Color and Scale: The Spatial Structure of Color Images

Jan-Mark Geusebroek*, Rein van den Boomgaard, Arnold W. M. Smeulders, and Anuj Dev**

Intelligent Sensory Information Systems, Department of Computer Science, University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands; mark@wins.uva.nl

accepted by ECCV2000

Abstract. For grey-value images, it is well accepted that the neighborhood rather than the pixel carries the geometrical interpretation. Interestingly the spatial configuration of the neighborhood is the basis for the perception of humans. Common practise in color image processing, is to use the color information without considering the spatial structure. We aim at a physical basis for the local interpretation of color images. We propose a framework for spatial color measurement, based on the Gaussian scale-space theory. We consider a Gaussian color model, which inherently uses the spatial and color information in an integrated model. The framework is well-founded in physics as well as in measurement science. The framework delivers sound and robust spatial color invariant features. The usefulness of the proposed measurement framework is illustrated by edge detection, where edges are discriminated as shadow, highlight, or object boundary. Other applications of the framework include color invariant image retrieval and color constant edge detection.

1 Introduction

There has been a recent revival in the analysis of color in computer vision. This is mainly due to the common knowledge that more visual information leads to easier interpretation of the visual scene. A color image is easier to segment than a grey-valued image since some edges are only visible in the color domain and will not be detected in the grey-valued image. An area of large interest is searching for particular objects in images and image-databases, for which color is a feature with high reach in its data-values and hence high potential for discriminability. Color can thus be seen as an additional cue in image interpretation. Moreover, color can be used to extract object reflectance robust for a change in imaging conditions [4, 6, 17, 18]. Therefore color features are well suited for the description of an object.

^{*} This work is partly sponsored by Janssen Research Foundation, Beerse, Belgium,

^{**} and by the Dutch ministry of economic affairs (IOP BeeldVerwerking)

Colors are only defined in terms of human observation. Modern analysis of color has started in colorimetry where the spectral content of tri-chromatic stimuli are matched by a human, resulting in the well-known XYZ color matching functions [20]. However, from the pioneering work of Land [16] we know that a perceived color does not directly correspond to the spectral content of the stimulus; there is no one-to-one mapping of spectral content to perceived color. For example, a colorimetry purist will not consider brown to be a color, but as computer vision practisers would like to be able to define brown in an image when searching on colors. Hence, it is not only the spectral energy distribution coding color information, but also the spatial configuration of colors. We aim at a physical basis for the local interpretation of color images.

Common image processing sense tells us that the grey-value of a particular pixel is not a meaningful entity. The value 42 by itself tells us little about the meaning of the pixel in its environment. It is the local spatial structure of an image that has a close geometrical interpretation [13]. Yet representing the spatial structure of a color image is an unsolved problem.

The theory of scale-space [13, 19] adheres to the fact that observation and scale are intervened; a measurement is performed at a certain resolution. Differentiation is one of the fundamental operations in image processing, and one which is nicely defined [3] in the context of scale-space. In this paper we discuss how to represent color as a scalar field embedded in a scale-space paradigm. As a consequence, the differential geometry framework is extended to the domain of color images. We demonstrate color invariant edge detectors which are robust to shadow and highlight boundaries.

The paper is organized as follows. Section 2 considers the embedding of color in the scale-space paradigm. In Sect. 3 we derive estimators for the parameters in the scale-space model, and give optimal values for these parameters. The resulting sensitivity curves are colorimetrical compared with human color vision. Section 4 demonstrates the usefulness of the presented framework in physics based vision.

2 Color and Observation Scale

A spatio-spectral energy distribution is only measurable at a certain spatial resolution and a certain spectral bandwidth. Hence, physical realizable measurements inherently imply integration over spectral and spatial dimensions. The integration reduces the infinitely dimensional Hilbert space of spectra at infinitesimally small spatial neighborhood to a limited amount of measurements. As suggested by Koenderink [14], general aperture functions, or Gaussians and its derivatives, may be used to probe the spatio-spectral energy distribution. We emphasize that no essentially new color model is proposed here, but rather a theory of color measurement. The specific choice of color representation is irrelevant for our purpose. For convenience we first concentrate on the spectral dimension, later on we show the extension to the spatial domain.

2.1 The Spectral Structure of Color

From scale space theory we know how to probe a function at a certain scale; the probe should have a Gaussian shape in order to prevent the creation of extra details into the function when observed at a higher scale (lower resolution) [13]. As suggested by Koenderink [14], we can probe the spectrum with a Gaussian. In this section, we consider the Gaussian as a general probe for the measurement of spatio-spectral differential quotients.

Formally, let $E(\lambda)$ be the energy distribution of the incident light, where λ denotes wavelength, and let $G(\lambda_0; \sigma_\lambda)$ be the Gaussian at spectral scale σ_λ positioned at λ_0 . The spectral energy distribution may be approximated by a Taylor expansion at λ_0 ,

$$E(\lambda) = E^{\lambda_0} + \lambda E^{\lambda_0}_{\lambda} + \frac{1}{2}\lambda^2 E^{\lambda_0}_{\lambda\lambda} + \dots \quad (1)$$

Measurement of the spectral energy distribution with a Gaussian aperture yields a weighted integration over the spectrum. The observed energy in the Gaussian color model, at infinitely small spatial resolution, approaches in second order to

$$\hat{E}^{\sigma_{\lambda}}(\lambda) = \hat{E}^{\lambda_{0},\sigma_{\lambda}} + \lambda \hat{E}^{\lambda_{0},\sigma_{\lambda}}_{\lambda} + \frac{1}{2}\lambda^{2}\hat{E}^{\lambda_{0},\sigma_{\lambda}}_{\lambda\lambda} + \dots$$
(2)

where

$$\hat{E}^{\lambda_0,\sigma_\lambda} = \int E(\lambda)G(\lambda;\lambda_0,\sigma_\lambda)d\lambda \tag{3}$$

measures the spectral intensity,

$$\hat{E}_{\lambda}^{\lambda_0,\sigma_{\lambda}} = \int E(\lambda)G_{\lambda}(\lambda;\lambda_0,\sigma_{\lambda})d\lambda \tag{4}$$

measures the first order spectral derivative, and

$$\hat{E}^{\lambda_0,\sigma_\lambda}_{\lambda\lambda} = \int E(\lambda) G_{\lambda\lambda}(\lambda;\lambda_0,\sigma_\lambda) d\lambda \tag{5}$$

measures the second order spectral derivative. Further, G_{λ} and $G_{\lambda\lambda}$ denote derivatives of the Gaussian with respect to λ . Note that, throughout the paper, we assume scale normalized Gaussian derivatives to probe the spectral energy distribution.

Definition 1 (Gaussian Color Model). The Gaussian color model measures the coefficients $\hat{E}^{\lambda_0,\sigma_\lambda}$, $\hat{E}^{\lambda_0,\sigma_\lambda}_{\lambda}$, $\hat{E}^{\lambda_0,\sigma_\lambda}_{\lambda\lambda}$, ... of the Taylor expansion of the Gaussian weighted spectral energy distribution at λ_0 and scale σ_{λ} .

One might be tempted to consider a higher, larger than two, order structure of the smoothed spectrum. However, the subspace spanned by the human visual system is of dimension 3, and hence higher order spectral structure cannot be observed by the human visual system.

2.2 The Spatial Structure of Color

Introduction of spatial extent in the Gaussian color model yields a local Taylor expansion at wavelength λ_0 and position \mathbf{x}_0 . Each measurement of a spatio-spectral energy distribution has a spatial as well as spectral resolution. The measurement is obtained by probing an energy density volume in a three-dimensional spatio-spectral space, where the size of the probe is determined by the observation scale σ_{λ} and σ_{x} , see Fig. 1. It is directly clear that we do not separately consider spatial scale and spectral scale, but actually probe an energy density volume in the 3D spectral-spatial space where the "size" of the volume is specified by the observation scales.



Fig. 1. The probes for spatial color consists of probing the product of the spatial and the spectral space with a Gaussian aperture.

We can describe the observed spatial-spectral energy density $\hat{E}(\lambda, \boldsymbol{x})$ of light as a Taylor series for which the coefficients are given by the energy convolved with Gaussian derivatives:

$$\hat{E}(\lambda, \boldsymbol{x}) = \hat{E} + \begin{pmatrix} \boldsymbol{x} \\ \lambda \end{pmatrix}^T \begin{bmatrix} \hat{E}_{\boldsymbol{x}} \\ \hat{E}_{\lambda} \end{bmatrix} + \frac{1}{2} \begin{pmatrix} \boldsymbol{x} \\ \lambda \end{pmatrix}^T \begin{bmatrix} \hat{E}_{\boldsymbol{x}\boldsymbol{x}} & \hat{E}_{\boldsymbol{x}\lambda} \\ \hat{E}_{\lambda\boldsymbol{x}} & \hat{E}_{\lambda\lambda} \end{bmatrix} \begin{pmatrix} \boldsymbol{x} \\ \lambda \end{pmatrix} + \dots$$
(6)

where

$$\hat{E}_{\boldsymbol{x}^{i}\lambda^{j}}(\lambda,\boldsymbol{x}) = E(\lambda,\boldsymbol{x}) * G_{\boldsymbol{x}^{i}\lambda^{j}}(\lambda,\boldsymbol{x};\sigma_{\lambda},\sigma_{x}) \quad .$$
(7)

Here, $G_{\boldsymbol{x}^i\lambda^j}(\lambda, \boldsymbol{x}; \sigma_\lambda, \sigma_x)$ are the spatio-spectral probes, or color receptive fields. The coefficients of the Taylor expansion of $\hat{E}(\lambda, \boldsymbol{x})$ represent the local image structure completely. Truncation of the Taylor expansion results in an approximate representation, optimal in least squares sense.

For human vision, it is known that the Taylor expansion is spectrally truncated at second order [10]. Hence, higher order derivatives do not affect color as observed by the human visual system. Therefore, three receptive field families should be considered; the luminance receptive fields as known from luminance scale-space [15] extended with a yellow-blue receptive field family measuring the first order spectral derivative, and a red-green receptive field families are illustrated in Fig. 2. For human vision, the Taylor expansion for luminance is spatially truncated at fourth order [21].



Fig. 2. A diagrammatic representation of the various color receptive fields, here truncated at second order. The spatial luminance only yields the well-known [15] receptive fields from grey-value scale-space theory (column denoted by 0). For color vision, the luminance family is extended by a yellow-blue family (column 1) measuring the first-order spectral derivatives, and a red-green family (column 2) measuring the second-order spectral derivatives.

3 Colorimetric Analysis of the Gaussian Color Model

The eye projects the infinitely dimensional spectral density function onto a 3D 'color' space. Not any 3D subspace of the Hilbert space of spectra equals the subspace that nature has chosen. Any subspace we create with an artificial color model should be reasonably close in some metrical sense to the spectral subspace spanned by the human visual system.

Formally, the infinitely dimensional spectrum e is projected onto a 3D space c by $c = A^T e$, where $A^T = (XYZ)$ represents the color matching matrix. The subspace in which c resides, is defined by the color matching functions A^T . The range $\Re(A^T)$ defines what spectral distributions e can be reached from c, and the nullspace $\aleph(A^T)$ defines which spectra e cannot be observed in c. Since any spectrum $e = e_{\Re} + e_{\aleph}$ decomposed into a part that resides in $\Re(A^T)$ and a part that resides in $\aleph(A^T)$, we define

Definition 2. The observable part of the spectrum equals $e_{\Re} = \Pi_{\Re} e$ where Π_{\Re} is the projection onto the range of the human color matching functions A^T .

Definition 3. The non-observable (or metameric black) part of the spectrum equals $e_{\aleph} = \Pi_{\aleph} e$ where Π_{\aleph} is the projection onto the nullspace of the human color matching functions A^T .

The projection on the range $\Re(A^T)$ is given by [1]

$$\Pi_{\Re} : A^T \mapsto \Re \left(A^T \right) = A \left(A^T A \right)^{-1} A^T \tag{8}$$

and the projection on the nullspace

$$\Pi_{\aleph} : A^T \mapsto \aleph \left(A^T \right) = I - A \left(A^T A \right)^{-1} A^T = \Pi_{\Re}^{\perp} \quad . \tag{9}$$

Any spectral probe B^T that has the same range as A^T is said to be colorimetric with A^T and hence differs only in an affine transformation. An important property of the range projector Π_{\Re} is that it uniquely specifies the subspace. Thus, we can rephrase the previous statement into:

Proposition 4. The human color space is uniquely defined by $\Re(A^T)$. Any color model B^T is colorimetric with A^T if and only if $\Re(A^T) = \Re(B^T)$.

In this way we can tell if a certain color model is colorimetric with the human visual system. Naturally this is a formal definition. It is not well suited for a measurement approach where the color subspaces are measured with a given precision. A definition of the difference between subspaces is given by [9, Section 2.6.3],

Proposition 5. The largest principle angle θ between color subspaces given by their color matching functions A^T and B^T equals

$$\theta(A^T, B^T) = \arcsin\left(\left\|\Re\left(A^T\right) - \Re\left(B^T\right)\right\|_2\right)$$

Up to this point we did establish expressions describing similarity between different subspaces. We are now in a position to compare the subspace of the Gaussian color model with the human visual system by using the XYZ color matching functions. Hence, parameters for the Gaussian color model may be optimized to capture a similar spectral subspace as spanned by human vision, see Fig. 3. Let the Gaussian color matching functions be given by $G(\lambda_0, \sigma_\lambda)$. We have 2 degrees of freedom in positioning the subspace of the Gaussian color model; the mean λ_0 and scale σ_λ of the Gaussian. We wish to find the optimal subspace that minimizes the largest principle angle between the subspaces, i.e.:

$$B(\lambda_0, \sigma_\lambda) = (G(\lambda; \lambda_0, \sigma_\lambda)G_\lambda(\lambda; \lambda_0, \sigma_\lambda)G_{\lambda\lambda}(\lambda; \lambda_0, \sigma_\lambda))$$

$$\sin \theta = \operatorname*{argmin}_{\lambda_0, \sigma_\lambda} \left(\left\| \Re \left(A^T \right) - \Re \left(B(\lambda_0, \sigma_\lambda)^T \right) \right\|_2 \right)$$

An approximate solution is obtained for $\lambda_0 = 520 \text{ nm}$ and $\sigma_{\lambda} = 55 \text{ nm}$. The corresponding angles between the principal axes of the Gaussian sensitivities and

the 1931 and 1964 CIE standard observers are given in Tab. 1. Figure 4 shows the different sensitivities, together with the optimal (least square) transform from the XYZ sensitivities to the Gaussian basis, given by

$$\begin{bmatrix} \hat{E} \\ \hat{E}_{\lambda} \\ \hat{E}_{\lambda\lambda} \end{bmatrix} = \begin{pmatrix} -0.019 & 0.048 & 0.011 \\ 0.019 & 0 & -0.016 \\ 0.047 & -0.052 & 0 \end{pmatrix} \begin{bmatrix} \hat{X} \\ \hat{Y} \\ \hat{Z} \end{bmatrix} .$$
(10)

Since the transformed sensitivities are a linear (affine) transformation of the original XYZ sensitivities, the transformation is colorimetric with human vision. The transform is close to the Hering basis for color vision [10], for which the yellow-blue pathway indeed is found in the visual system of primates [2].



Fig. 3. Cohen's fundamental matrix \Re for the CIE 1964 standard observer, and for the Gaussian color model ($\lambda_0 = 520 \text{ nm}, \sigma_{\lambda} = 55 \text{ nm}$), respectively.

	Gauss – XYZ 1931	Gauss – XYZ 1964	XYZ 1931 – 1964
θ_1	26°	23.5°	9.8°
θ_2	21.5°	17.5°	3.9°
θ_3	3°	3°	1°

Table 1. Angles between the principal axes for various color systems. For determining the optimal values $\lambda_0, \sigma_\lambda$, the largest angle θ_1 is minimized. The distance between the Gaussian sensitivities for the optimal values $\lambda_0 = 520 \text{ nm}, \sigma_\lambda = 55 \text{ nm}$ and the different CIE colorimetric systems is comparable. Note the difference between the CIE systems is 9.8°.



Fig. 4. The Gaussian sensitivities at $\lambda_0 = 520$ nm and $\sigma_{\lambda} = 55$ nm (left). The The best linear transformation from the CIE 1964 XYZ sensitivities (middle) to the Gaussian bases is shown right. Note the correspondence between the transformed sensitivities and the Gaussian color model.

4 Results

4.1 The Gaussian Color Model by a RGB-Camera

A RGB-camera approximates the CIE 1931 XYZ basis for colorimetry by the linear transform [11]

$$\begin{bmatrix} \hat{X} \\ \hat{Y} \\ \hat{Z} \end{bmatrix} = \begin{pmatrix} 0.621 & 0.113 & 0.194 \\ 0.297 & 0.563 & 0.049 \\ -0.009 & 0.027 & 1.105 \end{pmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} .$$
(11)

The best linear transform from XYZ values to the Gaussian color model is given by (Eq. 10). A better approximation to the Gaussian color model may be obtained for known camera sensitivities. Figure 5 shows an example image and its Gaussian color model components.



Fig. 5. The example image (left) and its color components \hat{E} , \hat{E}_{λ} , and $\hat{E}_{\lambda\lambda}$, respectively. Note that for the color component \hat{E}_{λ} achromaticity is shown in grey, negative bluish values are shown in dark, and positive yellowish in light. Further, for $\hat{E}_{\lambda\lambda}$ achromaticity is shown in grey, negative greenish in dark, and positive reddish in light.

4.2 Color Invariant Edge Detection

An interesting problem in the segmentation of man made objects is the segmentation of edges into the "real" object edges, or "artificial" edges caused by shadow boundaries or highlights [8]. Consider an image captured under white illumination. A common model for the reflection of light by an object to the camera is given by the Kubelka-Munk theory [12, 18, 20]

$$E(\lambda, \boldsymbol{x}) = i(\boldsymbol{x}) \left\{ \rho_{\rm f}(\boldsymbol{x}) + (1 - \rho_{\rm f}(\boldsymbol{x}))^2 R_{\infty}(\lambda, \boldsymbol{x}) \right\}$$
(12)

where *i* denotes the intensity distribution, $\rho_{\rm f}$ the Fresnel reflectance at the object surface, and R_{∞} the object spectral reflectance function. The reflected spectral energy distribution in the camera direction is denoted by *E*. The quantities *i* and $\rho_{\rm f}$ depend on both scene geometry and object properties, where R_{∞} depends on object properties only. Edges may occur under three circumstances:

- shadow boundaries due to edges in $i(\mathbf{x})$
- highlight boundaries due to edges in $\rho_{\rm f}(\boldsymbol{x})$
- material boundaries due to edges in $R_{\infty}(\lambda, \boldsymbol{x})$.

For the model given by (Eq. 12), material edges are detected by considering the ratio between the first and second order derivative with respect to λ , or

$$\frac{\partial}{\partial x} \left\{ \frac{E_{\lambda}}{E_{\lambda\lambda}} \right\}$$

where E represents $E(\lambda, \boldsymbol{x})$ and indices denote differentiation. Further, the ratio between $E(\lambda, \boldsymbol{x})$ and its spectral derivative are independent of the spatial intensity distribution. Hence, the spatial derivative

$$\frac{\partial}{\partial x} \left\{ \frac{E_{\lambda}}{E} \right\}$$

and

$$\frac{\partial}{\partial x} \left\{ \frac{E_{\lambda\lambda}}{E} \right\}$$

depend on Fresnel and material edges. Finally, the spatial derivatives of E, E_{λ} , and $E_{\lambda\lambda}$ depend on intensity, Fresnel, and material edges. Measurement of these expressions is obtained by substitution of E, E_{λ} , and $E_{\lambda\lambda}$ for the measured values \hat{E} , \hat{E}_{λ} , and $\hat{E}_{\lambda\lambda}$ (Eq. 10) at scale σ_x , together with their spatial derivatives. Combining these expressions in gradient magnitudes yields Fig. 6.

5 Conclusion

In this paper, we have established the measurement of spatial color information from RGB-images, based on the Gaussian scale-space paradigm. We have shown



Fig. 6. Edge detection in a color image. The left figure shows edges due to object reflectance; the second figure includes highlight boundaries, whereas the figure on the right also exhibits shadow boundaries. Spatial scale $\sigma_x = 1$ pixel. The image is captured by a Sony XC-003P camera, white balanced under office lightning, gamma turned off.

that the formation of color images yield a spatio-spectral integration process at a certain spatial and spectral resolution. Hence, measurement of color images implies probing a three-dimensional energy density at a spatial scale σ_x and spectral scale σ_{λ} . The Gaussian aperture may be used to probe the spatiospectral energy distribution.

We have achieved a spatial color model, well founded in physics as well as in measurement science. The parameters of the Gaussian color model have been estimated such that a similar spectral subspace as human vision is captured. The Gaussian color model solves a fundamental problem of color and scale by integrating the spatial and color information. The model measures the coefficients of the Taylor expansion of the spatio-spectral energy distribution. Hence, the Gaussian color model describes the local structure of color images. As a consequence, the differential geometry framework is extended to the domain of color images.

Spatial differentiation of expressions derived from the Gaussian color model is inherently well-posed, in contrast with often ad-hoc methods for detection of hue edges and other color edge detectors. The framework is successfully applied to color edge classification, labeling edges as material, shadow, or highlight boundaries. Other application areas include physics-based vision [6], image database searches [7], color constant edge detection [5], and object tracking.

References

- J. B. Cohen and W. E. Kappauff. Color mixture and fundamental metamer: Theory, algebra, geometry, application. Am. J. Psych., 98:171–259, 1985.
- D. M. Dacey and B. B. Lee. The "blue-on" opponent pathway in primate retina originates from a distinct bistratified ganglion cell type. *Nature*, 367:731–735, 1994.
- L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Cartesian differential invariants in scale-space. *Journal of Mathematical Imaging* and Vision, 3(4):327–348, 1993.

- R. Gershon, D. Jepson, and J. K. Tsotsos. Ambient illumination and the determination of material changes. J. Opt. Soc. Am. A, 3:1700–1707, 1986.
- J. M. Geusebroek, A. Dev, R. van den Boomgaard, A. W. M. Smeulders, F. Cornelissen, and H. Geerts. Color invariant edge detection. In *Scale-Space Theories* in Computer Vision, pages 459–464. Springer-Verlag, 1999.
- T. Gevers and A. W. M. Smeulders. Color based object recognition. *Pat. Rec.*, 32:453–464, 1999.
- T. Gevers and A. W. M. Smeulders. Content-based image retrieval by viewpointinvariant image indexing. *Image Vision Comput.*, 17(7):475–488, 1999.
- T. Gevers and H. Stokman. Reflectance based edge classification. In *Proceedings of Vision Interface*, pages 25–32. Canadian Image Processing and Pattern Recognition Society, 1999.
- G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins Press Ltd., London, 1996.
- E. Hering. Outlines of a Theory of the Light Sense. Harvard University Press, Cambridge, MS, 1964.
- ITU-R Recommendation BT.709. Basic parameter values for the HDTV standard for the studio and for international programme exchange. Technical Report BT.709 [formerly CCIR Rec. 709], ITU, 1211 Geneva 20, Switzerland, 1990.
- D. B. Judd and G. Wyszecki. Color in Business, Science, and Industry. Wiley, New York, NY, 1975.
- 13. J. J. Koenderink. The structure of images. Biol. Cybern., 50:363-370, 1984.
- J. J. Koenderink and A. Kappers. *Color Space*. Utrecht University, The Netherlands, 1998.
- J. J. Koenderink and A. J. van Doorn. Receptive field families. *Biol. Cybern.*, 63:291–297, 1990.
- 16. E. H. Land. The retinex theory of color vision. Sci. Am., 237:108-128, 1977.
- K. D. Mielenz, K. L. Eckerle, R. P. Madden, and J. Reader. New reference spectrophotometer. *Appl. Optics*, 12(7):1630–1641, 1973.
- S. A. Shafer. Using color to separate reflection components. Color Res. Appl., 10(4):210–218, 1985.
- B. M. ter Haar Romeny, editor. Geometry-Driven Diffusion in Computer Vision. Kluwer Academic Publishers, Boston, 1994.
- G. Wyszecki and W. S. Stiles. Color Science: Concepts and Methods, Quantitative Data and Formulae. Wiley, New York, NY, 1982.
- R. A. Young. The gaussian derivative theory of spatial vision: Analysis of cortical cell receptive field line-weighting profiles. Technical Report GMR-4920, General Motors Research Center, Warren, MI, 1985.