Tri and Tetrachromatic Metamerism

Alfredo Restrepo

Laboratorio de Señales
Departamento de Ing. Eléctrica y Electrónica de la Universidad de los Andes
Carrera 1 No. 18A-70; oficina ML-427, Bogotá 111711, Colombia

ABSTRACT

Two light beams that are seen of the same colour and possibly having different spectra are said to be metameric. The colour of a light beam is based on the reading of several photodetectors with different spectral responses and metamerism results when a set of photodetectors is unable to resolve two spectra. From a mathematical point of view, metamerism can be characterized in terms of the L, M and S responses of the retina, in terms of the CIE X, Y and Z curves, etc. We are interested in exploring the concept of metamerism both in the trichromatic and in the tetrachromatic cases. Applications are in computer vision, computational photography and satellite imagery, for example. We make a case for chromatic metamerism, when the luminance is different and also hue metamerism when both the luminance and the saturation are different.

Keywords: Metamerism, trichromacy, tetrachromacy.

1. INTRODUCTION

The fact that most vision systems "aperture sample" the electromagnetic spectrum on each of several, usually overlapping, wavelength intervals, implies that there will be some loss of information and ambiguity. This inherent ambiguity is the cause of metamerism: light beams physically different, may be seen as of the same colour; for example, a spectral (i.e. of energy at a unique wavelength) yellow light beam and an appropriate combination of spectral beam lights, green and red. The alternative, to take more energy samples that are more localized in frequency, makes the vision system slower and bigger.

In our case, frugivorous primates, the aperture sampling of the spectrum is achieved by means of the three types of photoreceptor cells L, M and S that are active in day light in the receptor layer of our retinas. With only starlight illumination at night, under scotopic vision conditions, only one type of a photoreceptor is active in our retinas (the rods), and differences in wavelength are confused with differences in light intensity, and the wiring in our retinas precludes us from seeing colours.

1.1 The Mathematics of Metamerism

We work here under the paradigm of unrelated colours, i.e. we do not take into account contrast effects. Assume $P$ photodetectors with spectral responses $\varphi_p(\lambda)$, $p \in \{1, 2, \ldots, P\}$, that to a given light beam with spectrum $s(\lambda)$ produce response $c = [c_1, c_2, \ldots, c_P]$, with $c_p = \int s(\lambda)\varphi_p(\lambda)d\lambda$; we say that $c$ is a colour. This map $s \mapsto c$, gives a linear transformation $R_{[\lambda_{\min}, \lambda_{\max}]} \rightarrow \mathbb{R}^P$ where $R_{[\lambda_{\min}, \lambda_{max}]}$ is the set of functions $s : [\lambda_{\min}, \lambda_{\max}] \rightarrow \mathbb{R}$, that includes the set of physical spectra, plus many others. In computer vision, as well as in biological vision, both the $\varphi_p$ and the $s$ are nonnegative; the generalization pays off though and not nonnegative spectra are called virtual spectra. Given that in the darkest of nights we do see a colour that we call black, that we assimilate to the black seen under photopic conditions, we call the origin of $\mathbb{R}^P$ the black (colour) point, and the spectra in the kernel of the linear transformation $s \mapsto c$ (i.e. the spectra that are mapped to black), metameric black (spectra). It turns out then that two spectra $s_1$ and $s_2$ are metameric, i.e. they are mapped to the same colour $c$, if and only if their difference is a metameric black.
It is not unreasonable to discretize the interval \([\lambda_{\min}, \lambda_{\max}]\) of visible wavelengths, as the vector \([\lambda_1 = \lambda_{\min}, \lambda_2, \ldots, \lambda_N = \lambda_{\max}]\) \(\in \mathbb{R}^N\) of \(N\) (assume \(N\) to be much larger than \(P\)) representative wavelengths \(\lambda_i\), reducing the dimension of the domain of the linear transformation and making the linear transformation representable as the product by a matrix \(M\), as in \(c = Ms\), where the integrals become dot products of vectors in \(\mathbb{R}^N\). Thus, if the linear transformation in question is full rank (which we assume), the kernel \(K\) of \(M\) has dimension \(K = N - P\) and it may be of interest to find a basis for it.

Clearly, the detector spectral responses \(\varphi_p\) determine the kernel of the transformation and you may ask the reverse question: given certain desirable properties of the kernel of the transformation, what restrictions should the \(\varphi_p\) meet? What spectra can we afford to confuse? What spectra should we be able to differentiate? Since two spectra are metameric if and only if their difference is a metameric black, you can group the spectra into classes of metameric spectra, each class (or coset) of dimension \(N - P\). Thus, you think in terms of cosets of metameric spectra when designing a vision system in such a fashion.

Two theoretical, archetypical types of photodetector are detectors that respond to a unique wavelength on the one hand, and detectors that respond equally well to any wavelength in a given interval of wavelengths. In the first case, the responses of the photodetectors are of the form \(c_p = \int s(\lambda)d\lambda - \lambda_p\) (a Riemann-Stieltjes integral where \(u\) is Heaviside’s step function), where a pointwise (instead of aperture) sampling of the spectrum \(s\); the Nyquist criterion could be used to find out how many such samples would be needed so that no loss of information would occur. The kernel \(K\) consists then of the spectra with zeros at \({\lambda_i : i = 1, \ldots, p}\}. Two spectra are then metameric if and only if they coincide at wavelengths \(\lambda_i\), regardless of the behavior at the open intervals in between the \(\lambda_i\)’s. In the second case, for detector spectral responses that are rectangular, \(c_p = \int_{\lambda_p}^{\lambda_{p+1}} s(\lambda)d\lambda\) and the kernel \(K\) consists of spectra that have zero area in the intervals; correspondingly, two spectra are metameric if and only if they have the same area (energy) within each of the intervals \((\lambda_p, \lambda_{p+1})\), \(p \in \{1, \ldots, P\}\). These intervals may overlap. A physical detector response curve typically is bell shaped and resembles more the second case. In any case, a linear combination of colour responses \(\sum \alpha_p c_p\) corresponds to a hypothetical detector with spectral response given by the corresponding linear combination \(\sum \alpha_p \varphi_p(\lambda)\) of the spectral responses of the original photodetectors. The typical bell shaped curve results from a probabilistic quantum photon catch and is not really a linear combination of elementary detectors. In this way, the luminance channel \(L + M + S\) of our trichromatic vision may be thought as the response of a detector with spectral response \(S(\lambda) + M(\lambda) + L(\lambda)\). In general: \(\int [\sum_{p \in Q} \varphi_p(\lambda)]s(\lambda) = \sum_{p \in Q} [\int d_p(\lambda)s(\lambda)]\); the right-hand side is the sum of the responses while in the left hand side you find the spectral response \(\sum_{p \in Q} \varphi_p(\lambda)\) of a hypothetical achromatic detector.

Assume that a flat (constant) spectrum gives equal readings of the photodetectors: \(c_i = c_j\), \(i, j \in \{1, \ldots, p\}\), i.e. that the area \(\int \varphi_p(\lambda)\) of the spectral responses of the photodetectors is normalized: it is the same, and equal to 1. Call one such a colour \(c\) an achromatic colour, and one such spectrum \(s\), a flat achromatic spectrum. Call any spectrum that gives rise to an achromatic colour, an achromatic spectrum. \(\sum_{p=1}^{P} \varphi_p(\lambda)\) is the hypothetical achromatic detector. The set of the achromatic colours is the achromatic line of \(\mathbb{R}^P\).

Given a spectrum, the flat spectrum of level equal to the minimum (as you consider all wavelengths) is called the flat achromatic component of the given spectrum. The achromatic spectrum that when subtracted from a given spectrum gives the physical (nonnegative) spectrum with minimal maximum is called the achromatic component of the given spectrum. Likewise, given a colour \(c\), its achromatic component is given by the constant colour of value equal to the minimum of \(c\). The inverse under the linear transformation \(s \mapsto c\) of the achromatic line of \(\mathbb{R}^P\) is also a subspace of \(\mathbb{R}^N\), called the set of the metameric gray and it contains the flat spectra; the spectra in each corresponding coset are said to be equiluminant and are linearly ordered by luminance.

The projection of a colour \(c\) onto the achromatic line of \(\mathbb{R}^P\) is given by the average \(\bar{c} := \frac{1}{P} \sum c_p\) of the components of \(c\); however, \(c - [\bar{c}, \ldots, \bar{c}]\) may have negative components and be a virtual colour.

Because the \(\varphi_p(\lambda)\) are nonnegative, the spectra in \(K\), with the exception of the zero spectrum, must be virtual: they must have both positive and negative components. One way to visualize a spectrum in \(K\), i.e. a metameric black, and in fact any virtual spectrum, is to add to it a metameric gray spectrum that yields a
nonnegative spectrum of minimal maximum value, and obtain the corresponding colour \( c \); this corresponds to a desaturation of sorts of the metameric black.

Metameric black spectra have zero crossings at wavelengths that correspond to the intersection points of metameric spectra. Such nodes have been considered, for the case of spectral reflectances, in.\(^1\)

### 1.2 3-metamerism

In our case, under photopic conditions, we have three types of cone photopigment \( S(\lambda) \), \( M(\lambda) \) and \( S(\lambda) \). Each of the corresponding linear, scalar transformations has a kernel, and the intersection of the three kernels \( K_L \cap K_M \cap K_S \) is the kernel \( K \) of the linear, vector transformation. A protanope (a type of colour blindness) has no cones of type \( L \) and can confuse more colours as the intersection of the kernels \( K_M \cap K_S \) is bigger than \( K \).

From the space \( L, M, S \), by means of a linear transformation \( \mathbb{R}^3 \rightarrow \mathbb{R}^3 \), three other channels \( L + M + S \), \( (L + S) - M \) and \( (L + M) - S \) are derived that optimally extract visual information, given the statistics of natural scenes; see for example.\(^2\) A question is then, what does \( LMS \) metamerism implies for these new three channels and, in particular, what does colour blindness implies for these channels.

With Cohen’s method,\(^3\) based on CIE data, you obtain a matrix of base vectors for \( K \), via \([M^T(MM^T)^{-1}]M\); instead, we derive a basis for \( K \) consisting of virtual spectra with localized support.

### 1.3 4-metamerism

Given four photodectectors with spectral sensitivity curves \( w(\lambda) \), \( x(\lambda) \), \( y(\lambda) \), and \( z(\lambda) \), and the spectrum \( s(\lambda) \) of a light beam that falls on the surface of each, the corresponding responses are given by \( c_w = \int s(\lambda)w(\lambda), c_x = \int s(\lambda)x(\lambda), c_y = \int s(\lambda)y(\lambda) \) and \( c_z = \int s(\lambda)z(\lambda) \). The integrals measure the area below the spectrum curve (i.e. the radiant energy) as “seen through” each of the sensitivity curves. In a sense, the sensitivity curves *aperture sample* the spectrum. In such a tetrachromatic a vision system\(^*\), two spectra giving rise to the same responses \( c_w, c_x, c_y \) and \( c_z \) will be undistinguishable by the photodetectors and will be said to be metameric. The point \( c = [c_w, c_x, c_y, c_z] \in \mathbb{R}^4 \) will be called a *colour point*; here, \( \mathbb{R} \) denotes the set of the real numbers.

Thus, in going from \( s(\lambda) \) to \( c = [c_w, c_x, c_y, c_z] \), you take four *aperture samples* of \( s \) and metamerism results when the photoreceptors are unable to resolve two spectra. This is unavoidable if you consider that a set of four photoreceptors linearly\(^\dagger\) maps the graph curve of each spectrum function \( s : [\lambda_{\text{min}}, \lambda_{\text{max}}] \rightarrow [0, \infty) \) into a point on the “16-tant”\(^\ddagger\) \( \mathbb{R}^{4+} \) that we denote also as \([+, +, +, +] := \{(t_1, t_2, t_3, t_4) : t_i \geq 0 \} \) of \( \mathbb{R}^4 \).

![Figure 1](image-url)  
*Figure 1. The rectangle is meant to be the domain space of a linear transformation; the point at the center is the origin of the space. The light green line is meant to be the kernel of the transformation, the green lines are cosets and the red line is the orthogonal complement of the kernel.*

---

\(^*\)In our case, we are old-world, frugivore, trichromatic primates and most of us have exactly three types of photopigment in the main receptor layer of our retinae, but many animals are tetrachromatic.

\(^\dagger\)The fact that the irradiance of the light beam is nonlinearly contracted to a bounded luminance is being overlooked here; nevertheless, regarding the hue, things are pretty much linear.

\(^\ddagger\)Analogously to the quadrants of the plane and the octants of 3-space.
In order to use the machinery of linear spaces with the transformation $s \mapsto c$, we must allow both spectra and color points to take on negative values as well. The resulting sets of spectra $R^{[\lambda_{\min}, \lambda_{\max}]}$ and of colours $R^4$ are now linear spaces and will be called sets of virtual spectra and of virtual colours, respectively. When a virtual spectrum is nonnegative we say that is is realizable and, likewise, when the components of a colour are nonnegative, we say that it is a realizable colour.

The kernel of the transformation $R^{[\lambda_{\min}, \lambda_{\max}]} \rightarrow R^4$ that maps $s \mapsto c$, is the set of spectra that are mapped to the colour point $[0, 0, 0, 0]$, called here black. The spectra in each corresponding coset of spectra are mapped to the same (colour) point on $R^4$; this is the main idea behind a mathematization of the phenomenon of metamerism. See Figure 1.

In our digital, technical world, magnitudes are made discrete; a camera aperture-samples the spectrum of the light at each of many small spatial regions or pixels. In addition to this, we assume that such sampling is done over an already sampled spectrum, sampled at a much more finer scale, e.g. every 10 nm, or so. Thus, the interval set of wavelengths $[\lambda_{\min}, \lambda_{\max}] \subset R$ is converted to a finite sequence $\{\lambda_1 = \lambda_{\min}, \lambda_2, \lambda_3, \ldots \lambda_N = \lambda_{\max}\}$ of wavelengths, and the transformation from spectra to colors is now of the form $R^N \rightarrow R^4$, considerably simpler and yet a good approximate model. Integrals become dot products, i.e. $c_w = w_s = \sum w_is_i$, $c_x = x_s = \sum x_is_i$, $c_y = y_s = \sum y_is_i$ and $c_z = z_s = \sum z_is_i$.

Most mammals are dichromatic\textsuperscript{8} and it has been argued that the L photopigment evolved in old-world monkeys as it resulted advantageous in the appraisal of the ripeness of fruits when seen from the distance; or, since male dichromacy is common in such primates, that it evolved providing females with health cues regarding potential mates. In biological vision, tetrachromacy is found in fish, birds, reptiles and in many invertebrates; the mantis shrimp is 12-chromatic and sees in the range from 300 nm to 700 nm. In satellite imagery, the bands may be many but you may restrict to R, G, B plus either NIR or UV. In both cases, the bands include some amount of overlap but are mostly disjoint; unlike the case of a recently developed detector for photography that, in addition to the usual R, G and B pixels of a Bayer, array sensor, it includes unfiltered (other than by the glass of the lens of the camera and the package of the sensor) "panchromatic" pixels.

Spectral lights are perceived as more saturated than more wide-band lights; thus, the spectral yellow might appear a bit more saturated, than the yellow that results from the mixture of green and red. In what we call hue metamerism, luminance differences of the light beams can be considered immaterial, as well as the saturation, up to a degree. Two colors may have the same hue but different luminance and different chromatic saturation; correspondingly, a relaxed type of metamerism may be also exploited in computer vision systems; differences in luminance may be due to differences in illumination intensity and differences in saturation may be due to atmospheric conditions but not to different spectral reflectances of surfaces. In this line, it is useful to consider a hypercube of photodetector responses and identify sets of constant hue or chromatic triangles, in it,\textsuperscript{5, 6}

2. TETRACHROMATIC METAMERISM

Let $w, x, y$ and $z$ be four linearly independent, $N$-vectors of samples of the spectral responses, at a set of wavelengths $\lambda_1, \ldots \lambda_N$, of four photodetectors; also, let the (corresponding samples of the) light spectrum be given by $s$. Denote the "colour" response of the photodetector set by $c = [c_w, c_x, c_y, c_z]$. Assume then that the response to a light beam falling on four such, nearly placed, photodetectors is given by

$$c^T = \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix}$$

or

\textsuperscript{5}Most mammals have cones $S$ (short wavelengths) and $M$ (medium wavelengths) but no cone $L$ (large wavelengths). See.\textsuperscript{4}
This provides a linear transformation $\mathbb{R}^N \rightarrow \mathbb{R}^4$, $s \mapsto r$, that reduces the dimensionality from $N$ to 4. $M$ has full rank and its entries are nonnegative and, typically, positive. Thus, for a nonnegative $s$, $c$ is nonnegative: $c \in \mathbb{R}^{4+}$. The kernel $K$ of this transformation is given by the set of vectors $k$ for which

$$\pmatrix{w \\ x \\ y \\ z} = \frac{1}{s_1} \pmatrix{0 \\ 0 \\ 0 \\ 0}$$

Thus, $K$, the set of the metameric blacks, is the space of vectors orthogonal to (each element of) the subspace $L := \text{span}\{w, x, y, z\} = \{aM : a \in \mathbb{R}^4\}$, which is isomorphic to $\mathbb{R}^4$. Also, any two spectra $s$ and $2s$ such that $s - 2s \in L^\perp = K$, produce the same colour response $c = [c_w, c_x, c_y, c_z]$. $L$ has dimension 4 and $L^\perp$ has dimension $N - 4$; also, $MM^T : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ is invertible. $K$ contains "spectra" (we might call them virtual spectra) that are neither nonnegative nor nonpositive. The cosets $s + K := \{s + k : s \in \mathbb{R}^N, k \in K\}$ provide a partition of $\mathbb{R}^N$. In a decomposition $s = f + k$, $f \in L$, $k \in K$, which is unique, $f$ is called a fundamental metamer and $k$ is called a metameric black. The spectra in the coset $f + K$ are said to be metameric and are mapped by $M$ to the same colour point $c \in \mathbb{R}^4$; only the nonnegative spectra in such coset are realizable, the remaining are merely virtual.

2.1 A basis for $K$

Calling the colour point $[0, 0, 0, 0]$ black, then $K$ is the set of spectra that "evolve" the colour black; call them metameric blacks. Since the components of $M$ are nonnegative, the only nonnegative spectrum that is a metameric black is the 0 spectrum; all other metameric black spectra include both positive and negative components.

\footnote{The spectra in $\mathbb{R}^N$ that are nonnegative are those in the wedge or “2$^N$-tant” $[+, +, ... +] := \mathbb{R}^{N+}$}
Cohen’s method,\(^3\) based on CIE data, consists of finding \(f\) as \(f = [M^T(MM^T)^{-1}M]s\) and then writing \(k = s - f\).

We derive a basis for \(K\) of narrow-band spectra in a 4-step process where 4 triangular, sparse matrices of row vectors of local support are derived. In the first matrix \(1A\) you have a basis for the orthogonal complement of \(\text{span}\{w\}\), in the second one \(2A\), a basis for the orthogonal complement of \(\text{span}\{w,x\}\), then, in \(3A\), a basis for \(\text{span}\{w,x,y\}\) and finally, in \(4A\), a basis for \(\text{span}\{w,x,y,z\}\). We assume that the components of \(w, x, y\) and \(z\) are positive so that the matrix \(1A\) below is computable and also that each of the matrices \(2A, 3A\) and \(4A\), as defined below, are computable.

Let \(1A\) be the \(N \times (N - 1)\) matrix with \(i^{\text{th}}\) row of the form \([0, ..., 0, 1, -w_i/w_{i+1}, 0, ..., 0]\); thus, \(1M\) has a diagonal of 1’s. Clearly, each row of \(1A\) is orthogonal to \(w\) and, since linearly independent, they provide a basis for \(\text{span}\{w\}\).

Let each row of \(3A\) result from linearly combining each pair of consecutive rows of from \(1A\). In this way, each row is still orthogonal \(\text{span}\{w\}\) and, by using appropriate weights in the combination, you can make it also orthogonal to \(\text{span}\{x\}\). In fact, let the \(i^{\text{th}}\) row of the \(N \times (N - 2)\) matrix \(2A\) be given by

\[
\begin{bmatrix}
0, ..., 0, 1, m_{i,i+1} + \beta m_{i+1,i+1}, m_{i,i+2} \\
+ \beta m_{i+1,i+2}, m_{i+1,i+3}, 0, ..., 0
\end{bmatrix}
\]

where the \(m_i\)’s are the components of \(1A\), and

\[
\beta_i = -\frac{m_{i+1,i+1}x_{i+1} + m_{i,i+2}x_{i+2}}{m_{i+1,i+1}x_{i+1} + m_{i,i+2}x_{i+2}};
\]

again, the diagonal of \(2A\) is a diagonal of 1’s. Likewise, by making sure a certain linear combination of each two consecutive rows in \(2A\) is orthogonal to \(y\), you get the \(N \times (N - 3)\)-matrix \(3A\) with \(i^{\text{th}}\) row of the form

\[
\begin{bmatrix}
0, ..., 0, 1, m_{i,i+1} + \beta m_{i+1,i+1}, m_{i,i+2} \\
+ \beta m_{i+1,i+2}, m_{i+1,i+3}, 0, ..., 0
\end{bmatrix}
\]

where the \(m_i\)’s are now the components of \(2A\) and

\[
\beta_i = \frac{-y_{i+1}m_{i,i+1} + y_{i+2}m_{i,i+2}}{y_{i+1}m_{i,i+1} + y_{i+2}m_{i,i+2}}.
\]

The diagonal of \(3A\) is a diagonal of ones. Finally, a linear combination of each two consecutive rows in \(3A\) that is orthogonal to \(z\), provides the \(N \times (N - 4)\)-matrix \(4A\) with \(i^{\text{th}}\) row

\[
\begin{bmatrix}
0, ..., 0, 1, m_{i,i+1} + \beta m_{i+1,i+1}, m_{i,i+2} \\
+ \beta m_{i+1,i+2}, m_{i+1,i+3}, 0, ..., 0
\end{bmatrix}
\]

where the \(m_i\)’s are now the components of \(3A\) and

\[
\beta_i = \frac{-x_{i+1}z_{i+1} + z_{i+2}m_{i,i+1} + z_{i+3}m_{i,i+2} + z_{i+4}m_{i,i+3}}{x_{i+1}z_{i+1} + z_{i+2}m_{i,i+1} + z_{i+3}m_{i,i+2} + z_{i+4}m_{i,i+3}}.
\]

See Fig. 6. Let the rows of the \(N \times (N - 4)\)-matrix \(B := 4A\) denote this resulting base for the kernel \(K\) of \(M\).

Note that the rows in each of these matrices are linearly independent due to their localized support. The rows in \(1A\) have support of length 2, those in \(2A\) have support of length 3, those in \(3A\) have support of length 4 and those in the basis \(B\) of \(L^\perp = K\) have support of length 5. The existence of \(\beta\) in each case is not so surprising due to this localization; that is, for example, it is not too difficult to find constants \(\alpha_1\) and \(\beta_1\) so that \((\alpha[1, -w_1/w_2, 0] + \beta[0, 1, -w_3/w_3])\] is \(\alpha x_1 + (\alpha w_1 + \beta x_2 - \beta w_2) x_3 = 0\); letting \(\alpha = 1\), you only need \(x_1 - w_1 x_2/w_2 + \beta(x_2 - w_2 x_3/w_3) = 0\), i.e., \(x_2 \neq w_2/w_3 x_3\) or \(x_3 \neq w_3/w_2 x_2\). In fact, when computing \(2A\), each \(\beta_i\) can be also expressed as \(\beta_i = -\frac{x_{i+1} - x_{i+2}}{x_{i+1} - x_{i+2}}\). Similar formulas of \(\beta\) can be derived for the remaining cases of \(3A\) and \(4A\).

The addition of a scaled element of the so-obtained basis for \(K\) is the addition of a metameric black that alters the spectrum in a very narrow region of it producing a new metamer. Large peaks in a base element of \(K\) may indicate indicating insensitivity to certain wavelengths.
2.2 Sets of Metameric Spectra

The set of spectra $\mathbf{R}^N$ is partitioned into cosets of the form $\mathbf{K} + \mathbf{s}$, $\mathbf{s} \in \mathbf{R}^N$. To each colour point $\mathbf{c} = [w, x, y, z]^T$ in the hypercube, there corresponds the coset $\mathbf{S}_c$ of dimension $N-4$, of spectra (not necessarily realizable as physical spectra) that are mapped by the matrix

$$
\mathbf{M} = 
\begin{bmatrix}
w \\
x \\
y \\
z
\end{bmatrix}
$$

to such colour point. To find $\mathbf{S}_c$, you find a spectrum vector $\mathbf{s}$ for which $\mathbf{M}\mathbf{s} = \mathbf{c}$ and then write $\mathbf{S}_c = \mathbf{K} + \mathbf{s}$. To find one such $\mathbf{s}$, choose a 4x4 matrix $\mathbf{N}$ given by four columns of $\mathbf{M}$, say the $i^{th}, j^{th}, k^{th}, l^{th}$ columns:

$$
\mathbf{N} = 
\begin{bmatrix}
w_i & w_j & w_k & w_l \\
x_i & x_j & x_k & x_l \\
y_i & y_j & y_k & y_l \\
z_i & z_j & z_k & z_l
\end{bmatrix}
$$

We assume that $\mathbf{N}$ is invertible; in fact we choose one such matrix $\mathbf{N}$ having highest absolute determinant so that the computation of its inverse is more accurate. Regarding the possible values of the absolute value of the determinant, there are $\binom{N}{4} = 1.282.975$ choices of $i, j, k, l$ to consider. Once one such matrix is chosen, put $\mathbf{t} = \mathbf{N}^{-1}\mathbf{c}$ and

$$
\mathbf{s} = [0, ...t_i, 0...t_j, 0, ...t_k, 0, ...t_l, 0, ...]^T
$$

$s$ is not necessarily in $\mathbf{L}$, i.e. it is not necessarily a fundamental metamer; also, $\mathbf{t}$ may have negative components; in such case, a spectrum $\mathbf{s}$ that is nonzero only at positions $i, j, k, l$ and produces colour $\mathbf{c}$, is not physically realizable. Only the nonnegative spectral photoreceptor vectors in $\mathbf{K} + \mathbf{s}$ are of realizable. It is possible that a realizable colour not be the image of a realizable spectrum.

If you are designing a tetrachromatic imaging system and do not want spectra $\mathbf{1}s$ and $\mathbf{2}s$ to be metameric, at least one of the pairs $\mathbf{w}_1\mathbf{s}$ and $\mathbf{w}_2\mathbf{s}$, or $\mathbf{x}_1\mathbf{s}$ and $\mathbf{x}_2\mathbf{s}$, or $\mathbf{y}_1\mathbf{s}$ and $\mathbf{y}_2\mathbf{s}$, or $\mathbf{z}_1\mathbf{s}$ and $\mathbf{z}_2\mathbf{s}$ should be different, in particular, $\mathbf{1}s - \mathbf{2}s$ should be in $\mathbf{L}$ and must not be in $\mathbf{K}$; i.e. $\mathbf{1}s - \mathbf{2}s \in \mathbf{L} - \mathbf{K}$.

2.3 A case example: RGB+Panchromatic

Besides satellites, a source of tetrachromatic images is computational photography. TrueSense Imaging inc. markets a digital image sensor that, in addition to R, G, and B pixels of a Bayer pattern, it includes as well panchromatic pixels in a pattern as shown in Fig. 3. The proportions are 1/4 of green pixels, 1/8 of red pixels, 1/8 of blue pixels and 1/2 panchromatic pixels. Even though the photosensitive transducers respond well into the UV, the microlens blocks wavelengths below 350 nm. The sensor responds in the infrared but the response is negligible above 1050 nm.

The pattern of the color filter array is

For our purposes, we do not need interpolate the data in the pattern array that give rise to the image shown at the top in Figure 4; instead, we downsample each 4 $\times$ 4 pixel block to a tetrachromatic pixel, by averaging the pixels in each band in the block. Thus, even though the original image is 1152 $\times$ 2044, the image we work with is only 287 $\times$ 510 pixels. Also, the bands we use are $w = P$, $x = R$, $y = G$ and $z = B$.

The data provided by TrueSense of the quantum efficiency of each sensor type, at each 10 nm from 350 to 1100 nm, provides 76 data per band. The basis elements in are shown in Figures 6, at the bottom.

The submatrix with largest determinant is given by

\(^{1}\)Pixels that are covered by the microlens but that otherwise do not receive filtered light.
Figure 3. Pattern in the array of the sensor Truesense Imaging KAI-01150: P B P G; B P G P; P G P R; G P R P.

Figure 4. Outdoors, 16-bit, RGBP image of a Macbeth chart; courtesy of Amy Enge. Below, RGB visualization without corrections.

Figure 5. Quantum efficiencies corresponding to Truesense sensors P, R, G and B.

\[
N = \begin{bmatrix}
p_5 & p_{13} & p_{20} & p_{29} \\
r_5 & r_{13} & r_{20} & r_{29} \\
g_5 & g_{13} & g_{20} & g_{29} \\
b_5 & b_{13} & b_{20} & b_{29}
\end{bmatrix}
\]

= \begin{bmatrix}
0.3544 & 0.5390 & 0.5110 & 0.3769 \\
0.0378 & 0.0337 & 0.0354 & 0.3688 \\
0.0261 & 0.1453 & 0.4382 & 0.0262 \\
0.1233 & 0.4566 & 0.0628 & 0.0013
\end{bmatrix}

and has determinant 0.0136 and inverse given by
\[ N^{-1} = \begin{bmatrix} -4.8879 & -5.0124 & -4.1153 & 5.1435 \\ 1.3590 & 1.0476 & 3.4252 & -1.4138 \\ -0.3459 & 2.2207 & -0.9022 & 0.1873 \\ 3.1216 & 0.2047 & 0.1949 & -0.4158 \end{bmatrix} \]

For example, corresponding to colour \( c = [0.25, 0.25, 0.25, 0.75]^T \) you get
\[ t = [0.3538, 0.3976, 0.3837, 0.5685]^T \] and
\[ s = [0...0.3538, 0...0.3976, 0...0.3837, 0...0.5685, ...] \], with nonzero values at coordinates 5, 13, 20 and 29.

2.4 Program code

In the MATLAB code below, vectors R, G, B, P are the quantum efficiencies. The base for the orthogonal complement is in AL4. Note: here, the matrix \( M \) used is \( R = [r; g; b; p]^T \)

\[
\begin{align*}
\text{AL1} & \text{= zeros(76,75);} \\
& \text{for ii=1:75} \\
& \quad \text{AL1(ii,ii)= 1;} \\
& \quad \text{AL1(ii,ii+1)= -R(ii)/R(ii+1);} \\
& \text{end} \\
& \text{figure; plot(L,AL1)} \\
\text{AL2} & \text{= zeros(76,74);} \\
& \text{for ii=1:74} \\
& \quad \text{AL2(ii,ii)= 1;} \\
& \quad \text{AL2(ii,ii+1)= -(G(ii)+G(ii+1)*AL1(ii,ii+1)/AL1(ii,ii)+G(ii+2))/...} \\
& \quad \quad \text{AL2(ii,ii+2)= AL2(ii,ii+1)/AL2(ii,ii)+BETA*AL2(ii+1,ii+1);} \\
& \text{end} \\
& \text{figure; plot(L,AL2)} \\
\text{AL3} & \text{= zeros(76,73);} \\
& \text{for ii=1:73} \\
& \quad \text{AL3(ii,ii)= 1;} \\
& \quad \text{AL3(ii,ii+1)= -(B(ii)+B(ii+1)*AL2(ii,ii+1)/AL2(ii,ii)+B(ii+2)*AL2(ii,ii+2)/AL2(ii,ii)+B(ii+3)*AL2(ii,ii+3)/AL2(ii,ii))/...} \\
& \quad \quad \text{AL3(ii,ii+2)= AL3(ii,ii+1)/AL3(ii,ii)+BETA*AL3(ii+1,ii+1);} \\
& \text{end} \\
& \text{figure; plot(L,AL3)} \\
\text{AL4} & \text{= zeros(76,72);} \\
& \text{for ii=1:72} \\
& \quad \text{AL4(ii,ii)= 1;} \\
& \quad \text{AL4(ii,ii+1)= -(P(ii)+P(ii+1)*AL3(ii,ii+1)/AL3(ii,ii)+P(ii+2)*AL3(ii,ii+2)/AL3(ii,ii)+P(ii+3)*AL3(ii,ii+3)/AL3(ii,ii)+P(ii+4)*AL3(ii,ii+4)/AL3(ii,ii))/...} \\
& \quad \quad \text{AL4(ii,ii+2)= AL4(ii,ii+1)/AL4(ii,ii)+BETA*AL4(ii+1,ii+1);} \\
& \text{end} \\
& \text{figure; plot(L,AL4)} \\
\text{AL5} & \text{= zeros(76,73);} \\
& \text{for ii=1:73} \\
& \quad \text{AL5(ii,ii)= 1;} \\
& \quad \text{AL5(ii,ii+1)= -(Q(ii)+Q(ii+1)*AL4(ii,ii+1)/AL4(ii,ii)+Q(ii+2)*AL4(ii,ii+2)/AL4(ii,ii)+Q(ii+3)*AL4(ii,ii+3)/AL4(ii,ii)+Q(ii+4)*AL4(ii,ii+4)/AL4(ii,ii))/...} \\
& \quad \quad \text{AL5(ii,ii+2)= AL5(ii,ii+1)/AL5(ii,ii)+BETA*AL5(ii+1,ii+1);} \\
& \text{end} \\
& \text{figure; plot(L,AL5)} \\
\text{AL6} & \text{= zeros(76,74);} \\
& \text{for ii=1:74} \\
& \quad \text{AL6(ii,ii)= 1;} \\
& \quad \text{AL6(ii,ii+1)= -(R(ii)+R(ii+1)*AL5(ii,ii+1)/AL5(ii,ii)+R(ii+2)*AL5(ii,ii+2)/AL5(ii,ii)+R(ii+3)*AL5(ii,ii+3)/AL5(ii,ii)+R(ii+4)*AL5(ii,ii+4)/AL5(ii,ii))/...} \\
& \quad \quad \text{AL6(ii,ii+2)= AL6(ii,ii+1)/AL6(ii,ii)+BETA*AL6(ii+1,ii+1);} \\
& \text{end} \\
& \text{figure; plot(L,AL6)} \\
\text{AL7} & \text{= zeros(76,75);} \\
& \text{for ii=1:75} \\
& \quad \text{AL7(ii,ii)= 1;} \\
& \quad \text{AL7(ii,ii+1)= -(S(ii)+S(ii+1)*AL6(ii,ii+1)/AL6(ii,ii)+S(ii+2)*AL6(ii,ii+2)/AL6(ii,ii)+S(ii+3)*AL6(ii,ii+3)/AL6(ii,ii)+S(ii+4)*AL6(ii,ii+4)/AL6(ii,ii))/...} \\
& \quad \quad \text{AL7(ii,ii+2)= AL7(ii,ii+1)/AL7(ii,ii)+BETA*AL7(ii+1,ii+1);} \\
& \text{end} \\
& \text{figure; plot(L,AL7)} \\
\text{AL8} & \text{= zeros(76,76);} \\
& \text{for ii=1:76} \\
& \quad \text{AL8(ii,ii)= 1;} \\
& \quad \text{AL8(ii,ii+1)= -(T(ii)+T(ii+1)*AL7(ii,ii+1)/AL7(ii,ii)+T(ii+2)*AL7(ii,ii+2)/AL7(ii,ii)+T(ii+3)*AL7(ii,ii+3)/AL7(ii,ii)+T(ii+4)*AL7(ii,ii+4)/AL7(ii,ii))/...} \\
& \quad \quad \text{AL8(ii,ii+2)= AL8(ii,ii+1)/AL8(ii,ii)+BETA*AL8(ii+1,ii+1);} \\
& \text{end} \\
& \text{figure; plot(L,AL8)} \\
\text{DETMAX} & \text{= 0;} \\
& \text{for ii = 1:73} \\
& \quad \text{DETMAX= max(DETMAX,det(AL4x4));} \\
& \text{end} \\
\text{figure; plot(L,AL8)} \\
\text{AL= transpose(AL);} \\
& \text{iii=1} \\
\text{for jj= 1:75} \\
& \quad \text{plot(L,AL(jj,iii))} \\
\text{end} \\
\text{figure; plot(L,AL8)} \\
\text{end}
\]

In the MATLAB code below, get 4x4 submatrix with largest determinant, then get \( t \) (METAMER) for, e.g.
\[ c = [3 \ 3 \ 7 \ 7]. \]
3. OTHER ASPECTS OF METAMERISM

Incidentally, the so-called optimal Schrödinger’s colours, result from reflectance spectra \( r(\lambda) \) with either \( r(\lambda) = 1 \) (maximal) or \( r(\lambda) = 0 \) (minimal), having at most two transitions in the visible spectra. Therefore, the spectra can be classified as low-pass with only one transition downwards, high-pass with only one transition upwards, band-pass, or Mittelpigmente with an up transition followed by a down transition, and or stop-band = Mittelfehlpigmente with a down transition followed by an up transition; four possibilities, and also four unique colours. Fur unique colours that may be related to the equality in the responses of the channels L+M, L+S, L+M+S, M, S, L... These are spectra that are mapped to the boundary of the object-colour solid, i.e. the portion of tristimulus space mapped to by the surface of an object.

Perhaps it is not as easy to come up with arbitrary biological cone responses, they are usually bell shaped. Biologica vision systems, such as mammalian ones, transform in the retina the receptor responses L, M and S into the channels \( L + M + S \), achromatic, and \( \pm [(L + M) - S] \) and \( \pm [(L + S) - M] \) (cromatic). Perhaps the maxima and minima of the two chromatic channels give rise to the four uniques. Perhaps for dichromatic animals the corresponding channels are \( L + M \), \( M - S \) and \( L - M \).

Even thogh the kernel of the separate linear transforms \( s \rightarrow L, s \rightarrow M, \) and \( s \rightarrow S, \) as their intersection, give the kernel of \( M, \) it is more revealing to study the kernel of \( L + M + S, [L + M] - S \) and \( [L + S] - M, \) as they are more related to the luminance and chrominance visual functions, and allow the study of a separate-from-luminance chromatic metamerism.

Consider the transformation

\[
\mathbf{f} := \begin{bmatrix}
\frac{1}{2} & -1 & \frac{1}{2} \\
\frac{1}{2} & 1 & -1 \\
\frac{1}{3} & 1 & 1
\end{bmatrix}
\begin{bmatrix}
L \\
M \\
S
\end{bmatrix}
\]

Let

\[
\mathbf{T} = \begin{bmatrix}
\frac{1}{2} & -1 & \frac{1}{2} \\
\frac{1}{2} & 1 & -1 \\
\frac{1}{3} & 1 & 1
\end{bmatrix}
\]

Thus

\[
\mathbf{f} = \mathbf{Tm}
\]
Two spectra with same $L + M + S$ are said to be luminance metameric, or to have the same luminance. Likewise, two spectra with same $0.5(L+S) - M$ and same $0.5(L+M) - S$ are said to have the same chrominance or to be chrominance metameric. This provides a description of the kernel $K$ as the intersection of the kernels of the 3 scalar linear transformations in $F$. Because $T$ is invertible, the kernel of $F$ is also $K$.

Chromatic metamerism has two aspects, Violet-Green and Ylellow-Blue metamers.

For a tetrachromatic system with 6 unique hues, on a 2-sphere of hues, like the vertices of an octahedron, there may be 8 families of trinary hues (the faces of the octahedron), just as for a trichromatic vision system (perhaps also a dichromatic system $S, M$) on a circle of hues there are 4 uniques and four families of binary hues.

### 3.1 Chromatic 3-meterism

For us humans, two colours may have the same saturation, the same luminance or the same hue. When studying the colour vision of a tetrachromatic animal, it may be interesting to design an experiment to find out if the animal can distinguish hue while disregarding luminance and saturation. In this sense, we call two spectra hue-metameric if they give rise to colour points on the same chromatic triangle; see Figure 9. In the colour hypercube you also have the achromatic segment and instead of a chromatic hexagon, you have a chromatic icositetrahedron; the triangles having as base the achromatic segment and as opposing a vertex a point in the chromatic icositetrahedron are called chromatic triangles and all colours in each such triangle are said to have the same hue.

If, in the trichromatic case, we take for the luminance $\lambda = L + M + S$ and for the chromatic components $Yb = (L + M) - S$ and $Vg = (L + S) - M$ then, on the one hand, we have a linear invertible transformation $L, M, S \rightarrow \lambda, Yb, Vg$.

\[
\begin{bmatrix}
Yb \\
Vg \\
\lambda
\end{bmatrix} =
\begin{bmatrix}
1 & -1 & 1 \\
1 & 1 & -1 \\
1 & 1 & 1
\end{bmatrix}
\begin{bmatrix}
L \\
M \\
S
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & -1 & 1 \\
1 & 1 & -1 \\
1 & 1 & 1
\end{bmatrix}^{-1} =
\begin{bmatrix}
1/2 & 1/2 & 0 \\
-1/2 & 0 & 1/2 \\
0 & -1/2 & 1/2
\end{bmatrix}
\]

The kernel of $spectra \rightarrow LMS$ is also the kernel of $spectra \rightarrow \lambda, Yb, Vg$. The kernel are those spectra that give a zero dot product with each of the three $L, M, S$ or $\lambda, Vg, Yb$ and therefore if we consider the second case, those orthogonal to $\lambda$ give luminance me tamers and those orthogonal both to $Yb$ and $Vg$ are the chromatic me tamers, including both hue and saturation. The spectra that are both luminance metameric and chromatic metameric are then the (fully) metameric spectra.

Wyszecki and others have wondered about the so called nodes or crossing points of metameric spectra. They correspond to the zeros of the virture spectra int the kernel of the lms transformation. Such elements have a null dot product with each of $L$, $M$ and $S$ and it is therefor that, unless it is the zero spectrum, they include both positive and negative components. Since the Y CIE component measures luminance, a plot in the XZ CIE plane could be illuminating regarding chrominance metamerism.

The $\lambda$ vector is nonnegative while the $Yb$ and $Vg$ vectors have a support of negative and positive connected components.
4. FINAL COMMENTS

Metamerism is an important aspect of the relationship between colours and light spectra. Metamerism is important e.g. in the textile and printing industries, in photography, etc.

At lifeless places both on earth and in the solar system, the spectral reflectance in the wavelength range from 100 nm to 1000 nm, of most surfaces is rather flat; with life, most earth surfaces have a rich and varied set of spectral reflectances. To infer cues from such spectra is an important aspect of seeing.

Tetrachromacy and in particular tetrachromatic metamerism is a subject well worth of attention. It has applications both in computer vision and in the modeling of biological vision systems. The subject of tetrachromatic metamerism in computer vision has applications for example in detection, in satellite images.

It is an interesting fact of our colour vision that we have four perceptual unique hues red, green, blue and yellow; perhaps they are in a one-to-one correspondence with the three channels \(S + L\) (“band-stop”), \(M\) (band-pass), \(S\) (low-pass) and \(L + M\) (high-pass). The fact that the \(L\) channel does not appear in an isolated form here, might have to do with the fact that it was the last to evolve.

It would be interesting to know how these facts extrapolate in cases of the vision systems of tetrachromatic animals. One possibility is that they might perceive 6 unique hues, corresponding to the cases \(W + X\) (low-pass), \(Y + Z\) (high-pass), \(X + Y\) (band-pass), \(W + Z\) (stop-band) and, \(W + Y\) and \(X + Z\) (alternate band).

In a trichromatic context,\(^9\) has shown how the reflectance spectra (sampled at N=40 wavelengths) of a set of 150 Munsell chips, turned out to be nearly three-dimensional; i.e. each spectrum is nearly a linear combination of certain three spectra. The analysis of large sets of natural reflectance spectra surely gives interesting results.

In a tetrachromatic vision system, the use detector with a bell response curve having a peak between those of the \(S\) and \(M\) detectors, could prove to be useful to differentiate between certain types of cyan allowing the perception a certain type of cyan as a unique and not as a combination of green and blue. This would be certainly useful for marine vision since short-wavelength light penetrates water more than other wavelengths.

Typically, receptor curves are unimodal. In biology, although not always in engineering as the example in Section 2.3 shows, each receptor curve ”aperure-samples” the visible spectrum, each sampling the energy in a, maybe overlapping, interval. In engineering, the use of detectors of comb spectra might be useful as well.

REFERENCES

Figure 6. Obtention of the kernel $K$ (in bottom row) of $[p, r, g, b]^T : \mathbb{R}^{76} \rightarrow \mathbb{R}^4$. In each graph, respectively from above, 75, 74, 73 and 72 row vectors are plotted.
Figure 7. Four types of spectra corresponding to Schrödinger’s optimal object colours: band-pass, stop-band, high-pass and low-pass.

Figure 8. The kernels of L+M+S (in yellow) and of VG and YB (in blue). The intersection of the three kernels (the yellow dark blue region) is the kernel of M; the dark blue together with the yellow dark blue region is the chrominance kernel.
Figure 9. Chromatic triangles in RGB cube and hypercube.