Most Luminous Tetrachromatic Surfaces

Alfredo Restrepo (Palacios) and Edisson Maldonado

Laboratorio de señales, Dept. de Ing. Eléctrica y Electrónica, Universidad de los Andes, Bogotá, Colombia http://labsenales.uniandes.edu.co, arestrep@uniandes.edu.co, eomaldonado10@uniandes.edo.co

Abstract

A surface of a given chromaticity is seen as more colourful, the more luminous it is; if the surface is black, its color is seen more vivid the less luminous it is. In the trichromatic case, the spectral reflectance of a most luminous surface is binary and may have at most 2 transitions or discontinuities, in the wavelength domain. Assuming a spectrally flat illuminant, we give analog conditions for the spectral reflectance of a surface to be most luminous, when observed through photoreceptors having four types of spectral function.

Introduction

It is well known that jacquets of a certain citrine (lemon yellow) hue, corresponding to a spectral colour of a wavelength of about 550 nm, are seen as having a very vivid colour, and the explanation must be different from that of the luminosity of spectral beams, as in the component \bar{y} [2] of the CIE tristimulus values \bar{x} , \bar{y} and \bar{z} .

The shape of the spectral reflectance of a surface giving rise to a *most luminous colour* of a given chromaticity was first advanced by Schrödinger [1], while MacAdam [2] studied the concept in the context of the CIE's chromaticity diagram. West and Brill [3] revised MacAdam's proof stressing the need for the convexity of the chromaticity diagram, and of a certain property that the ordering of spectral colours in the chromaticity diagram must have. The important work of MacAdam [2] and of Brill [3], which followed upon the pre-war work of Ostwald and Schrödinger, is given in the context of the CIE chromaticity diagram and the illuminant is a variable in their work.

We generalise Schrödinger's characterisation of most luminous colours (of non light-emitting surfaces) to the case of tetrachromacy, for which we work in a hypercube of colours; we find analogous results. This may have applications in computer vision, for 4-spectral imaging systems, for the detection of most luminous objects of a given tetrachromatic hue; also, in the study of animal vision, for the case of tetrachromatic animals, where luminous colours can be designed for their use in experiments for testing their vision.

We review the trichromatic case first and then consider the tetrachromatic case. In the trichromatic case, we do not work in the context of CIE's chromaticity diagram but with a chromaticity space derived in the trichromatic, *lms cube*: we assume a vision system with three types of photoreceptor spectral absorbances, $l(\lambda)$, $m(\lambda)$ and $s(\lambda)$. So, we do not speak of *tristimulus specifications*, rather, we speak of the spectral reflectance curve of a surface and of its corresponding *colour* in the lms cube space (we assume a spectrally flat illuminant). In the case of most luminous colours, we speak of *Schrödinger colours* (points in the *lms* cube) and of the corresponding *Schrödinger reflectances*

bourful, they are in a one-to-one correspondence. The map $\rho \mapsto [l, m, s]$ from reflectances to colours is given by *lse, the nd may* $\phi(\rho) = [L, M, S] = [\int \rho(\lambda) l(\lambda) d\lambda, \int \rho(\lambda) m(\lambda) d\lambda, \int \rho(\lambda) s(\lambda) d\lambda]$ *belength*

and we assume that $\int l(\lambda)d\lambda = \int m(\lambda)d\lambda = \int s(\lambda)d\lambda = 1$.

giving rise to them; for nonnegative spectral receptoral functions,

In contrast to the colour of a light (a beam colour [4]), the colour of a surface (an object colour [4]) reflecting light of a continuous and broad spectrum, for only one wavelength would be perceived as black, since very few photons would bounce back from the surface. Surfaces of a light colour reflect light at relatively large wavelength intervals. It was discovered by Schrödinger, and foreseen by Ostwald, that the spectral reflectance curve of theoretically most luminous surfaces, of a given hue, is binary (taking only the values 0 and 1) and has at most two transitions or discontinuities. Spectral lights (those having their energy concentrated at a very narrow wavelength interval) are perceived as being fully saturated and they fall on the curved boundary of the CIE chromaticity diagram. As there are no spectral purples, maximally saturated purples result from a combination of a large-wavelength, spectral red and a small-wavelength, spectral violet; they fall on the straight portion of the boundary of the CIE chromaticity diagram.

A chromaticity space for the trichromatic case

As mentioned, *colours* are taken to be points in the cube $[0,1]^3 \subset \mathbf{R}^3$. Each point $[x,y,z] \in \mathbf{R}^3$ can be uniquely expressed as the sum of the point $[\frac{x+y+z}{3}, \frac{x+y+z}{3}, \frac{x+y+z}{3}]$ on the *achromatic line* (i.e. the line "x=y=z") plus the point $[\frac{2x-y-z}{3}, \frac{-x+2y-z}{3}, \frac{-x-y+2z}{3}]$ on the plane perpendicular to that line.

We take the luminance of a colour point to be the l_1 distance from the point to the origin, that is, the luminance Λ of the colour point [x, y, z] is taken to be x + y + z. The colour points with a luminance of 1 determine the triangle with vertices [100], [010] and [001] that is (in a plane that is an affine space) also perpendicular to the achromatic segment. We call this the *chromaticity triangle*.

Although you could define the chromaticity of a colour point as the projection $\frac{1}{3}[2x - y - z, -x + 2y - z, -x - y + 2z]$ of the colour point on the plane Π^2 through the origin that is perpendicular to the achromatic line, which has the advantage of being a linear transformation $\mathbf{R}^3 \to \Pi^2$, we work with the chromaticity triangle and define the *chromaticity* κ of each nonblack (\neq [000]) colour point [x, y, z] to be the radial projection with respect to the origin of the colour point on the *chromaticity triangle*, see Fig. 1; thus, $\kappa = \frac{1}{\Lambda}[x, y, z] = [\frac{x}{x+y+z}, \frac{y}{x+y+z}, \frac{z}{x+y+z}]$. The definition has the disadvantage that radial projection is not a linear transformation and the advantage that chromaticities are points in the cube; we will further abbreviate and say that the chromaticity is the projection $\left[\frac{x}{x+y+z}, \frac{y}{x+y+z}, 0\right] = "\left[\frac{x}{x+y+z}, \frac{y}{x+y+z}\right]$ " of this point on the chromaticity triangle on the "plane xy".

All achromatic points with a defined chromaticity have the same chromaticity, $\kappa = [1/3, 1/3, 1/3] = "[1/3, 1/3]$ ", and we say that that is also the chromaticity of black.

Now let us relate this with the reflectances-to-colours map ϕ . The chromaticity of the colour [L, M, S] due to a reflectance function ρ is given by

$$\begin{split} \kappa &= \frac{1}{\Lambda} [L, M, S] = \frac{1}{L + M + S} [L, M, S] \\ &= \frac{1}{\int \rho(\lambda) [l(\lambda) + m(\lambda) + s(\lambda)] d\lambda} . \\ &\quad . \ [\int \rho(\lambda) l(\lambda) d\lambda, \int \rho(\lambda) m(\lambda) d\lambda, \int \rho(\lambda) s(\lambda) d\lambda] \end{split}$$



Figure 1. The point marked \times on the chromaticity triangle is the chromaticity of the colour point [L, M, S].

The object-colour solid

As said, the colour points in the cube result from a linear transformation $[0,1]^{[\lambda_0,\lambda_1]} \rightarrow [0,1]^3$ of the space of reflectance functions, that we assume to be full-rank, with functional image given therefore by a solid (i.e. a 3D) subset of the cube, see Fig. 6. In this solid, known as the *object-colour solid*, for each given chromaticity, we are interested in the colour point with highest luminance, and a corresponding reflectance function it may have arisen from. Clearly, these colour points of highest luminance (Schrödinger's colours) lie on the boundary of the object-colour solid.

More precisely, the space of all spectral reflectance functions $A \subset [0,1]^{[\lambda_0,\lambda_1]}$, given by the reflectance functions taking values in the interval [0, 1], for which the integrals in the definition of ϕ exist, gives rise to the object-colour solid $\phi([0,1]^{[\lambda_0,\lambda_1]}) \in [0,1]^3$ which consists of the colours of non-luminous surfaces.

Being the linear image of a convex set, the object-colour solid is a convex set. Also, ϕ is a closed map (e.g. with respect to the product metric for the space of reflectances) and the object colour solid is a closed set as well; also, since the integrals of *l*, *m* and *s* are finite, the object-colour solid is bounded. The black point is an exception in the sense that it belongs to the boundary of the object colour solid yet it has zero luminance. Being the object-colour solid convex, for each non-black point on the boundary of the solid, a line containing the point and the black point intersects no other point on the boundary of the solid. Assuming the *lms* functions to be nonnegative, the object-colour solid lives in the

first octant on \mathbf{R}^3 and then, for each chromaticity there is at most one point in the object-colour solid with a maximal luminance.

A hypothetical case

Consider the, unrealistic but illustrative, case of three Gaussian photoreceptor functions l, m and s, such as those shown in Fig. 2. The resulting map ϕ is then full rank, and the binary (taking only the values 0 or 1) reflectance functions, with at most two transitions, are then precisely those that are mapped to the boundary of the solid, as shown in Fig. 6.

For Gaussians of a small "variance", with a very low degree of overlap, bandpass reflectance functions (see Fig. 9) of small bandwidth will map to points pretty much near the three axes, as illustrated in Fig. 3; thus, three such bandpass reflectance functions, centred near the peak lambdas of the Gaussians, will span the whole cube; also, near the origin, the map ϕ will be locally onto. This points to a complete set of chromaticities near black. On the other hand, if the Gaussians have a fair degree of overlap so that the L and M, as well as the M and S, bells overlap and not so much the L and S bells, then, near 0, thin bandpass functions will map to the L and S axes and to diagonals in the LM and MS planes.

Likewise, the chromaticities of the spectral lights corresponding to the spectral colour loci of Fig. 3 are shown in Fig. 5.

For stop-band functions (see Fig. 10) it is more interesting to consider the viewpoint of the white colour point [111]. The narrowness of the bell curves then makes thin stop-band functions to map to points near the white point along the edges of that corner of the colour cube.



Figure 2. For illustration purposes, a set of three Gaussian photo receptor functions is assumed.

The high- and low-pass reflectance functions (see Figs. 7 and 8) are mapped by ϕ into the boundary of the solid in a more dense fashion than in the case of the band- and stop-band reflectance functions, as seen in Fig. 6, due to the fact that approximately the same number of functions was considered for each of the four types, in a Monte Carlo simulation. A low-pass type reflectance function can be seen as an extreme case of either a band-pass function or of a stop-band function and the same holds for highpass functions. In Fig. 6, at the blue and red lines, the curvature of the boundary of the solid is higher than at the green and violet caps; these lines of high curvature are the intersections of two cap surfaces. These surfaces could be extended: a low pass function can be seen as well as a transition from a band-pass function to a similar function that takes the value -1 at the large-wavelength rejection part (not a physical reflectance function, an element of $\mathbf{R}^{[350,750]}$ but not of $[0,1]^{[350,750]}$).

Even though clearly the reflectance-to-colour function ϕ is not injective, it is very interesting that the {0,1}-valued reflectance functions with at most two transitions are injectively mapped onto the boundary of the solid; since through each chromaticity point in the chromaticity triangle in Figure 1, precisely one ray from the origin passes, the farthest intersection with the boundary of the solid will correspond to a unique reflectance function. In this sense, there is no metamerism for colours of largest luminance.

The solid of the colours of reflectance functions (The object-colour solid)

The functional image under (1) of the set $[0,1]^{[350,750]}$ of the set of functions with domain [350nm,700nm] and range [0,1], is a 3D subset of $[0,1]^3$, called the object-colour solid; that is, the set of colours corresponding to arbitrary reflectance functions.

The linear image of a convex set is convex and the set of reflectance functions is convex (if $\rho_0(\lambda)$ and $\rho_1(\lambda)$ are two reflectance functions, then $\alpha \rho_0(\lambda) + (1 - \alpha)\rho_0(\lambda)$, $\alpha \in [0, 1]$ is a



Figure 3. The spectrum locus in the lms cube: the colours corresponding to very small-bandwithd pass-band reflectance functions, for a set of very thin (left), and wider (right), gaussian lms functions.



Figure 4. The spectrum locus can also be obtained, even if scaled and very close to the origin, as the colours corresponding to band-pass reflectance functions of very small band-width. Above, for bandwidths of 1 nm and, below, for bandwidths of 10 nm. Notice the difference in scale.

IS&T International Symposium on Electronic Imaging 2016 Image Processing: Algorithms and Systems XIV line of reflectance functions), therefore, the object-colour solid is a convex subset of the colour cube.

Reflectance functions with only one down transition are said to be of the *low-pass* type, those with only one up transition are said to be of the *high-pass* type; those with an up transition followed (i.e. at a larger wavelength) by a down transition are are said to be of the *pass-band* type, while those with a down transition followed by an up transition are are said to be of the *stop-band* type. Low-pass type and high-pass type reflectance functions are each described by a unique parameter (their "cut wavelength") and correspondingly are mapped to lines (1D sets) in the boundary of the solid (the blue line in Figure 6), while pass-band and stop-band type reflectance functions are each described by precisely two parameters (their high and low "cut wavelengths") and correspondingly are mapped to surfaces (2D sets) in the boundary of the solid (the green and purple surfaces in Figure 6).

Schrödinger reflectance functions.

According to Schrödinger, the (pigment) reflectance functions of surfaces that produce the most luminous colours (*grösster Leuchtkraft*) are binary (taking values in the doubleton $\{0,1\}$), and having at most two transitions; we call them *Schrödinger reflectance functions*. There are five types of Schrödinger reflectance functions: (degenerate) all-pass, of which there are only two: the zero reflectance which baps to the black colour point and



Figure 5. The chromaticities (projection on the L, M plane) corresponding to the spectrum loci shown in Fig. 3, in red for wider bell functions Ims and in blue for thinner bells.



Figure 6. Consider the object colour solid shown, arising from hypothetical gaussian functions l, m and s. All-pass type reflectance functions are sent to achromatic colours; among those only two take values in $\{0, 1\}$ and are mapped to the colours black and white, which are at vertices of the object-colour solid. Low-pass type reflectance functions (illustrated in blue) form a line of high curvature of the boundary of the solid, likewise for high-pass functions, which are illustrated in red. Band-pass type functions, illustrated in green, and stop-band type functions, illustrated in violet, form the remaining 2D caps of the solid.

the one reflectance function which maps to the white colour point, low-pass (with respect to the wavelength domain rather than in the electromagnetic-frequency domain), as in Fig. 7, high-pass as in Fig. 8, band-pass as in Fig. 9 and stop-band, as in Fig. 10.

High-pass and low-pass reflectance functions are oneparameter, reflectance-function families, with parameter given by the transition wavelength at which the function jumps; they give rise to the so-called *end colours*, they lack energy contents either from a low-wavelength range or a large-wavelength range of the spectrum. On the other hand, the band-pass and stop-band types of reflectance functions are two-parameter families, the parameters being given by the two transition wavelengths, and correspond to the Mittelpigmente and the Mittelfehlpig*mente* colours, respectively. For nonconstant l, m, s functions, Schrödinger reflectance functions map in a one-to-one fashion to the object-colour solid, in fact to the boundary of the solid unless the lms happen to take on negative values as well. Thus, the end colours correspond to lines in the boundary of the solid and the Mittelpigmente and Mittelfehlpigmente correspond to surfaces on the boundary of the object-colour solid, see Fig. 6.



Figure 7. A low-pass type reflectance function; low-pass reflectance functions correspond to end colours of the blue kind.



Figure 8. A high-pass type of reflectance function; high-pass reflectance functions correspond to end colours of the red kind.



Figure 9. A band-pass-pass type of reflectance function; band-pass reflectance functions correspond to Mittelpigmente.

Schrödinger's proof

Assuming a flat spectral illuminant, so that the product illuminanat-times-reflectance is essentially the reflectance, and in the context of the receptoral colour cube *lms*, a version of Schrödinger's proof that the *grösster Leuchtkraft* pigments of optimally coloured surfaces have binary spectral reflectances, with at most two transitions goes along the following lines.

Let $\rho(\lambda)$ be the spectral reflectance of the optimally coloured surface and let $l(\lambda), m(\lambda)$ and $s(\lambda)$ be the (nonnegative) spectral receptoral curves of a trichromatic colour vision system e.g. the three cone responses of a human or a bee, or the three types of pixel in a Bayer mosaic. The transformation ϕ that maps $\rho \mapsto [L, M, S] \in \mathbb{R}^3$, from spectra into colours, is given by

$$[L, M, S] = \left[\int l(\lambda)\rho(\lambda)d\lambda, \int m(\lambda)\rho(\lambda)d\lambda, \int s(\lambda)\rho(\lambda)d\lambda\right]$$
(1)

or, for short, $[L,M,S] = \int [l(\lambda), m(\lambda), s(\lambda)]\rho(\lambda) d\lambda$, and it is assumed that $\int [l(\lambda), m(\lambda), s(\lambda)] d\lambda = [1, 1, 1]$.

The chromaticity $\kappa = [\bar{L}, \bar{M}, \bar{S}] = \frac{1}{L+M+S} [L, M, S]$ of a colour point [L, M, S] is a point that lies in the first octant of \mathbf{R}^3 , in fact in the triangle with vertices [100], [010] and [001]; see Figure 1; it is the *radial projection* of the colour point on the chromaticity triangle, with respect to the black point or origin of \mathbf{R}^3 .

Since $\overline{L} + \overline{M} + \overline{S} = 1$, for short, the chromaticity can be unambiguously shortened to the pair $[\overline{L}, \overline{M}]$, which is the orthogonal projection of the chromaticity on the *L*-*S* plane. It is claimed that a reflectance function $\rho(\lambda)$ giving rise to a colour [L, M, S] with a given chromaticity κ and maximal luminosity L + M + S must meet the two conditions:

- $\forall \lambda \in [\lambda_{min}, \lambda_{max}] \rho(\lambda) \in \{0, 1\}.$
- The maximal number of discontinuities of $\rho(\lambda)$ is 2.

That is, provided that ρ is a Schrödinger reflectance function. To each low-pass and to each high-pass reflectance functions there corresponds a transition frequency λ_t ; likewise, to each band-pass and to each stop-band reflectance functions there corresponds two transition frequencies λ_L and λ_H . A low-pass function is a band-pass function with $\lambda_L = \lambda_{min}$ and a stop-band function with $\lambda_H = \lambda_{max}$; meanwhile a high-pass function is a stop-band function with $\lambda_L = \lambda_{min}$ and a band-pass function with $\lambda_H = \lambda_{max}$.

The proof is by contradiction. Assume that the reflectance curve ρ is mapped to the colour point $\mathbf{c} := [L, M, S] = \phi(\rho)$ on the



Figure 10. A band-pass-pass type of reflectance function; band-pass reflectance functions correspond to Mittelfehpigmente.

boundary of the object-colour solid and that, at three wavelengths λ_1 , λ_2 and λ_3 , ρ is not $\{0,1\}$ -valued, that is $0 < \rho(\lambda_i) < 1$, see Fig. 11; assume further that the matrix

$$\left[\begin{array}{ccc} l(\lambda_1) & l(\lambda_2) & l(\lambda_3) \\ m(\lambda_1) & m(\lambda_2) & m(\lambda_3) \\ s(\lambda_1) & s(\lambda_2) & s(\lambda_3) \end{array}\right]$$

is invertible and that (1) can be expressed in a discrete fashion as

$$[L,M,S] = [\sum_{n} l(nT)\rho(nT), \sum_{n} m(nT)\rho(nT), \sum_{n} s(nT)\rho(nT)]$$
(2)

for some small length T of the "sampling interval"; to get this approximation you may use Poisson's summation formula and assume that the integrands in (1) are bandlimited. A modified colour αc

$$\begin{aligned} \alpha[L,M,S] &= (1+\varepsilon)[L,M,S] \\ &= [L,M,S] + \varepsilon[L,M,S] \\ &= [L,M,S] + [\delta_L,\delta_M,\delta_S], \end{aligned}$$

with the same chromaticity as that of [L, M, S], and a larger luminance (assuming $\alpha > 1$) can then be achieved by slightly changing the reflectance function ρ in neighbourhoods of λ_1 , λ_2 and λ_3 , in accordance to the results of





Figure 11. At the indicated wavelengths, the reflectance function ρ is strictly in between 0 and 1.

Thus, with suitable changes in the corresponding terms of the sums in (2), a new colour $\mathbf{c}' := [L', M', S'] = \alpha[L, M, S]$ is obtained that has the same chrominance as \mathbf{c} and larger luminance; since the object colour solid is convex and contains the black point, this shows that the colour point was not on the boundary, contradicting the assumption. Since the linear transformation (1) is continuous, if the δ 's are small enough, you still have $0 < \rho(\lambda_i) < 1$, i = 1, 2, 3. This shows that ρ must be

 $\{0,1\}$ -valued.

As for the number of discontinuities having a maximum of two, there is not a corresponding linear transformation to be inverted, but a smooth transformation that can be locally linearised. In fact, for a fixed, binary $\rho(\lambda)$ with three discontinuities, at λ_1 , λ_2 and λ_3 , consider the smooth function $f : \mathbf{R}^3 \to \mathbf{R}^3$ given by $f(\lambda_1, \lambda_2, \lambda_3) =$

$$\int (\lambda_1, \lambda_2, \lambda_3) = [\int l(\lambda) \rho(\lambda) d\lambda, \int m(\lambda) \rho(\lambda) d\lambda, \int s(\lambda) \rho(\lambda) d\lambda] = [L, M, S]$$

let **J** be the Jacobian matrix of the transformation at $[\lambda_1, \lambda_2, \lambda_3]$, then, by inverting **J** (which assumes the function is not locally constant), as in the previous case, you can find δ_i 's such that

$$f(\lambda_1 + \delta_1, \lambda_2 + \delta_2, \lambda_3 + \delta_3) = \alpha[L, M, S], \alpha > 1.$$



Figure 12. The case of a $\{0,1\}$ -valued reflectance function with more than two transitions. A reflectance function can then be found such that the corresponding color has the same chromaticity and the higher luminance.



Figure 13. The points at the boundary of the object-colour solid with their corresponding colours.

Comments on Schrödinger's proof

Schrödinger's proof is based on the linearity of the reflectance-to-colour transformation, on the linear independence of the photoreceptor spectral responses and on the convexity of the set of reflectance functions.

On the linear independence of the *lms* functions

• The assumption that for each three λ 's the corresponding three vectors $[l(\lambda_i), m(\lambda_i), s(\lambda_i)]$, $i \in \{1, 2, 3\}$, be linearly independent is equivalent to requiring that the vectors $[l(\lambda_1), l(\lambda_2), l(\lambda_3)]$, $[m(\lambda_1), m(\lambda_2), m(\lambda_3)]$ and $[s(\lambda_1), s(\lambda_2), s(\lambda_3)]$ be linearly independent, and is stronger than requiring that the functions $l, m, s : [370, 730] \rightarrow \mathbf{R}$ be linearly independent.

• Thus, we say that the functions l, m and s are strongly linearly independent if for each three λ 's the vectors $[l(\lambda_1), l(\lambda_2), l(\lambda_3)]$, $[m(\lambda_1), m(\lambda_2), m(\lambda_3)]$ and $[s(\lambda_1), s(\lambda_2), s(\lambda_3)]$ are linearly independent. This is equivalent to saying that no plane through the origin contains more than two points of the curve $\{(l(\lambda), m(\lambda), s(\lambda)) \in \mathbf{R}^3 | \lambda \in [370, 730]\}$, called the spectral locus. In our proof, the functions *lms* are assumed to be strongly linearly independent.

On the convexity of the set bounded by the image of Schrödinger reflectance functions

Consider the images $\phi(\rho)$ of Schrödinger reflectance functions ρ . Suppose that ρ is a low pass function with cut off wavelength λ_c . Then,

$$\phi(oldsymbol{
ho}) = \int_{oldsymbol{\lambda}_{min}}^{oldsymbol{\lambda}_{c}} [l(oldsymbol{\lambda}), m(oldsymbol{\lambda}), s(oldsymbol{\lambda})] doldsymbol{\lambda}$$

Likewise, for a band-pass reflectance function with cut off frequencies λ_{LO} and λ_{HI} ,

$$egin{aligned} \phi(oldsymbol{
ho}) &= \int_{\lambda_{LO}}^{\lambda_{HI}} [l\left(\lambda
ight), m\left(\lambda
ight), s\left(\lambda
ight)] d\lambda \ \end{aligned}$$
 $&= \int_{\lambda_{min}}^{\lambda_{HI}} [l\left(\lambda
ight), m\left(\lambda
ight), s\left(\lambda
ight)] d\lambda - \int_{\lambda_{min}}^{\lambda_{LO}} [l\left(\lambda
ight), m\left(\lambda
ight), s\left(\lambda
ight)] d\lambda \end{aligned}$

that is, the response to a band-pass function is also the difference of responses to corresponding low-pass functions.

First, consider separately the first and second components of the resulting colour point. For the case of a low-pass function you get, as a function of λ_c ,

$$[L(\lambda_c), M(\lambda_c)] = [\int_{\lambda_{min}}^{\lambda_c} l(\lambda) d\lambda, \int_{\lambda_{min}}^{\lambda_c} m(\lambda) d\lambda]$$

This implicitly defines a function $\mu(L)$ for $L \in \{L(\lambda) | \lambda \in \{\lambda_{min}, \lambda_{max}\}\}$, given by $\mu(L(\lambda_c)) = M(\lambda_c)$. The derivative of this function is given by

$$\mu'(L(\lambda)) = \frac{M'(\lambda)}{L'(\lambda)} = \frac{m(\lambda)}{l(\lambda)}$$

and has second derivative

$$\mu''(L(\lambda)) = \frac{M''(\lambda) - \mu'(L(\lambda))L''(\lambda)}{[L'(\lambda)]^2} = \frac{m'(\lambda) - \frac{m(\lambda)}{l(\lambda)}l'(\lambda)}{l^2(\lambda)}$$

For high-pass functions $\int_{\lambda_c}^{\lambda_{max}} [l(\lambda), m(\lambda)] d\lambda$, the case is entirely similar and, since

$$\int_{\lambda_c}^{\lambda_{max}} [l(\lambda), m(\lambda)] d\lambda + \int_{\lambda_{min}}^{\lambda_c} [l(\lambda), m(\lambda)] d\lambda = [1, 1]$$

or

$$\int_{\lambda_{min}}^{\lambda_c} [l(\lambda), m(\lambda)] d\lambda = [1 - \int_{\lambda_c}^{\lambda_{max}} [l(\lambda), 1 - m(\lambda)] d\lambda]$$

the derivatives of the function μ for the high-pass case are the negative of the derivatives for the low-pass case; in particular, the second derivative. Therefore, if one of the functions is convex

IS&T International Symposium on Electronic Imaging 2016 Image Processing: Algorithms and Systems XIV the other is concave. also, since

$$[L(\lambda_{min}), M(\lambda_{min})] = [0,0]$$
 and $[L(\lambda_{max}), M(\lambda_{max})] = [1,1],$

for the low-pass case and

$$[L(\lambda_{min}), M(\lambda_{min})] = [1, 1]$$
 and $[L(\lambda_{max}), M(\lambda_{max})] = [0, 0],$

for the high-pass case. Therefore, the curves that are the graph of μ for the low-pass and high-pass cases have common endpoints and (for bell shaped functions *l*, *m* and *s*) bound a convex set of the plane.

Now consider the general, tridimensional case, with band-pass and stop-band functions; note in passing that, for $\{0, 1\}$ -valued functions, stop-band = 1 - band-pass. The relation

$$\begin{bmatrix} L(\lambda_{LO}, \lambda_{HI}), M((\lambda_{LO}, \lambda_{HI}), S((\lambda_{LO}, \lambda_{HI}))] := \\ \int_{\lambda_{HO}}^{\lambda_{HI}} [l(\lambda), m(\lambda), s(\lambda)] d\lambda,$$

with $(\lambda_{LO}, \lambda_{HI}) \in [\lambda_{min}, \lambda_{max}] \times [\lambda_{min}, \lambda_{max}]$, defines a surface in tridimensional space; in fact, the relation implicitly defines two functions v(L, M), with $v(L(\lambda_{LO}, \lambda_{HI}), M(\lambda_{LO}, \lambda_{HI})) =$ $S(\lambda_{LO}, \lambda_{HI})$, one for the case of band-pass functions and the other for the stop-band functions; one of this two functions is convex and the other concave and their graphs bound a convex, solid subset. Thus, let v(L, M), with $(L, M) \in$

 $\{(L(\lambda_{LO}, \lambda_{HI}), M(\lambda_{LO}, \lambda_{HI})), \lambda_{min} \leq \lambda_{LO} \leq \lambda_{HI} \leq \lambda_{max}\}$

be given by

$$\mathcal{W}(L(\lambda_{LO}, \lambda_{HI}), M(\lambda_{LO}, \lambda_{HI})) = S(\lambda_{LO}, \lambda_{HI})$$

The Hessian indicates that for band-pass reflectance functions v is concave and for stop-band reflectance functions v is convex, together, they bound a convex solid.

• The assumption that the colour αc is achievable in turn assumes that the colour object solid is convex, which it is.

Abusing the notation again, put

$$[L(\lambda), M(\lambda), S(\lambda)] := \int_{\lambda_{min}}^{\lambda} [l(\eta), m(\eta), s(\eta)] d\eta$$

In the illustration given above of gaussian functions l, m and s, these functions are textiterror functions. Clearly, for low pass and high-pass reflectance functions, the function ϕ is not only locally injective but injective as well since, in these cases $\phi(\rho) = [L(\lambda), M(\lambda), S(\lambda)]$, or $\phi(\rho) = [1 - L(\lambda), 1 - M(\lambda), 1 - S(\lambda)]$, respectively. We *conjecture* that for band-pass reflectance functions ϕ is also injective for the human photoreceptor functions l, m and s. In fact, we give an equivalent condition:

For no set $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4 | \lambda_1 < \lambda_2, \lambda_3 < \lambda_4\}$ it holds that: $L(\lambda_2) - L(\lambda_1) = L(\lambda_4) - L(\lambda_3),$ $M(\lambda_2) - M(\lambda_1) = M(\lambda_4) - M(\lambda_3),$ and $S(\lambda_2) - S(\lambda_1) = S(\lambda_4) - S(\lambda_3).$

This condition should be checked empirically for the human photopigment spectral absorbance functions. For stopband reflectance functions, the condition is the same since $\phi(stopband) = [1,1,1] - \phi(band pass)$, completing the *conjecture*. If we imagine more curves like those in Fig. 4, for all possible bandwidths, the idea is that they cover the band-pass cap of the object colour solid with paths that have the shape of the spectrum locus, without intersecting; see Fig. 14.

• The boundary of the object colour solid is given by the image of the set of the Schrödinger reflectance functions ({0,1}-valued functions with at most two transitions); they map in a one-to-one fashion into the boundary of the object-colour solid, given to the fact that the three photoreceptor functions are never simultaneously constant at two wavelengths. Notice that the straight "line" segment $\alpha \rho_1(\lambda) + (1-\alpha)\rho_2(\lambda)$, between two Schrödinger's reflectance functions ρ_1 and ρ_2 has in general 3 levels (0, 1, α and $1-\alpha$) and as many as 4 transitions.



Figure 14. The bandpass cap of the boundary of the colour object solid is a disjoint union of the images of bandpass functions of all bandwidths. Likewise for the stop-band cap.

• For lights, you can achieve highly discontinuous spectra using nonlinear optics. Less so (as far as the authors know) for pigmente surfaces which, for natural objects, are rather smooth.

An Alternate Definition of Chromaticity

The definition of chromaticity that we have been using is based on the assumption that all colours on a ray from the origin [000], the *black colour point*, have the same chromaticity; i.e. that all colours on a shortest path from a given colour point to the origin have the same chromaticity.

Alternatively, consider the analogous definition, with respect to the white colour point [111]. In the previous case, you want the point in the triangle of the plane L+M+S=1, in the first octant, given by the intersection with the ray from the origin trough a given (non-black: in fact, notice that black is not a chromaticity in this chromaticity space) colour point; now, consider the point in the triangle of the plane L+M+S=2, that fits in the cube $[0,1]^3$, given by the intersection with the ray from the white point [111] trough a given (non-white: white will not be a chromaticity is this space) colour point.

The point in question is given by $\alpha[L,M,S]$ + $[1-\alpha, 1-\alpha, 1-\alpha]$, with α given by $\frac{1}{3-(L+M+S)}$, and the

chromaticity is given by

$$\kappa' = rac{1}{3 - (L + M + S)} \left[2 - M - S, 2 - L - S, 2 - L - M
ight]$$

Now white is not a chromaticity. Call the triangles with vertices [100], [010] and [001], and, [011], [101] and [110], the *dark* and *light chromaticity triangles*.

In the RGB cube, the luminance is a measure of the distance from a color point to the black point, and the chromatic saturation is a measure of the distance to the achromatic segment (the intersection of the cube and the ray from the origin through the white point). Both chromaticity triangles are orthogonal to the achromatic segment.

How oblong the color solid is, in particular near the white and black points, depends on the width and overlap of the bell shaped curves of the photoreceptors. If near the origin the boundary of the solid is parallel to the walls of the cube, all colours, except black have a defined chromaticity, likewise for the light chromaticity: if the boundary of the colour solid near the white point is parallel to the walls of the cube, all colour points, except white, have a defined light chromaticity.

MOST COLOURFUL TETRACHROMATIC PIG-MENTS

We extend the concept of most colourful surfaces to the tetrachromatic case where four photoreceptor types are assumed. Applications may be found in tetrachromatic computer vision and satellite imaging.

Let $w(\lambda)$, $x(\lambda)$, $y(\lambda)$ and $z(\lambda)$, be the spectral responses of four photoreceptor types. To a given reflectance function $\rho(\lambda)$, assuming an illumination with a flat spectrum, there corresponds the point $[W, X, Y, Z] \in \mathbf{R}^4$, called its tetrachromatic colour, as follows

$$[W, X, Y, Z] = \left[\int w(\lambda) \rho(\lambda) d\lambda, \int x(\lambda) \rho(\lambda) d\lambda, (3)\right]$$

$$\int y(\lambda) \rho(\lambda) d\lambda, \int z(\lambda) \rho(\lambda) d\lambda$$

where it is assumed that $[\int w\rho, \int x\rho, \int y\rho, \int z\rho] = [1,1,1,1]$. As before, the map $\rho \mapsto [W,X,Y,Z]$, given by (3), is a linear transformation $[0,1]^{[300,800]} \subset \mathbb{R}^{[300,800]} \to [0,1]^4 \subset \mathbb{R}^4$. (We are abusing notation here as we are assuming that the integrals exist.)

In the hypercube $[0,1]^4$ of colour points, the luminance of a colour [W,X,Y,Z] is taken to be the sum $\Lambda = W + X + Y + Z$ of its components, and the chromaticity is taken to be the radial projection with respect to the origin, on the tetrahedron with vertices [1000], [0100], [0010] and [0001], that is, the point $\kappa = \frac{1}{\Lambda}[W,X,Y,Z]$. Since only 3 of the components of the chromaticity specify it, we write for short $[\kappa_1, \kappa_2, \kappa_3] := [\frac{W}{\Lambda}, \frac{X}{\Lambda}, \frac{Y}{\Lambda}]$. The chrominance is the intersection of the line from the origin to the colour point and the 3-space given by

$$\{[W, X, Y, Z] : W + X + Y + Z = 1\}.$$

The intersection of this affine 3-space and the ++++ 16-tant is the tetrahedron with vertices [1000], [0100], [0010] and [0001]; it is indicated in Figure 17.

In the trichromatic case, the chromaticity of a colour is a (2D, planar) triangular variable and the hue is a cyclic [5], [9], 1D variable; in the tetrachromatic case, the hue is a spherical [6], 2D variable and chromaticity is a (3D, spatial) tetrahedrical variable.



Figure 15. At the wavelengths where a reflectance function takes values strictly less than 1 and strictly above 0, a new reflectance function of higher luminosity and of the same chromaticity can be obtained.

As in the trichromatic case, assume that the *colour function* (3): $\mathbf{R}^{[\lambda_{min}, \lambda_{max}]} \rightarrow \mathbf{R}^4$ can be approximated as

$$[W,X,Y,Z] = \begin{bmatrix} w(\lambda_1) & w(\lambda_2) & \dots & w(\lambda_N) \\ x(\lambda_1) & x(\lambda_2) & \dots & x(\lambda_N) \\ y(\lambda_1) & y(\lambda_2) & \dots & y(\lambda_N) \\ z(\lambda_1) & z(\lambda_2) & \dots & z(\lambda_N) \end{bmatrix} \begin{bmatrix} \rho(\lambda_1) \\ \rho(\lambda_2) \\ \dots \\ \rho(\lambda_N) \end{bmatrix}$$

for a suitably large N.



Figure 16.

The reflectance functions that, for a given chrominance, give rise to most luminous surfaces are binary, that is, for each wavelength, the surface reflects all the energy or absorbs all the energy. To show this, assume that at 4 wavelengths λ_1 , λ_2 , λ_3 and λ_4 , and a small wavelength interval around each of them, the reflectance function is strictly between 0 and 1. Then, as in the trichromatic case, the luminance can be increased without changing the chrominance. Let $\rho(\lambda)$ be such that $0 < \rho(\lambda_{i(1)}), \rho(\lambda_{i(2)}), \rho(\lambda_{i(3)}), \rho(\lambda_{i(4)}) < 1$ and have colour [W, X, Y, Z]; a $\rho'(\lambda)$ can be found by slightly changing the value of ρ at the λ_i 's that has the same chrominance as ρ and a larger luminance. To do this write, for some $\alpha > 1$,

$$\begin{aligned} \alpha[W, X, Y, Z] &=: (1 + \varepsilon)[W, X, Y, Z] \\ &=: [W, X, Y, Z] + [\delta_W, \delta_X, \delta_Y, \delta_Z], \end{aligned}$$
where

$$[\delta_W, \delta_X, \delta_Y, \delta_Z] = [\varepsilon W, \varepsilon X, \varepsilon Y, \varepsilon Z],$$

Then write

 $\left[egin{array}{c} \Delta
ho'(\lambda_{i(1)}) \ \Delta
ho'(\lambda_{i(2)}) \ \Delta
ho'(\lambda_{i(3)}) \ \Delta
ho'(\lambda_{i(3)}) \ \Delta
ho'(\lambda_{i(4)}) \end{array}
ight]$

	$w(\lambda_{i(1)})$	$w(\lambda_{i(2)})$	$w(\lambda_{i(3)})$	$w(\lambda_{i(4)})$] -1 [δ_W
=	$x(\lambda_{i(1)})$	$x(\lambda_{i(2)})$	$x(\lambda_{i(3)})$	$x(\lambda_{i(4)})$		δ_X
	$y(\lambda_{i(1)})$	$y(\lambda_{i(2)})$	$y(\lambda_{i(3)})$	$y(\lambda_{i(4)})$		δ_Y
	$z(\lambda_{i(1)})$	$z(\lambda_{i(2)})$	$z(\lambda_{i(3)})$	$z(\lambda_{i(4)})$		δ _Z

and put $\rho'(\lambda_k) = \rho(\lambda_k) + \Delta \rho(\lambda_k)$ for k = i(1), i(2), i(3), i(4), and $\rho'(\lambda_k) = \rho(\lambda_k)$ otherwise.



Figure 17. The tetrahedron indicated in red is part of the 3-affine space that contains the orthonormal stardard basis for \mathbf{R}^4 .

In the tetrachromatic case as well, reflectance functions corresponding to the boundary of the object-colour hyper-solid are $\{0,1\}$ -valued and may have up to 3 transitions, so they may be of 7 types: all-pass, low-pass, high-pass, pass-band and stop-band as in the trichromatic case, but also low-pass with a large-wavelength pass band, and high-pass with an extra pass band at small wavelengths. Thus, there are 0-parameter reflectance functions, i.e. the constant (all-pass) 0 (giving rise to the black colour point) and 1 (giving rise to the white colour point) functions, that give rise to two vertices of the color hyper solid; one-parameter reflectance functions, i.e. the low-pass and high-pass functions, that give rise to two lines of high 3D curvature; the two parameter reflectance functions given by the pass-band and stop-band functions, which give rise to 2D, high-3D-curvature components of the boundary of the hyper solid, and the 3-parameter functions given by the highpass with lower band, and low-pass with higher band, which give rise to two 3D smooth caps at the boundary.

For tetrachromatic vision systems, the condition on the receptor functions w, x, y and z, for the injectiveness of ϕ for the Schrödinger reflectance functions reads as follows, for the case of bandpass-plus-highpass functions: For no set $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6 | \lambda_1 < \lambda_2 < \lambda_3, \lambda_4 < \lambda_5 < \lambda_6, \}$ it holds that:

 $\begin{array}{l} W(\lambda_2) - W(\lambda_1) + W(\lambda_{max}) - W(\lambda_3) = W(\lambda_5) - W(\lambda_4) + W(\lambda_{max}) - W(\lambda_6), \\ X(\lambda_2) - X(\lambda_1) + X(\lambda_{max}) - X(\lambda_3) = X(\lambda_5) - X(\lambda_4) + X(\lambda_{max}) - X(\lambda_6), \\ Y(\lambda_2) - W(\lambda_1) + Y(\lambda_{max}) - Y(\lambda_3) = Y(\lambda_5) - Y(\lambda_4) + Y(\lambda_{max}) - X(\lambda_6), \\ m Z(\lambda_2) - Z(\lambda_1) + Z(\lambda_{max}) - Z(\lambda_3) = Z(\lambda_5) - Z(\lambda_4) + Z(\lambda_{max}) - Z(\lambda_6), \end{array}$

where, abusing the notation, $[W(\lambda), X(\lambda), Y(\lambda), Z(\lambda)] = \int_{\lambda_{min}}^{\lambda} [w(\eta), x(\eta), y(\eta), z(\eta)] d\eta$

Conclusion and Further Work

It is very interesting that there is no metamerism for colours of largest luminance (assuming the conjecture of the injectiveness of the reflectance-to-colour map ϕ for the Schrödinger reflectance functions stated above). This speaks highly of our vision system and says that for highly luminous surfaces we are "identifying" uniquely the spectral reflectance of the surface in the sense that to each chromaticity there corresponds a unique reflectance function. This comment must be sobered in two respects; for one thing, even though we discriminate many chromaticities, surely there is a limit to that, also, Schrödinger reflectances are perhaps very idealistic to be found in practice. Nevertheless it may have happened that the reason for frugivorous primates to be an exception among the mammals in being trichromatic be due to the importance of a pigment reflectance (of the skin? of a fruit?) having two step transitions, see Fig. 18.

The same holds for a dichromat but in that case he/she can only identify low-pass and high-pass reflectance functions giving rise to most luminous colours; a dichromat will not see as most luminous certain colours that a trichromat sees as such. On the other hand, a tetrachromat could then identify more Schrödinger (i.e. binary, $\{0,1\}$ -valued) reflectance functions: functions with up to 3 transitions.

A vision system, natural or artificial, equipped with four types of photoreceptor functions, each type responding best to different wavelength ranges, even if overlapping, can be modeled analogously to the trichromatic case. It may, or may not, be a good model of the corresponding qualia in the case of an animal.

Again, for tetrachromatic vision systems that meet the corresponding condition, most luminous colours of each chromaticity point out in a unique way to a binary spectral reflectance function having up to 3 transitions. See again Fig. 18

One way to describe a colour point in the cube is to give both its luminance and its chromaticity. Chromaticity and hue are different properties of colour. Hue is a simpler magnitude in the sense that it has one dimension less than chromaticity; however, achromatic colours do not have a hue yet *achromatic* is a chromaticity. For Runge-type, spherical colour spaces [7], [8], two variables can also describe a colour point: colourfulness and chromaticness; the cube (or hypercube) is deformed into a round ball (or hyperball) and using "polar" coordinates (r, θ), where the *colorfulness* r is the distance from the *central gray* colour point ([1/3, 1/3, 1/3] in the trichromatic case and [1/4, 1/4, 1/4, 1/4] in the tetrachromatic case) and the *chromaticness* θ is a point on the spherical boundary of the ball. The hue sphere sits in codimension 1, as an equator of the chromaticness sphere.

In computer vision, tetrachromatic colour gives a way to code spectral reflectances with four-coordinate vectors. Even though spectral reflectances of the surfaces of objects in nature are rather smooth in the range from 300 nm to 800 nm, that is

IS&T International Symposium on Electronic Imaging 2016 Image Processing: Algorithms and Systems XIV not always the case. These systems ought to be specialised and the spectral responses of the photoreceptors should be designed accordingly. In turn, the 4-number colours in the cube can be described in a shortened way as a combination of chrominance and luminance. More luminous colours are more salient for us humans and the design of most luminous surfaces may be useful in dealing with tetrachromatic animals. Likewise, the search for most luminous objects of a given tetrachromatic chromaticity may have applications in the recognition e.g.of flowers.

We are studying the property given by West and Brill of the convexity of the convexity of the chromaticity diagram, using our definition of chromaticity.

References

- [1] E.Schrödinger, *Theorie der Pigmente von grösster Leuchtkraft*, Annalen der Physik 4, 62 (1920) pp. 603-622.
- [2] D.L. MacAdam, Color Measurement, Theme and variations, 2nd ed., Springer, Berlin, 1985
- [3] G. West and M. Brill, Conditions under which Schrodinger object colors are optimal. JOSA letters, Vol. 73, No. 9; pp. 1223-1225, September, 1983.
- [4] J. Koenderink., Color for the Sciences. MIT press, 1990.
- [5] A. Restrepo, *Colour processing in Runge Space*. Proceedings SPIE Vol. 8295: Image Processing: Algorithms and Systems X; and Parallel Processing for Imaging Applications II, San Jose, CA, 2011.
- [6] A. Restrepo, *Tetrachromatic Colour Space*. Proceedings SPIE and IS&T Electronic Imaging, Vol. 8295: Image Processing: Algorithms and Systems X; and Parallel Processing for Imaging Applications II, San Francisco, February, 2012.
- [7] A. Restrepo *Hue Processing in Tetrachromatic Spaces*, SPIE and IS&T Electronic Imaging, San Francisco, Feb. 3-7, 2013
- [8] A. Restrepo Colour Processing in Tetrachromatic Spaces Uses of tetrachromatic colour spaces Visapp, Barcelona, Feb. 21-24, 2013
- [9] A. Restrepo and V. Estupiñán, Colour visualization of cyclic signals, Proceedings of the 2014 SPIE Electronic Imaging Conference, San Francisco, 2014.
- [10] A. Restrepo and J. Garzón, Metamerism in the context of aperture sampling reconstruction. Proceedings SPIE Vol. 9399: Image Processing: Algorithms and Systems XIII, San Francisco, February, 2015.
- [11] R. Menzel, Spectral sensitivity and color vision in invertebrates. In Handbook of Sensory Physiology Vol. VII/6A, Comparative Physiology and Evolution of Vision in Invertebrates A: Invertebrate Photoreceptors, H. Autrum, ed. (Springer-Verlag, New York, 1979), pp. 564-565.



Figure 18. The spectral absorbances of hemoglobin (taken from http://biochimica.bio.uniroma1.it/didattica/Fcourse/L02_Methods/LabMed.htm), and of chlorophyll (taken from https://en.wikipedia.org/wiki/Chlorophyll).

- [12] W. A. H. Rushton, *Pigments and signals in color vision*. J. Physiol. (London) 220, 1-31 (1972).
- [13] A. Restrepo and Edisson Maldonado. Con Respecto al Muestreo por Apertura. Simposio de Seales, Imagenes y Visin Artificial (STSIVA), 20th Symposium, September 2nd-4th, 2015.

Author Biography

Edisson Maldonado was born at Bogotá, Colombia in 1991; he obtained his High School diploma from Colegio Gustavo Restrepo and his B.Sc. in Electronics Engineering from the Pontificia Universidad Javeriana, at Bogotá, Colombia in 2014. Currently he is in enrolled in the M.Sc. program in Electronics and Computer Engineering, of the Universidad de los Andes at Bogotá, where he is a teaching assistant. His work and research interests include image processing and color theory.

Author Biography

Alfredo Restrepo received his B.Sc. in electronics engineering from the Pontificia Universidad Javeriana, at Bogotá, Colombia (1982) and his M.Sc. and Ph.D. in electrical engineering the University of Texas at Austin (1986, 1990). Since then he has worked at the Universidad de los Andes, in Bogotá. He has been a researcher at the ENST, Paris, France and at the Universitá degli Studi di Trieste. His work has been on nonlinear signal processing, image processing and colour; also on geometric topology.