Color Differential Structure

Bart M. ter Haar Romeny[1], Jan-Mark Geusebroek[2], Peter Van Osta[3], Rein van den Boomgaard[2], and Jan J. Koenderink[4]

[1] Image Sciences Institute, Utrecht University, The Netherlands
[2] Intelligent Sensory Information Systems, University of Amsterdam, The Netherlands
[3] Biological Imaging Laboratory, Janssen Research Foundation, Beerse, Belgium
[4] Dept. Physics of Man, Utrecht University, The Netherlands

Abstract. The scale-space approach to the differential structure of color images was recently introduced by Geusebroek et al. [1, 2], based on the pioneering work of Koenderink's Gaussian derivative color model [5]. To master this theory faster, we present the theory as a practical implementation in the computer algebra package Mathematica for the extraction of color differential structure. Many examples are given, the practical code examples enable easy extensive experimentation. High level programming is now fast: all examples run with 5-15 seconds on typical images on a typical modern PC.

1 Color Image Formation and Color Invariants

Color is an important extra dimension. Information extracted from color is useful for almost any computer vision task, like segmentation, surface characterization, etc. The field of color science is huge [6], and many theories exist. It is far beyond the scope of this paper to cover even a fraction of the many different approaches. We will focus on a single recent theory, based on scale-space models for the color sensitive receptive fields in the front-end visual system. We are especially interested in the extraction of multi-scale differential structure (derivatives) in the spatial and the wavelength domain of color images. What is color invariant structure? To understand that notion, we first have to study the process of color image formation.

The light spectrum falling onto the eye results from interaction between a light source, the object, and the observer. Color may be regarded as the measurement of spectral energy, and will be handled in the next section. Here, we only consider the interaction between light source and material. Before we see an object as having a particular color, the object needs to be illuminated. After all, in darkness objects are simply black. The emission spectra $l(\lambda)$ of common light sources are close to Planck's formula [6] (NB: λ in nm):

$$\mathtt{h} = 6.626176\,10^{-34};\ \mathtt{c} = 2.99792458\,10^{8};\ \mathtt{k} = 2.9979245810^{8};$$
$$\mathtt{l}[\lambda\_, \mathtt{T}\_] = 8\pi\mathtt{hc}(10^{-9}\lambda)^{-5}$$

where h is Planck's constant, k Boltzmann's constant, and c the velocity of light in vacuum. The *color temperature* of the emitted light is given by T, and typically ranges from $2500K$ (warm red light) to $10,000K$ (cold blue light). Note that the terms "warm" and "cold" are given by artists, and refer to the sensation caused by the light. Representative white light is, by convention, chosen to be at a temperature of $6500K$. However, in practice, all light sources between $10,000K$ and $2500K$ can be found. Planck's equation is adequate for incandescent light and halogen. The spectrum of daylight is slightly different, and is represented by a *correlated color temperature*. Daylight is close enough to the Planckian spectrum to be characterized by a equivalent parameter.

The part of the spectrum reflected by a surface depends on the surface *spectral reflection function*. The spectral reflectance is a material property, characterized by a function $c(\lambda)$. For planar, matte surfaces, the spectrum reflected by the material $e(\lambda)$ is simplified as the multiplication between the spectrum falling onto the surface $l(\lambda)$ and the surface spectral reflectance function $c(\lambda)$: $e(\lambda) = c(\lambda)l(\lambda)$.

At this point it is meaningful to introduce spatial extent, hence to describe the spatio-spectral energy distribution $e(x, y, \lambda)$ that falls onto the retina. Further, for three-dimensional objects the amount of light falling onto the object's surface depends on the energy flux, thus on the local geometry. Hence *shading* (and shadow) may be introduced as being a wavelength independent multiplication factor $m(x, y)$ in the range $[0, 1]$: $e(x, y, \lambda) = c(x, y, \lambda)l(\lambda)m(x, y)$. Note that the illumination $l(\lambda)$ is independent of position. Hence the equation describes spectral image formation of matte objects, illuminated by a single light source. For shiny surfaces the image formation equation has to be extended with an additive term describing the Fresnel reflected light, see [3] for more details.

The *structure* of the spatio-spectral energy distribution is due to the three functions c, l, and m. By making some general assumptions, these quantities may be derived from the measured image. Estimation of the object reflectance function c boils down to deriving material properties, the "true" color invariant which does not depend on illumination conditions. Estimation of the light source l is well known as the color constancy problem. Determining m is in fact estimating the shadows and shading in the image, and is closely related to the shape-from-shading problem.

For the extraction of color invariant properties from the spatio-spectral energy distribution we search for algebraic or differential expressions of e, which are independent of l and m. Hence the goal is to solve for differential expressions of e which results in a function of c only.

To proceed, note that the geometrical term m is only a function of spatial position. Differentiation with respect to λ, and normalization reduces the problem to only two functions: $e(\mathbf{x}, \lambda) = c(\mathbf{x}, \lambda)l(\lambda)m(\mathbf{x}) \Rightarrow \frac{1}{e(\mathbf{x},\lambda)}\frac{\partial e(\mathbf{x},\lambda)}{\partial \lambda} = \frac{l_\lambda}{l} + \frac{c_\lambda}{c}$ (indices indicate differentiation). After additional differentiation to the spatial

variable x or y, the first term vanishes, since l only depends on λ

$$\texttt{e[x\_,}\lambda\texttt{\_]} = \texttt{c[x, }\lambda\texttt{] l[}\lambda\texttt{] m[x]};$$
$$\partial_{\texttt{x}}\left(\frac{\partial_\lambda \texttt{e[x,}\lambda\texttt{]}}{\texttt{e[x,}\lambda\texttt{]}}\right)\texttt{//shortnotation}$$
$$\frac{\texttt{cc}_{\texttt{xy}} - \texttt{c}_{\texttt{x}}\texttt{c}_{\texttt{y}}}{\texttt{c}^2}$$

The spatial derivative of the normalized wavelength derivative, after applying the chain rule,

$$\texttt{D}\left[\frac{\texttt{D[e[x,y,}\lambda\texttt{], }\lambda\texttt{]}}{\texttt{e[x,y,}\lambda\texttt{]}}, \texttt{x}\right]\texttt{//shortnotation}$$
$$\frac{\texttt{e e}_{\texttt{x}\lambda} - \texttt{e}_{\texttt{x}}\,\texttt{e}_\lambda}{\texttt{e}^2}$$

is completely expressed in spatial and spectral derivatives of the observable spatio-spectral energy distribution.

We develop the differential properties of the invariant color-edge detector $\mathcal{E} = \frac{1}{e}\frac{\partial e}{\partial \lambda}$, where the measured spectral intensity $e = e(x, y, \lambda)$. Spatial derivatives of $\mathcal{E}$, like $\frac{\partial \mathcal{E}}{\partial c}$, contain derivatives to the spatial as well as to the wavelength dimension due to the chain rule. In the next section we will see that the zero-th, first and second order derivative-to-λ kernels are acquired from the transformed RGB space of the image directly. The derivatives to the spatial coordinates are acquired in the conventional way, i.e. convolution by a spatial Gaussian kernel.

2 Koenderink's Gaussian Derivative Color Observation Model

Spatial structure can be extracted from the data in the environment by measuring the N-jet of (scaled) derivatives to some order. For the spatial domain this has led to the family of Gaussian derivative kernels, sampling the spatial intensity distribution. These derivatives naturally occur in a local Taylor expansion of the signal.

Koenderink [5] proposed to take a similar approach to the sampling of the color dimension, i.e. the spectral information contained in the color. If we construct the Taylor expansion of the spatio-spectral energy distribution $e(x, y, \lambda)$ of the measured light to wavelength, in the fixed spatial point (x_0, y_0), and around a central wavelength λ_0 we get (to second order):

$$\texttt{Series[e[x0, y0, }\lambda\texttt{], \{}\lambda\texttt{, }\lambda\texttt{0, 2\}]}$$
$$\texttt{e[x0, y0, }\lambda\texttt{0]} + \texttt{e}^{(0,0,1)}\texttt{[x0, y0,}\lambda\texttt{0]}(\lambda - \lambda 0) +$$
$$\frac{1}{2}\texttt{e}^{(0,0,2)}\texttt{[x0, y0, }\lambda\texttt{0]}(\lambda - \lambda 0)^2 + O[\lambda - \lambda 0]^3$$

A physical measurement with an aperture is mathematically described with a convolution. So for a measurement of the luminance a with aperture function $G(x, \sigma)$ in the (here in the example 1D) spatial domain we get: $L(x; \sigma) =$

$\int_{-\infty}^{\infty} L(x-y)G(y;\sigma)dy$ where y is the dummy spatial shift parameter running over all possible values. For the temporal domain we get $L(t;\sigma) = \int_{-\infty}^{\infty} L(t-s)G(s;\sigma)ds$ where s is the dummy temporal shift parameter running over all possible values in time. Based on this analogy, we might expect a measurement along the color dimension to look like: $L(\lambda;\sigma) = \int_{-\infty}^{\infty} L(\lambda-\mu)G(\mu;\sigma)d\mu$ where λ is the wavelength and μ is the dummy wavelength shift parameter.

In the scale-space model for vision the front-end visual system has implemented the shifted spatial kernels with a grid on the retina with receptive fields, so the shifting is implemented by the simultaneous measurement of all the neighboring receptive fields. The temporal kernels are implemented as time-varying lateral geniculate nucleus (LGN) and cortical receptive fields. However, in order to have a wide range of receptive fields which shift over the wavelength axis in sensitivity, would require a lot of different photo-sensitive dyes (rhodopsines) in the receptors with these different—shifted—color sensitivities. The visual system has opted for a cheaper solution: The convolution is calculated at just a single position on the wavelength axis, at around $\lambda_0 = 520$nm, with a standard deviation of the Gaussian kernel of about $\sigma_\lambda = 55$ nm. The integration is done over the range of wavelengths that is covered by the rhodopsines, i.e. from about 350 nm (blue) to 700 nm (red). The values for λ_0 and σ_0 are determined from the best fit of a Gaussian to the spectral sensitivity as measured phychophysically in humans, i.e. the Heering model.

So we get for the spectral intensity

$$e(\mathbf{x},\lambda_0;\sigma_0) = \int_{\lambda_{\min}}^{\lambda\max} e(\mathbf{x},\lambda)G(\lambda,\lambda_0;\sigma_\lambda)d\lambda$$

. This is a 'static' convolution operation (i.e. inner product in function space). It is not a convolution in the familiar sense, because we don't shift over the whole wavelength axis. We just do a single measurement with a Gaussian aperture over the wavelength axis at the position a. Similarly, the derivatives to λ:

$$\frac{\partial e(\mathbf{x},\lambda_0)}{\partial\lambda} = \sigma_\lambda \int_{\lambda_{\min}}^{\lambda\max} e(\mathbf{x},\lambda)\frac{\partial^2 G(\lambda,\lambda_0,\sigma_\lambda)}{\partial\lambda^2}d\lambda$$

and

$$\frac{\partial^2 e(\mathbf{x},\lambda_0)}{\partial\lambda^2} = \sigma_\lambda^2 \int_{\lambda_{\min}}^{\lambda\max} e(\mathbf{x},\lambda)\frac{\partial^2 G(\lambda,\lambda_0,\sigma_\lambda)}{\partial\lambda^2}d\lambda$$

describe the first and second order spectral derivative respectively. The factors σ_λ and σ_λ^2 are included for the normalization, i.e. to make the Gaussian spectral kernels dimensionless.

In Fig. 1 the graphs of the 'static' normalized Gaussian spectral kernels to second order as a function of wavelength are given.

Color sensitive receptive fields come in the combinations red-green and yellow-blue center-surround receptive fields. The subtraction of yellow and blue in these receptive fields is well modeled by the first order derivative to λ, the subtraction

```
gaussλ[λ_,σ_] = D[gauss[λ,σ],λ];
gaussλλ[λ_,σ_] = D[gauss[λ,σ],{λ,2}];
```
$\lambda 0 = 520; \sigma 0 = 55;$

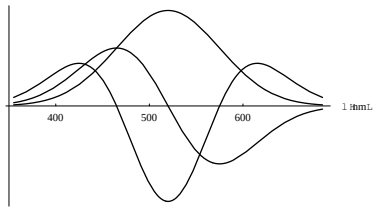

Fig. 1. The zero-th, first and second derivative of the Gaussian function with respect to wavelength as models for the color receptive field's wavelengths sensitivity in human color vision. After [Koenderink 1998a]. The central wavelength is 520 nm, the standard deviation 55 nm.

of red and green minus the blue is well modeled by the second order derivative to λ. Alternatively, one can say that the zero-th order receptive field measures the luminance, the first order the 'blue-yellowness', and the second order the 'red-greenness'.

Note: the wavelength axis is a half axis. It is known that for a *half axis* (such as with positive-only values) a logarithmic parameterization is the natural way to 'step along' the axis. E.g. the scale axis is logarithmically sampled in scale-space (remember the 'orders of magnitudes'), the intensity is logarithmically transformed in the photoreceptors, and the time axis can only be measured causally when we sample it logarithmically. We might conjecture here a better fit to the Heering model with a logarithmic wavelength axis.

The Gaussian color model needs the first three components of the Taylor expansion of the Gaussian weighted spectral energy distribution at λ_0 and scale σ_0. An RGB camera measures the red, green and blue component of the incoming light, but this is not what we need for the Gaussian color model. We need a method to extract the Taylor expansion terms from the RGB values.

An RGB camera approximates the CIE 1964 XYZ basis for colorimetry by the linear transformation matrix `rgb2xyz`, while Geusebroek et al. [2] give the best linear transform from the XYZ values to the Gaussian color model, i.e. matrix `xyz2e`:

$$\texttt{rgb2xyz} = \begin{pmatrix} 0.621 & 0.113 & 0.194 \\ 0.297 & 0.563 & 0.049 \\ -0.009 & 0.027 & 1.105 \end{pmatrix} \qquad \texttt{xyz2e} = \begin{pmatrix} -0.019 & 0.048 & 0.011 \\ 0.019 & 0.000 & -0.016 \\ 0.047 & -0.052 & 0.000 \end{pmatrix}$$

The resulting transform from the measured RGB input image to the sampling ' la human vision' is the product of the above matrices: `colorRF = xyz2e . rgb2xyz`.

The Gaussian color model is an approximation, but has the attractive property of fitting very well into Gaussian scale-space theory. The notion of image structure is extended to the wavelength domain in a very natural and coherent way. The similarity with human differential-color receptive fields is more than a coincidence.

Now we have all the tools to come to an actual implementation. The RGB values of the input image are transformed into Gaussian color model space, and plugged into the spatio-spectral formula for the color invariant feature. Next to the derivatives to wavelength we need spatial derivatives, which are computed in the regular way with spatial Gaussian derivative operators. The full machinery of e.g. gauge coordinates and invariance under specific groups of transformations is also applicable here. The next section details the implementation.

3 Implementation

In *Mathematica* a color image is represented as a two dimensional array of color triplets. The RGB triples are converted into measurements through the color receptive fields in the retina with the transformation matrix `colorRF` defined in the previous section. To transform every RGB triple we map the transformation to our input image as a pure function at the second list level: `observedimage = Map[Dot[colorRF,#]&, im, 2];`.

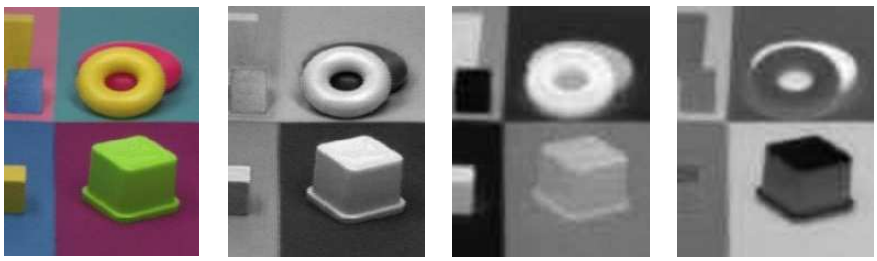

Fig. 2. The input image (left) and the observed images e, e_λ and $e_{\lambda\lambda}$ with the color differential receptive fields. Image resolution 228x179 pixels.

The color image data set can be smartly resliced by a reordering `Transpose`:

```
obs = Transpose[observedimage, 2,3,1];.
```

The resulting data set is a list of three scalar images, allowing us to access the measurements e, e_λ and $e_{\lambda\lambda}$ individually as scalar images (see Fig. 2).

We now develop the differential properties of our invariant color-edge detector $\mathcal{E} = \frac{1}{e}\frac{\partial e}{\partial \lambda}$, where the spectral intensity $e = e(x, y, \lambda)$. The derivatives to the spatial and spectral coordinates are easily found with the chainrule. Here are

the explicit forms:

$$\mathcal{E} := \frac{\texttt{D[e[x,y,}\lambda\texttt{], }\lambda\texttt{]}}{\texttt{e[x,y,}\lambda\texttt{]}};$$

$$\texttt{shortnotation[} \partial_x\mathcal{E}\texttt{, } \partial_y\mathcal{E}\texttt{, } \partial_\lambda\mathcal{E} \texttt{]}$$

$$\left\{\frac{e e_{x\lambda} - e_x e_\lambda}{e^2}, \frac{e e_{y\lambda} - e_y e_\lambda}{e^2}, \frac{-e_\lambda^2 + e e_{\lambda\lambda}}{e^2}\right\}$$

The gradient magnitude (detecting yellow-blue transitions) becomes:

$$\mathcal{G} = \texttt{Simplify[} \sqrt{(\partial_x\mathcal{E})^2 + (\partial_y\mathcal{E})^2} \texttt{]}$$

$$\mathcal{G} \texttt{ // shortnotation}$$

$$\sqrt{\frac{(e\, e_{x\lambda} - e_x\, e_\lambda)^2 + (e\, e_{y\lambda} - e_y\, e_\lambda)^2}{e^4}}$$

The second spectral order gradient (detecting purple-green transitions) becomes:

$$\mathcal{W} = \texttt{Simplify[} \sqrt{(\partial_{x,\lambda}\mathcal{E})^2 + (\partial_{y,\lambda}\mathcal{E})^2} \texttt{]}$$

Finally, the total edge strength $\mathcal{N}$ (for all color edges) in the spatio-spectral domain becomes:

$$\mathcal{N} = \texttt{Simplify[} \sqrt{(\partial_x\mathcal{E})^2 + (\partial_y\mathcal{E})^2 + (\partial_{x,\lambda}\mathcal{E})^2 + (\partial_{y,\lambda}\mathcal{E})^2} \texttt{];}$$

As an example, we implement this last expression for discrete images. First we replace each occurrence of a derivative to λ with the respective plane in the observed image `rf` (by the color receptive fields). Note that we use `rf[[n`λ`+1]]` because the zero-th list element is the `Head` of the list.

We will look for derivative patterns in the *Mathematica* expression for $\mathcal{N}$ and replace them with another pattern. We do this pattern matching with the command `/.` (`ReplaceAll`). We call the observed image at this stage `rf`, without any assignment to data, so we can do all calculations symbolically first:

```
Clear[rf0, rf1, rf2, σ]; rf = {rf0, rf1, rf2};
𝒩 = 𝒩 /. { Derivative[nx_, ny_, nλ_][e][x, y, λ]
  :→ Derivative[nx, ny][rf[[nλ+1]]][x,y], e[x,y,λ]]
    :→ rf[[1]] } // Simplify;
```

Note that we do a delayed rule assignment here ($:\!\!\rightarrow$ instead of $\rightarrow$) because we want to evaluate the right hand side only after the rule is applied. We finally replace the spatial derivatives with the spatial Gaussian derivative convolution `gD` at scale σ:

```
𝒩 = 𝒩 /. { Derivative[nx_,ny_][rf_][x, y]
  :→ gD[rf,nx,ny,σ], rf1[x, y] :→ rf1, rf2[x, y] :→ rf2}
```

The resulting expression for the total edge strength can now safely be calculated on the discrete data (see Fig. 3). Equivalent expressions can be formulated for the yellow-blue edges $\mathcal{G}$ and the red-green edges $\mathcal{W}$, the results of these detectors are given in Fig. 4.

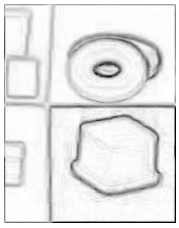

Fig. 3. The color-invariant $\mathcal{N}$ calculated for our input image at spatial scale $\sigma = 1$ pixel. Primarily the red-green color edges are found, as expected, with little edge detection at intensity edges. Image resolution 228x179 pixels.

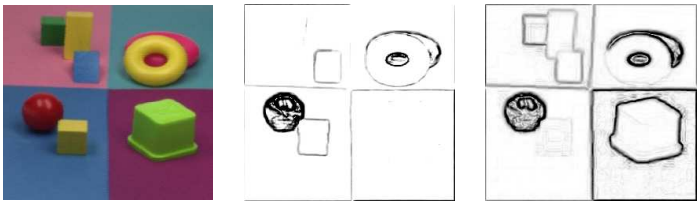

Fig. 4. Left: original color image. Middle: The yellow-blue edge detector $\mathcal{E}$ calculated at a spatial scale $\sigma = 1$ pixel. Note that there is hardly any red-green edge detection. Right: output of the red-green edge detector $\mathcal{W}$. Image resolution 249x269.

4 Combination with Spatial Constraints

Interesting combinations can be made when we combine the color differential operators with the spatial differential operators. E.g. when we want to detect specific blobs with a specific size and color, we can apply feature detectors that are best matching the shape to be found. We end the paper with one examples: locating PAS stained material in a histological preparation. This examples illustrates the possible use of color differential operators and spatial differential operators in microscopy.

Blobs are detected by calculating those locations (pixels) where the Gaussian curvature $\texttt{lgc} = L_{xx}L_{yy} - L_{xy}^2$ on the black-and-white version (`imbw`) of the image is greater then zero. This indicates a convex 'hilltop'. Pixels on the boundaries of the 'hilltop' are detected by requiring the second order directional derivative in the direction of the gradient $(L_x^2 L_{xx} + 2L_x L_{xy} + L_y^2 L_{yy})/(L_x^2 + L_y^2)$ to be positive. Interestingly, by using these invariant shape detectors we are largely independent of image intensity. For the color scheme we rely on $\mathcal{E}$ and its first and second order derivative to λ.

In Fig. 5 we detect stained carbohydrate deposits in a histological application using this combined color and spatial structure detection mechanism.

The Mathematica functions not described in the text and the images are available from the first author.

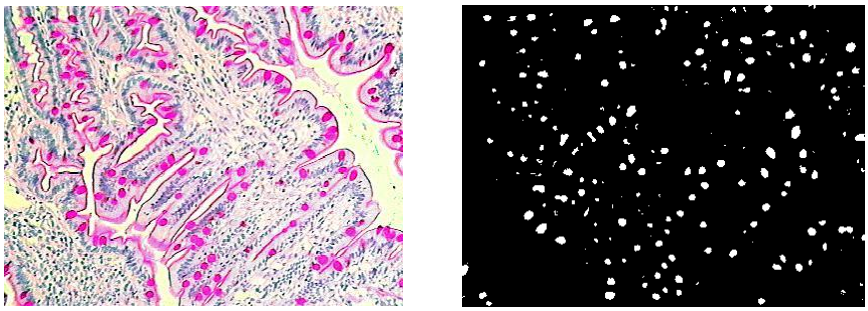

Fig. 5. Detection of carbohydrate stacking in the mucus in intestinal cells, that are specifically stained for carbohydrates with periodic acid Schiff (P.A.S.). The carbohydrate deposits are in magenta, cell nuclei in blue. The blob-like areas are detected with positive Gaussian curvature and positive second order directional derivative in the gradient direction of the image intensity, the magenta with a boolean combination of the color invariant $\mathcal{E}$ and its derivatives to λ. Scale = 4 pixels. Example due to P. Van Osta. Image taken from http://www.bris.ac.uk/Depts/PathAndMicro/CPL/pas.html.

References

1. Geusebroek, J.M., Dev, A., van den Boomgaard, R., Smeulders A.W.M., Cornelissen, F., and Geerts H., Color Invariant edge detection. In: Scale-Space theories in Computer Vision, Lecture Notes in Computer Science, vol. 1252, pp. 459—464, Springer-Verlag, 1999.
2. Geusebroek, J.M., van den Boomgaard, R., Smeulders A.W.M., and Dev, A., Color and scale: the spatial structure of color images. In: Eur. Conf. on Computer Vision 2000, Lecture Notes in Computer Science, Vol. 1842, Springer, pp. 331—341, June 26 - July 1, 2000.
3. Geusebroek, J.M., Smeulders A.W.M., and van den Boomgaard, R., Measurement of Color Invariants. Proc. CVPR, vol. 1, pp. 50—57, June 13-15, 2000.
4. Hering, E., Outlines of a theory of the light sense, Harvard University Press, Cambridge, 1964.
5. Koenderink, J.J., Color Space. Utrecht University, the Netherlands, 1998.
6. Wyszecki, G., and Stiles, W.S., Color science: concepts and methods, quantitative data and formulae, Wiley, New York, NY, 2nd edition, 2000.