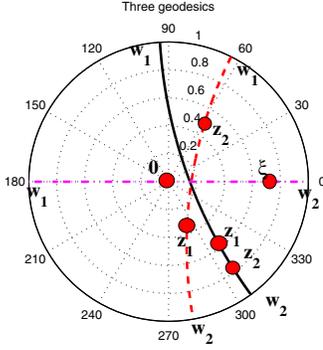


Figure 1. Three hyperbolic straight lines

is $\mathfrak{C} = \{z \in \mathbb{C} : |z| = 1\}$ and $\mathfrak{D} = \{z \in \mathbb{C} : |z| < 1\}$ is the unit disk. The metric on \mathfrak{D} is introduced as follows (see [6, 15] for more information on the disk model of hyperbolic geometry): consider a curve $\gamma(t) = x(t) + iy(t)$ on the disk. Differentiating this curve gives tangent vectors $\gamma'(t)$ and the length of the curve segment between parameter values 0 and 1 is (for details see [15], Chapter 3):

$$L(\gamma) = 2 \int_0^1 \frac{\sqrt{x'(t)^2 + y'(t)^2}}{1 - x(t)^2 - y(t)^2} dt \quad (5)$$

where the conventional euclidean metric is changed by a correction factor $(1 - |z|^2)^{-1}$. Euclidean distances grow thus without bound the nearer the points come to the boundary (given by the unit circle \mathfrak{C}).

A geodesic (or a hyperbolic straight line) is the shortest curve between two points. For the metric defined in Eq. (6) the geodesics are those Euclidean circle arcs passing through z_1, z_2 that are orthogonal to \mathfrak{C} . If $z_1, z_2, 0$ are collinear then the geodesic is the diameter. Furthermore if w_1, w_2 are the points where the geodesic through z_1, z_2 intersects the boundary circle \mathfrak{C} , then the distance between the points z_1, z_2 is given by

$$d(z_1, z_2) = 2 \operatorname{atanh} \left| \frac{z_1 - z_2}{1 - \bar{z}_1 z_2} \right| = \ln(w_1, w_2, z_1, z_2) \quad (6)$$

where (w_1, w_2, z_1, z_2) is the cross ratio

$$(w_1, w_2, z_1, z_2) = \frac{(z_2 - w_1)(z_1 - w_2)}{(z_2 - w_2)(z_1 - w_1)} \quad (7)$$

Three examples of geodesics are shown in Figure 1. For a geodesic we show two points z_1, z_2 on the disk and their corresponding boundary points w_1, w_2 .

The conical structure of the space ensures that for every triple (α, x, y) and all positive constants $\gamma > 0$ the descriptors $(\gamma\alpha, x, y), \gamma > 0$ are elements in this space. For fixed (x, y) the points $(\gamma\alpha, x, y), \gamma > 0$ define a ray. We call such a ray a chromatic-ray since it represents all colors with constant chromatic properties but varying intensity. The positivity of the spectral distributions imply that

all chromatic rays pass through the unit disk (after suitable scaling of the axes). This geometrical structure is identical to models in computer vision that describe cameras by perspective projections [2, 11]. In these models the plane containing the unit disk represents the sensor array, and the projection rays connect the pixels in the image to the points in the scene.

In the space of chromatic rays, represented by points on the unit disk, we select the group of transformations that leave the hyperbolic distance invariant. All such mappings are specified by complex numbers a, b with $|a|^2 - |b|^2 = 1$. It operates on the unit disk as follows. Define first the matrix

$$M = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \quad (8)$$

and for such a matrix the fractional linear transform:

$$Mz = \frac{az + b}{\bar{b}z + \bar{a}} \quad \text{for } z \in \mathfrak{D} \quad (9)$$

These matrices define the (Lorentz) group $SU(1,1)$ and act on the unit disk. Occasionally we will use the shorthand notation $M_{(a,b)}$ to specify the matrix defined in Eq.(8). For constants a, b with $|a|^2 - |b|^2 \neq 1$ we define the convention that we first construct the matrix M of the same form as in Eq.(8) and then we enforce $\det M = 1$ by dividing with $\sqrt{|a|^2 - |b|^2}$.

These matrices form a group under ordinary matrix multiplication and matrix multiplication acts in the same way as a concatenation of the fractional linear mappings:

$$M_1 M_2 z = M_1 (M_2 z); z \in \mathfrak{D}; M_1, M_2 \in SU(1,1) \quad (10)$$

As mentioned above, these transforms are the motions of this geometry, ie. they preserve the distance in Eq.(6):

$$d(z_1, z_2) = d(Mz_1, Mz_2) \quad (11)$$

for all $z_1, z_2 \in \mathfrak{D}$ and all $M \in SU(1,1)$.

We now use these properties to show: for every pair $z_1, z_2 \in \mathfrak{D}$, there is an $M \in SU(1,1)$, such that the images of these points under this fractional linear mapping M are the origin and a positive real number: $Mz_1 = 0$ and $Mz_2 > 0$. We define first the matrix $L = M_{(1, -z_1)}$ (recall the determinant convention in connection with Eq.(8)) mapping the first point to the origin. By construction L is in $SU(1,1)$ and $Lz_1 = 0$. Now apply L to z_2 and get $w = Lz_2$. The matrices $M_{(e^{i\varphi}, 0)}$ are also in $SU(1,1)$ and act as ordinary rotations with rotation angle 2φ . We combine L and a rotation $K = M_{(e^{i\varphi}, 0)}$ to get

$$KLz_2 = Kw = \xi > 0 \quad (12)$$

The geodesic through 0 and ξ is the real line, the points where the geodesic crosses the unit circle are -1 and 1, and

therefore

$$\begin{aligned} d(z_1, z_2) &= d(\mathbf{KL}z_1, \mathbf{KL}z_2) \\ &= d(0, \xi) = 2\operatorname{atanh} \xi = \ln \left(\frac{1+\xi}{1-\xi} \right) \end{aligned} \quad (13)$$

For any two points z_1, z_2 we characterize their geometric relation by the hyperbolic distance between them, and the two points w_1, w_2 where the unique geodesic through the points crosses the unit circle. We compute these three parameters by applying the procedure just described.

1. Given two points z_1, z_2 , compute the matrices \mathbf{L}, \mathbf{K} such that z_1, z_2 are mapped to 0 and ξ .
2. The distance between z_1, z_2 is $\ln \left(\frac{1+\xi}{1-\xi} \right)$.
3. The points $-1, 0, \xi, 1$ are collinear and the geodesic through $0, \xi$ is thus the real axis. The points w_1, w_2 are obtained by applying the inverse $\mathbf{N} = (\mathbf{KL})^{-1}$ to $-1, 1$: $w_1 = \mathbf{N}(-1)$ and $w_2 = \mathbf{N}(1)$.

In the following we will investigate the relation between two descriptors $z_1, z_2 \in \mathcal{D}$, representing the same point in the scene under different imaging conditions. For geometrical reasons we choose the hyperbolic geodesics to describe their relation. In our experiments with real data we found that this hyperbolic model provides a good description of actual color changes.

4. Experiments

In our experiments we use three types of color measurements:

1. color signals computed from multichannel measurements of reflectance spectra and blackbody radiators and measured illumination spectra;
2. simulated RGB vectors obtained from measured reflectance spectra, illumination spectra and an estimation of the spectral sensitivity functions of a consumer SLR camera and
3. measured color signals from a scene under different illuminations.

In all our experiments we use the same PCA basis b_0, b_1, b_2 (see Eq.(2)) computed from the reflection spectra in the Munsell color atlas (described by Parkkinen et. al. in [13]) combined with 100 blackbody radiators with temperatures from 4000K to 10000K. The temperatures represent typical daylight distributions. We use the mired or reciprocal megakelvin scale to select them since human color vision is more adapted to changes in inverse color temperature. More information on blackbody radiators and the mired scale can

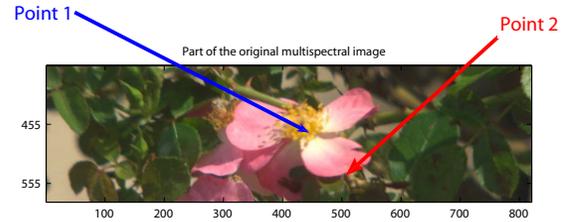


Figure 2. RGB image of a scene

be found in [17]. The spectra are sampled with 10nm resolution in the range from 400nm to 700nm resulting in 31-dimensional vectors. In the experiments we use the basis to compute the color descriptors α_0, x, y (as described in Section 3).

In the *first* series of experiments we use the reflectance spectra of natural scenes analyzed by Nascimento et. al. in [12]. One part of the RGB image of the scene is shown in Figure 2. A typical result obtained within the hyperbolic framework is shown in Figure 3. In this experiment we use two illuminants: blackbody radiators with 4000K and 10000K. The color signal is computed as the pointwise product between the reflection spectra of the object points and the illuminant. From these color signals the color descriptors were extracted. Chromatic changes are described by the geodesics and the hyperbolic distance. In Figure 3A the values of the angle belonging to the first intersection points w_1 of the geodesics with the unit circle are shown. Figure 3B contains the corresponding angular values of the second intersection points w_2 . The histogram of the distance values obtained is in Figure 3D. It shows that most distance values are located in the range 0.2 to 1. Very few object points generated distance values greater than one. This shows that the basis computed from the Munsell chips and the selected blackbody illuminants captures the chromaticity properties of the natural scene spectra very well. The pixel values in Figure 3C are proportional to the value of the hyperbolic distances truncated at value one.

Figure 4 contains a detailed analysis of two selected points representing object points with low and high hyperbolic distances in Figure 3. In this experiment we simulate 50 blackbody illuminants ranging from 4000K to 10000K (in the mired sampling) and compute the corresponding descriptors. The two selected points are located in the center and the boundary of the flower marked in Figure 2. The two dark curves in Figure 4A show the location of the original points. For two pairs of points from such a sequence we estimate the geodesics shown as light lines in the figure. The selected point pairs used to compute the geodesics are computed from the illuminants with 4500K and 5041K (for the first pair) and 5731K and 6746K (for the second geodesic).

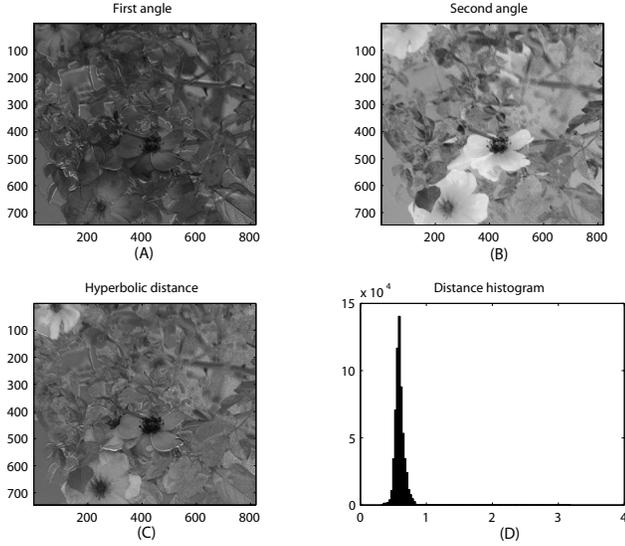


Figure 3. Natural scene and blackbody illuminants of 4000K and 10000K. The two angles (A) and (B), the hyperbolic distance (C) and the histogram of the distance values (D)

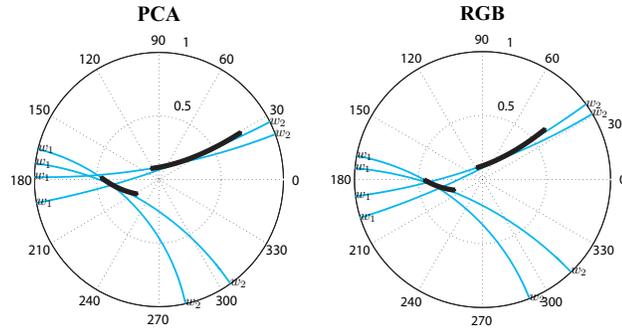


Figure 4. Chromaticity changes for two selected scene points (A) computed from PCA coefficients (B) computed from RGB images

Multichannel cameras are still expensive and difficult to handle. Therefore it is of great practical interest to see if the color descriptors can be estimated from conventional RGB images. The *second* experiment illustrates that this may be feasible, although RGB cameras and PCA-based basis functions describe different subspaces of the full space of color signals.

In this experiment we use the estimated sensitivity functions of a Canon-10D consumer camera (described by Solli et. al. in [16]) to compute simulated RGB images. We created color signals from the reflection spectra in the Munsell atlas and 100 blackbody radiators in the range from 4000K to 10000K (in mired sampling). For these color signals we computed both, the first three PCA coefficients in the selected basis and the simulated camera RGB vectors. From this collection of PCA/RGB vector pairs we obtained the 3×3 regression matrix that estimates the PCA coefficients from the RGB vectors. This regression matrix was then used to estimate the positions of the color descriptors on

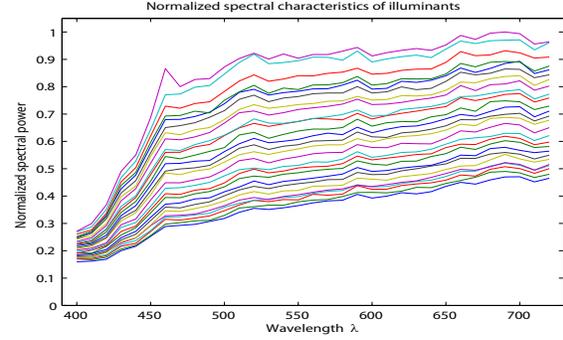


Figure 5. Spectral distribution of the 26 illuminants

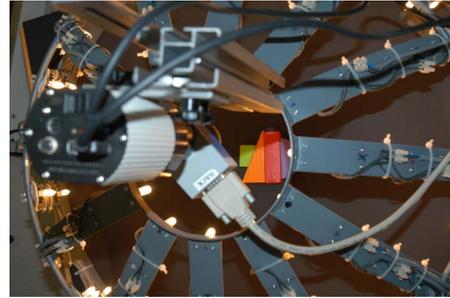


Figure 6. Multichannel illumination and image acquisition system

the unit disk. The estimates obtained from the coefficients and the RGB vectors are very similar as seen in Figure 4B showing the traces of the same two points as shown in Figure 4A.

In the *third* experiment we used a multichannel camera to obtain images of four wooden toys under the 26 different illuminants shown in Figure 5. The distributions were scaled with a common factor to the maximum value of one. Figure 6 shows the half-dome with the light sources, the camera mounted on top of the half-dome and the four wooden toys in the middle, located directly under the camera. The spectral resolution was from 400nm to 720nm in 10nm steps.

From this sequence of 26 images we computed the hyperbolic descriptors for all image pairs [Image(1) with Image(k)], $k = 2 \dots 26$. Figure 7A shows the hyperbolic distance values computed comparing the first and the last image in the sequence, ie. from the pair [(Image(1) - Image(26))]. The line marks the position of the points that will be selected in Figures 7B-D. In these figures the image pairs [Image(1) - Image(5)], [Image(1) - Image(11)] and [Image(1) - Image(26)] are used. Figure 7B shows the location of the points w_1 on the circle, Figure 7C the position of the points w_2 and Figure 7D the hyperbolic distance values.

From Figures 7B and 7C we see that the values for the angles are stable once the chromaticity points correspond to sufficiently different image pairs (Image(1) - Image(k)) with $k > 5$. The hyperbolic distance values, on the other

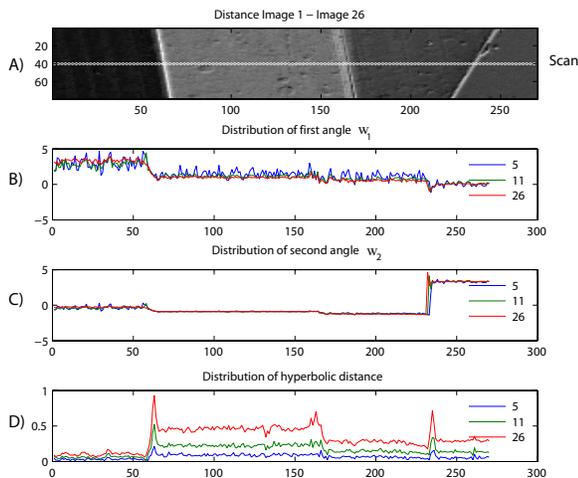


Figure 7. Evaluation of a series of illumination changes. A): Image of hyperbolic distances between the first and the last image in the sequence. B) Value of first angle. C) Value of second angle. D) Value of hyperbolic distance. See text for detailed explanation.

hand, are increasing monotonically. The interpretation in this case is that the chromaticity points for one material are approximately located on a hyperbolic straight line. We can thus see that in this experiment the material properties are encoded in the parameters of the hyperbolic line whereas the chromaticity changes of the illuminant is only changing the relative position of the points on such a line.

5. Discussion and Conclusion

We started from the simple observation that all wavelength dependent functions involved in color imaging have only non-negative function values. From this we concluded that chromaticity changes can be described by transformations of the unit disk. Using the natural hyperbolic geometry of the disk leads to the assumption that the geodesics curve segments are natural candidates for descriptors of the chromaticity changes defined by two points on the disk. In our experiments we used both, theoretical models and measured multispectral and RGB images, to demonstrate that this really is the case for chromaticity changes induced by varying illumination sources. The multispectral images indicate also that the properties of the material mainly defines the overall parameters of the chromaticity curve whereas the illumination changes lead to movements of the coordinate points along such a curve. This separation of the material properties and the illumination changes makes these methods suitable for such tasks as color-based segmentation, color correction and design of color invariants.

In this paper we described only the basic idea and some experimental results. One interesting question is the comparison between local and global estimation methods. The

method described here is local in the sense that only two points are used to compute the geodesic and the distance. Regression methods based on one-parameter subgroups [9] are global since they use the whole sequence of data points for the estimation of the curve. Global methods are usually superior since the curves are not limited to be straight lines. They are, however, unable to adapt to sudden changes like changes from daylight to a sunset. We also investigated the relation between the hyperbolic distance and the standard CIELAB distance used in color science. These results and applications to color based segmentation will be reported elsewhere.

References

- [1] M. Brill and G. West. Group theory of chromatic adaptation. *Die Farbe*, 31(1-3):4–22, 1983/84.
- [2] G. Csurka and O. Faugeras. Algebraic and geometric tools to compute projective and permutation invariants. *IEEE T-PAMI*, 21(1):58–65, 1999.
- [3] H. Grassmann. Zur Theorie der Farbenmischung. *Poggendorfs Annalen der Physik und Chemie*, 80:69–84, 1853.
- [4] J. J. Koenderink and A. J. van Doorn. The structure of colorimetry. *Proc. AFPAC, LNCS-1888*, pp. 69–77. 2000.
- [5] R. G. Kuehni. *Color space and its divisions: color order from antiquity to the present*. J. Wiley, Hoboken, NJ, 2003.
- [6] J. Lehner. *Discontinuous Groups and Automorphic Functions*. American Mathematical Society, Providence, USA, 1964.
- [7] R. Lenz. Two stage principal component analysis of color. *IEEE T-IP*, 11(6):630–635, June 2002.
- [8] R. Lenz and T. Bui. Statistical properties of color-signal spaces. *J. Opt. Soc. Am. A*, 22(5):820–7, May 2005.
- [9] R. Lenz, T. Hai Bui, and J. Hernandez-Andres. Group theoretical structure of spectral spaces. *J. Math. Im. Vis.2005*, 23(3):297–313, Nov. 2005.
- [10] R. Lenz and M. Solli. Lie methods in color signal processing: Illumination effects. In *Proc. ICPR 2006*, vol. III, pp. 738–741, 2006. IEEE-Press.
- [11] P. Meer, R. Lenz, and S. Ramakrishna. Efficient invariant representations. *Int. J. Comp. Vis.*, 26(2):137–152, 1998.
- [12] S. Nascimento, F. Ferreira, and D. Foster. Statistics of spatial cone-excitation ratios in natural scenes. *J. Opt. Soc. Am. A*, 19:1484–1490, 2002.
- [13] J. Parkkinen, J. Hallikainen, and T. Jaaskelainen. Characteristic spectra of Munsell colors. *J. Opt. Soc. Am. A.*, 6(2):318–322, 1989.
- [14] E. Schrödinger. Grundlinien einer Theorie der Farbenmetrik im Tagessehen. *Ann. Phys.*, 63:33–132, 1920.
- [15] C. Siegel. *Topics in Complex Function Theory, Automorphic Functions and Abelian Integrals*. Wiley Interscience, New York, 1971.
- [16] M. Solli, M. Andersson, R. Lenz, and B. Kruse. Color measurements with a consumer digital camera using spectral estimation techniques. *Proc. SCIA, LNCS-3540*, pp. 105–114. Springer, 2005.
- [17] G. Wyszecki and W. Stiles. *Color Science*. Wiley & Sons, London, England, 2 edition, 1982.