Vision Research 51 (2011) 2224–2233

Contents lists available at SciVerse ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

The effect of photopigment optical density on the color vision of the anomalous trichromat

P.B.M. Thomas [a,b,*], M.A. Formankiewicz [b,c], J.D. Mollon [b]

[a] Department of Ophthalmology, Addenbrooke's Hospital, Cambridge, United Kingdom
[b] Department of Experimental Psychology, University of Cambridge, Cambridge, United Kingdom
[c] Anglia Vision Research, Department of Vision and Hearing Sciences, Anglia Ruskin University, Cambridge, United Kingdom

ARTICLE INFO

Article history:
Received 6 February 2010
Received in revised form 13 August 2011
Available online 27 August 2011

Keywords:
Anomalous trichromacy
Photopigment
Optical density
Computational modeling
Color vision

ABSTRACT

We present a theoretical model to estimate the influence of photopigment optical density (OD) on the color vision of anomalous trichromats. Photopigment spectral sensitivities are generated using the Lamb (1995) template, which we correct for OD and pre-receptoral filters. Sixteen hyperspectral images (Foster, Nascimento, & Amano, 2004; Nascimento, Ferreira, & Foster, 2002) are analyzed, and the signals produced in the post-receptoral channels calculated. In the case of anomalous trichromats whose two longer-wavelength cones have peak sensitivities that lie close together in the spectrum, color vision can be substantially enhanced if the cones differ in optical density by a realistic amount.

1. Introduction

There is wide variation in color vision among anomalous trichromats: extreme anomalous trichromats exhibit similar levels of chromatic discrimination to the dichromat, whereas minimally affected anomalous trichromats perform near normally on pseudo-isochromatic plates and on discrimination tests (such as the Farnsworth-Munsell 100-hue test), but make color matches typical of simple anomalous trichromats (Vierling, 1935). Much of this variation has been attributed to the presence, within the human population, of photopigments that peak at many different spectral positions (Alpern & Moeller, 1977; Alpern & Wake, 1977; Asenjo, Rim, & Oprian, 1994; Merbs & Nathans, 1992). Since color vision relies ultimately on comparison of the output of different cone classes, any factor reducing the difference between these outputs should reduce the quality of color vision. Thus, an anomalous trichromat whose residual cone classes (M and M′ for the protanomal; L and L′ for the deuteranomal[1]) have similar peak sensitivities will have poorer color vision than another anomalous trichromat who has more widely separated peak sensitivities (Alpern & Moeller, 1977; Alpern & Wake, 1977). This argument has been called the "spectral proximity hypothesis" by Regan, Reffin, and Mollon (1994).

Variation in photopigment peak sensitivity is sufficient to explain some of the variation in the color matches made by anomalous trichromats, and in the quality of color vision they enjoy. In an earlier paper, we presented a model of the Rayleigh matching behavior of anomalous trichromats (Thomas & Mollon, 2004). By manipulating the peak sensitivities of modeled observers, typical deuteranomalous and protanomalous match mid-points and ranges could be predicted. However, manipulation of peak sensitivities alone does not explain the existence of observers who make match mid-points typical of anomalous trichromacy, but exhibit matching ranges that are paradoxically small (Hurvich, 1972). In agreement with earlier work (He & Shevell, 1995; Sanocki, Teller, & Deeb, 1997), Thomas and Mollon's model showed that an important determinant of matching performance is the concentration of photopigment within the cones and thus the optical density of the photopigment. The model generated theoretical observers with non-normal mid-points but relatively constrained matching ranges.

Photopigment optical density exerts its effect through a process known as "self-screening" (Alpern, Fulton, & Baker, 1987; Brindley, 1953; Knowles & Dartnall, 1977): the presence of many photopigment molecules within a cone alters the overall spectral sensitivity of the cone. As photons of many wavelengths pass axially through the cone, they are non-uniformly absorbed: those of wavelength close to the peak sensitivity of the photopigment are more likely to be absorbed by the superficial photopigment molecules and thus photons of other wavelengths will be over-represented deeper in the cone. In this way, the incident light is filtered by photopigment as it travels through the cone. As a result, the spectral sensitivity

* Corresponding author at: Department of Ophthalmology, Addenbrooke's Hospital, Cambridge, United Kingdom.
E-mail address: peter.thomas@doctors.org.uk (P.B.M. Thomas).

[1] According to this nomenclature, the protanomal possesses cone classes S, M and M′, and the deuteranomal possesses cone classes S, L′ and L. The peak sensitivities of M′ and L′ lie between the peaks of the normal M and L cones.

0042-6989/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2011.08.016

curve of the cone is broader than that of the dilute photopigment. As the optical density of the photopigment increases, so the sensitivity curve of the cone becomes broader (Fig. 1). Two cones expressing the same photopigment at different optical densities will, therefore, have different spectral sensitivities, and comparison of their output will yield a color signal. A cone with a higher optical density will also be more sensitive across the spectrum (more photopigment means more photoisomerisations). There are reports that observers exist who gain color discrimination by such a comparison: Neitz et al. (1999) describe protanomalous observers who, according to genetic analysis, possess only two spectrally distinct photopigments (M and S), but achieve trichromatic vision. The authors suggest that the M photopigment is expressed at two different optical densities, thus supporting a limited discrimination in the red–green range.

In the literature, there is substantial variation in the values reported for photopigment optical density (OD). It also remains a matter of debate whether the optical density is usually equal for the different cone classes of a given observer. Miller (1972) reports ODs of 0.4–0.5 for the M cone in protanopes and 0.5–0.6 for the L cone in deuteranopes. Smith and Pokorny (1973) place these values at 0.3 and 0.4 respectively, while Burns and Elsner (1993) find a greater disparity, with values of 0.27 and 0.48. Shevell and He (1997) suggest that the OD of L may be higher than that of L′ in the deuteranomalous observer. Berendschot, van de Kraats, and van Norren (1996) report values of 0.39 for M and 0.42 for L in dichromats. Renner et al. (2004) find no significant difference between the OD of M and L: 0.66 and 0.65 respectively. The literature does not, then, provide us with any consensus on the amount and nature of OD variation among normal or anomalous trichromats. It does, however, support the notion that such variation may exist, and that the differences in OD between cone classes may be as large as 0.2 or more. Multiple factors will underlie this variation in OD, including the length of the cone outer segment, the concentration at which photopigment is expressed, and the quantal efficiency of the individual photopigment molecules (Penn & Williams, 1986). The stability of the photopigment will also affect the OD, and could be of particular importance in anomalous trichromats whose "hybrid" photopigments may have reduced stability (Williams et al., 1992). Several of these factors may vary over time within an individual. For example, the density of photopigment expression and the length of the rod outer segment are, in part, determined by the ambient light levels in rats through a phenomenon known as photostasis (Penn & Williams, 1986), and such a mechanism may exist in humans (Beaulieu et al., 2009).

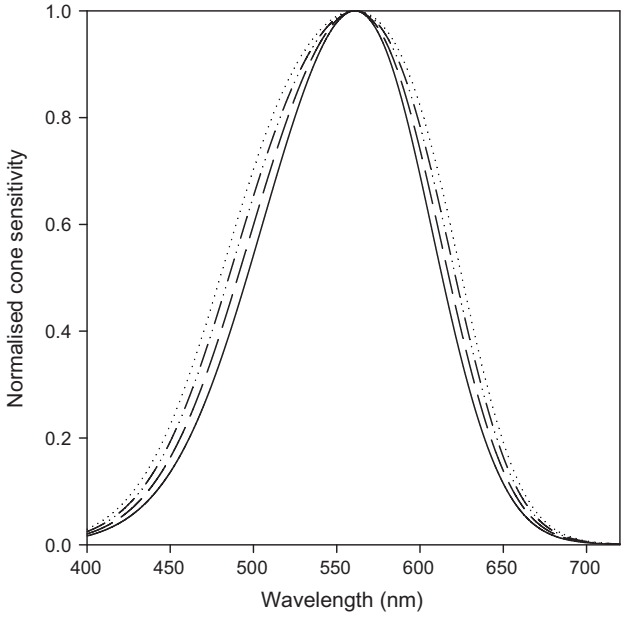

Fig. 1. The effect of photopigment optical density on the spectral sensitivity of a cone. Here, the sensitivity of a cone expressing photopigment at 561 nm is shown at OD of 0.1 (solid line), 0.3, 0.5 and 0.7 (from innermost to outermost). Note that the spectra here are normalized to their maxima. Without this normalization, the peaks of cones with greater OD would be higher than those of cones with a lower OD.

In this work we estimate the contribution that optical density variation could make to the real-world color vision of anomalous trichromats. In order to model this, we need to know the values of two factors: first, the cone sensitivities of the theoretical anomalous trichromat; second, the spectral composition of incident light from each point in real-world scenes.

1.1. Cone spectral sensitivities

There is no exhaustive database of human cone sensitivities that we can use to generate observers with photopigments of any peak wavelength and expressed at any optical density. Fortunately, it was noted by Dartnall (1953) that although photopigments vary in their wavelength of peak sensitivity (λ_{max}), they retain the same fundamental shape. He described this shape with a nomogram of sensitivity plotted against $1/\lambda - 1/\lambda_{max}$. Ebrey and Honig (1977) noted that the bandwidth of the sensitivity curves varied with λ_{max}, and introduced three separate nomograms to cover different parts of the spectrum. Mansfield (1985) found that description of a template on a normalized frequency axis allowed the return to a single template to cover the entire spectrum. Following this realization, a number of generalized templates have been developed (Baylor, Nunn, & Schnapf, 1987; Govardovskii et al., 2000; Lamb, 1995).

In this work, we use the Lamb (1995) template to define sensitivity spectra for photopigments of any given λ_{max} (in earlier work, we used the Baylor, Nunn, and Schnapf (1987) template, which gave very similar results). Lamb (1995) validated his template against data from eight psychophysical and electrophysiological studies on human, bovine, monkey, and squirrel subjects. Having generated the photopigment spectrum, we correct it for a given optical density. Thus, we can produce cone sensitivity triplets for all theoretical deuteranomalous and protanomalous observers (i.e. all combinations of wavelength of peak sensitivity and photopigment optical density).

1.2. The spectral composition of real world scenes

We use the hyperspectral images of Foster, Nascimento, and Amano (2004) and Nascimento, Ferreira, and Foster (2002). To construct these images, multiple photographs were taken of the same scene through narrowband filters centered on different wavelengths. In this way, the spectral flux from each point could be determined. The technique amounts to spectroradiometry with preservation of spatial information.

With knowledge of the cone spectral sensitivities and of the spectral reflectances of real world scenes, we can calculate the cone excitations produced in any observer by any of our scenes under any illuminant. We then use simple metrics to estimate the impact of small changes in peak sensitivity and optical density on the gamut of colors potentially available to the observer, and in doing so we assess the relative importance of peak separation and optical density to the color vision of the anomalous trichromat.

2. Methods

Matlab (The Mathworks Inc., Natick, USA) was used for all computational modeling.

2.1. Generation of spectral sensitivities

In this analysis, we vary the peak sensitivities and the optical densities of only the L and M photopigments and their anomalous equivalents. The corneal sensitivity of the S cone was taken from Smith and Pokorny (1975). M and L sensitivities were generated as follows:

(1) *Photopigment spectra* are generated according to the Lamb (1995) template for any value of λ_{max}:

$$S(\lambda) = \left(\exp a\left(A - \frac{\lambda_{max}}{\lambda}\right) + \exp b\left(B - \frac{\lambda_{max}}{\lambda}\right) + \exp c\left(C - \frac{\lambda_{max}}{\lambda}\right) + D\right)^{-1}$$

where $a = 70$; $b = 28.5$; $c = -14.1$; $A = 0.880$; $B = 0.924$; $C = 1.104$; $D = 0.655$; λ_{max} is the wavelength of peak sensitivity. We choose 531 nm as the peak sensitivity for M and 561 nm for L. Peaks for M′ and L′ lie between these two values.

(2) The photopigment spectra are then corrected for self screening to yield the *cone spectral sensitivity*:

$$\text{Cone sensitivity} = 1 - 10^{-(\text{POD}\times S(\lambda))}$$

where $S(\lambda)$ is the extinction coefficient (as a function of wavelength) of the photopigment and POD is the photopigment optical density.

(3) The cone spectral sensitivities are then corrected for the effect of the pre-retinal filters. We use the data of Stockman, Sharpe, and Fach (1999) for the lens density spectrum, and of Bone, Landrum, and Cains (1992) for the macular pigment density spectrum. In this way, the *corneal spectral sensitivities* are generated for the observer. The product of the corneal spectral sensitivity and an incident spectral light gives the stimulation of the photoreceptor.

(4) Finally, the sensitivity spectra of the cones are scaled. There are no experimental data from which we can extract suitable scaling factors for anomalous trichromats. As a default solution, we assume the same scaling factors apply as in the normal trichromat of Smith and Pokorny (1975), where L contributes twice as much to the post-receptoral channels as M. Therefore, we scale the longer wavelength sensitive cone (L in the deuteranomal, M′ in the protanomal) to have twice as large an area under its spectral sensitivity curve as the shorter wavelength sensitive cone (L′ in the deuteranomal, M in the protanomal). This assumes that anomalous trichromacy is produced by a simple substitution of photopigment, and that all retinal circuitry is unchanged. As an alternative possibility we assume that there is no mechanism by which scaling factors are fixed between observers, and that the signal passed from photoreceptor to downstream neurons is related to the number of photoisomerizations. In this case we apply no scaling to the cone sensitivity spectra after correction for optical density.

To verify the method of calculating spectral sensitivities described above, we compared the cone sensitivity spectra for an observer with M and L photopigment peaks of 531 and 561 respectively and ODs of 0.4 (whom we define as normal in the results section) to the cone fundamentals currently recommended by the CIE (Stockman & Sharpe, 2000). Fig. 2 shows that the L and M cone spectral sensitivities we generated agree well with those of Stockman and Sharpe.

Thus our model will generate a theoretical observer who would exhibit color-matching functions similar to those of a standard observer. We do not set out here, however, to fit the standard observer exactly, in part because there are known to be large individual differences in the normal population, but primarily because our purpose in this paper is to illustrate the *trajectories* of change in discrimination that would be expected as the optical density and the peak sensitivities of the photopigments are varied.

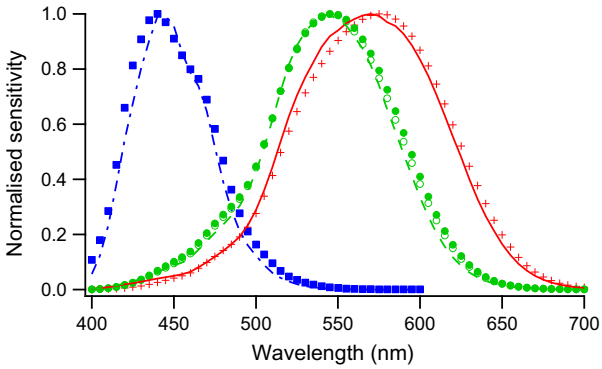

Fig. 2. A comparison of L (+), M (●) and S (■) cone sensitivity spectra of the observer defined as normal in this paper and the L (-), M (- - -) and S (· - · -) cone fundamentals of Stockman and Sharpe (2000). For the M cone, we also show the derived cone sensitivity spectrum (○) when OD has been changed from 0.4 as used in our model to 0.3 as suggested by Smith and Pokorny (1973).

2.2. Hyperspectral imaging details

Our hyperspectral images were drawn from the study of Foster, Nascimento, and Amano (2004) and Nascimento, Ferreira, and Foster (2002). We used the eight publicly available images from each of these sets. The former are depicted in Fig. 1 of Foster, Nascimento, and Amano (2004) and consist of five rural and three urban scenes. The latter are depicted in Fig. 2 of Nascimento, Ferreira, and Foster (2002) and consist of four rural scenes, three urban scenes, and one of colorful objects (primarily toys). Each hyperspectral image comprises 31–33 narrowband (minimum 7 nm, maximum 16 nm bandwidth) images centered on wavelengths at increments of 10 nm between 400 and 720 nm (400 nm and 720 nm images are often omitted owing to noise). The incident light was converted to the reflectance by comparison with the spectral energy distribution of an object in the scene of known spectral reflectance. Some corrections were applied to these images to correct for factors such as dark noise and stray light. We refer the reader to Nascimento, Ferreira, and Foster (2002) and Foster, Nascimento, and Amano (2004) for a full description of the techniques.

2.3. Calculating cone and channel excitations

The hyperspectral data provide us with the reflectance spectrum of each pixel in the image. The product of reflectance spectrum and an illuminant gives the spectral distribution of light incident on the observer's cornea. The illuminant was recorded individually for each of the Foster, Nascimento, and Amano (2004) scenes, while CIE D65 illuminant was used for the Nascimento, Ferreira, and Foster (2002) scenes. Photoreceptor stimulation is the product of the spectrum of incident light and the corneal sensitivities of the observer. For each of 10,000 randomly selected pixels from each scene, we determine a triplet of excitations, one for each class of cone possessed by the trichromat. The post-receptoral signals for each pixel are then calculated by the comparisons analogous to S/(L + M) and L/(L + M) in the normal. For the protanomal, these channels would correctly be labeled S/(M + M′) and M′/(M + M′); for the deuteranomal, S/(L + L′) and L/(L + L′).

2.4. Analysis of data

To assess the color information available in a given scene for a given observer, we can plot each pixel in a chromaticity diagram of the type that was introduced for normal observers by MacLeod

and Boynton (1979) and was generalized to other phenotypes by Regan et al. (1998). The abscissa is L/(L + M) for the normal, M′/(M′ + M) for the protanomal, and L/(L + L′) for the deuteranomal. The ordinate is S/(L + M) for the normal, S/(M′ + M) for the protanomal, and S/(L + L′) for the deuteranomal. Expressing the richness of the gamut in a single number is challenging since the level of color vision will be determined not only by the absolute bounds of this distribution, but also by how well the signals are spread within these bounds. As a simple method we calculate the variance of the signal along the abscissa. Our preferred metric, however, was "cell-counting". The color space was divided into cells of 0.0001 along the abscissa and 0.0005 along the ordinate. Discriminations based on the short-wave cones are known to be less sensitive than those based on long- and middle-wave cones, although the relative sensitivity of the short-wave channel depends on the spatial and temporal properties of the stimulus (Mollon, 1982). Here we adopt the ratio of 1:5 for relative sensitivity on the vertical and horizontal axes of the MacLeod–Boynton space, basing this value on Table 1 (7.4.1) of Wyszecki and Stiles (1982), which holds for a 1°, 200-ms target. The number of cells containing a signal was then calculated, and this single number was used as a measure of color vision. The number does not represent an estimate of the number of colors visible, but it should be related to this quantity.

In some parts of the analysis we reduce our four variables (two peak sensitivities and two optical densities) to two: peak separation and optical density disparity. We define the optical density disparity as the difference in OD between the photopigments in the anomalous trichromat equivalent to L and M, and calculate it as (OD of longer wavelength photopigment) – (OD of shorter wavelength photopigment). For the protanomal this is (OD of M′) – (OD of M), for the deuteranomal (OD of L) – (OD of L′). The peak separation is simply the difference in nm between the peak sensitivities of the two longer wavelength photopigments of the anomalous trichromat.

3. Results

3.1. Spectral sensitivities of anomalous cones

As argued in the Introduction, it is not the absolute shapes of the cone sensitivities that potentially determine the acuteness of color discrimination, but rather the extent to which the various cone sensitivities of a given observer differ. If OD variation is coupled in the cones of the anomalous trichromat (i.e. the OD values of M and M′ or of L and L′ are set equal), the shapes of the cone sensitivity spectra will change as OD is changed, but the relative difference between the spectra will remain largely unchanged: they will still differ in peak sensitivity, but not in bandwidth. However, if we allow uncoupled OD variation (i.e. non-equal values of OD in the M and M′ cones of the protanomal, or L and L′ cones of the deuteranomal), we can produce relative change between the sensitivity spectra which should effect a change in the quality of color vision. Fig. 3 illustrates this effect. When a positive optical density disparity is expressed, the long-wavelength limbs of the L and L′ cones of the deuteranomal become more

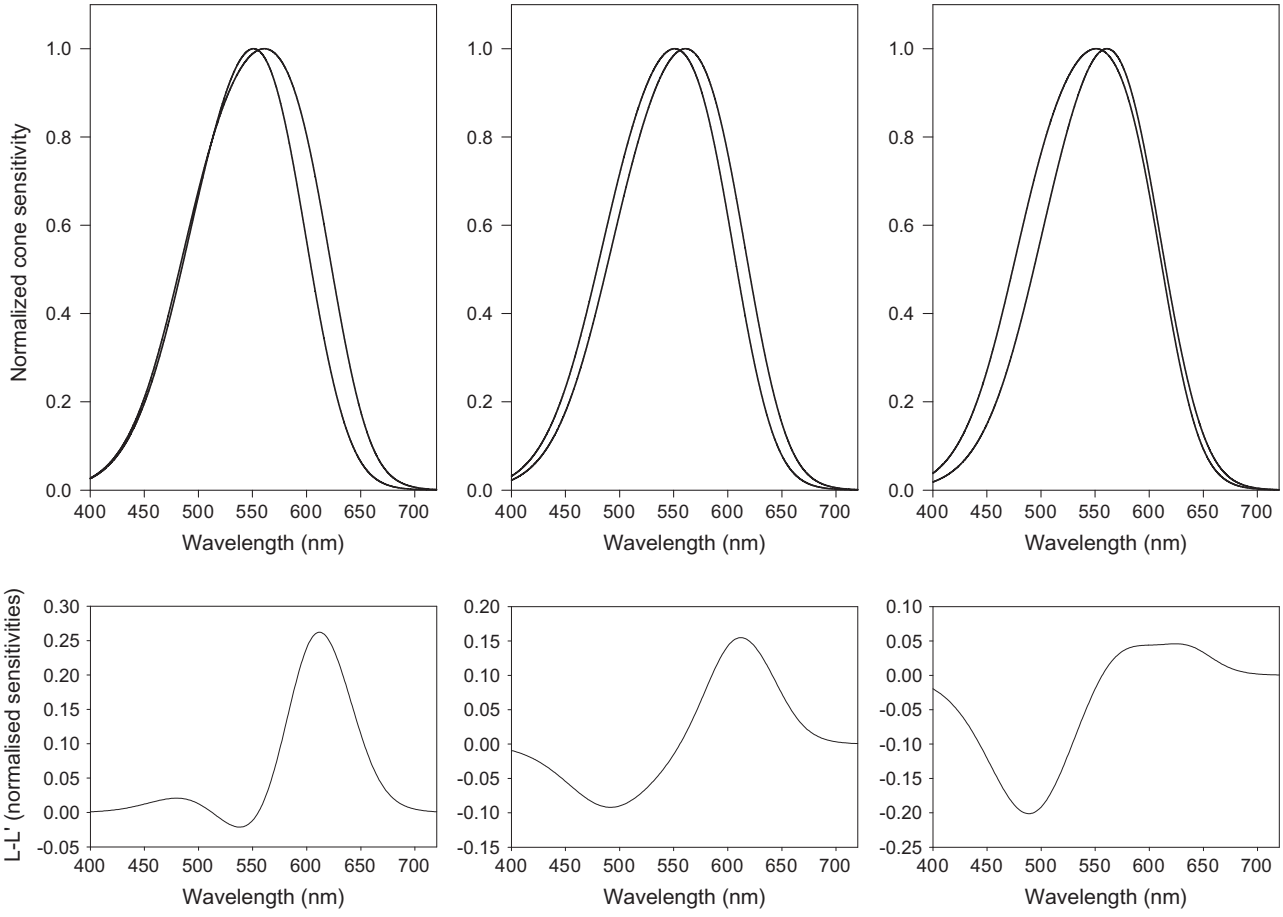

Fig. 3. The spectral sensitivities of the long-wave photopigments of a deuteranomalous trichromat with positive (above, left), no (above, middle), and negative (above, right) optical density disparity. The lower three plots show the difference between the spectral sensitivities plotted above them (difference = L – L′).

different at the cost of greater similarity of the short-wavelength limbs. The reverse holds true for negative disparities. The same pattern is observed for protanomals.

3.2. The effect of optical density on color vision

For any theoretical observer (possessing any combination of peak sensitivity and optical density) viewing any modeled scene, our cell-counting method yields a single number that is related to the richness of his color vision. However, this number will also be affected by the quantity of chromatic information potentially available in the scene (clearly some scenes will give rise to a more varied chromatic percept than others). Therefore, in order to summarize the data for many scenes we normalize the cell count for each theoretical observer to the cell count for a "normal" observer viewing the scene. We define our normal observer as having peaks of M and L at 531 nm and 561 nm respectively, and expressing both pigments at an OD of 0.4. We can then express the cell count of each theoretical observer as a fraction of the cell count of the normal observer viewing the scene, and take the mean of this fraction to summarize the data from many scenes. Fig. 4 shows this summary for theoretical deuteranomals and protanomals viewing the two sets of hyperspectral images. The default assumption of a 2:1 scaling of L:L′ and M′:M is made here – other combinations of cone scaling are considered in the next section.

Data for each individual scene produce similar plots to the summary data shown in Fig. 4 (i.e. the content of the scene has little effect on the relationship between cone ODs and cell count). A few general observations can be made. First, for positive OD disparities, increased peak separation increases the cell count. Second, where there is a non-zero peak separation, a positive OD disparity is preferable to a negative OD disparity of the same magnitude. Third, each plot shows a trough, i.e. a locus of ODs that gives the lowest cell count for any given peak separation. Fig. 5 allows clearer visualization of the movement of this trough to more negative OD disparities as peak separation increases. An important observation is that the trough for any given peak sensitivity falls along a locus described by a fixed OD disparity. It is therefore primarily the OD disparity that influences the cell count (and hence color vision) and not the absolute values of the OD.

3.3. Effect of cone scaling

The default assumption in our model is that the two longer wavelength sensitive cones of the anomalous trichromat show the same 2:1 scaling as do L:M in the normal (see Section 4). Since we cannot empirically justify this assumption, we also considered the effect of other scaling ratios. Fixed ratios of 1:1 and 1:2 (for L:L′ or M′:M) do not change the relationships described in this paper. We also considered the possibility that there is no mechanism that

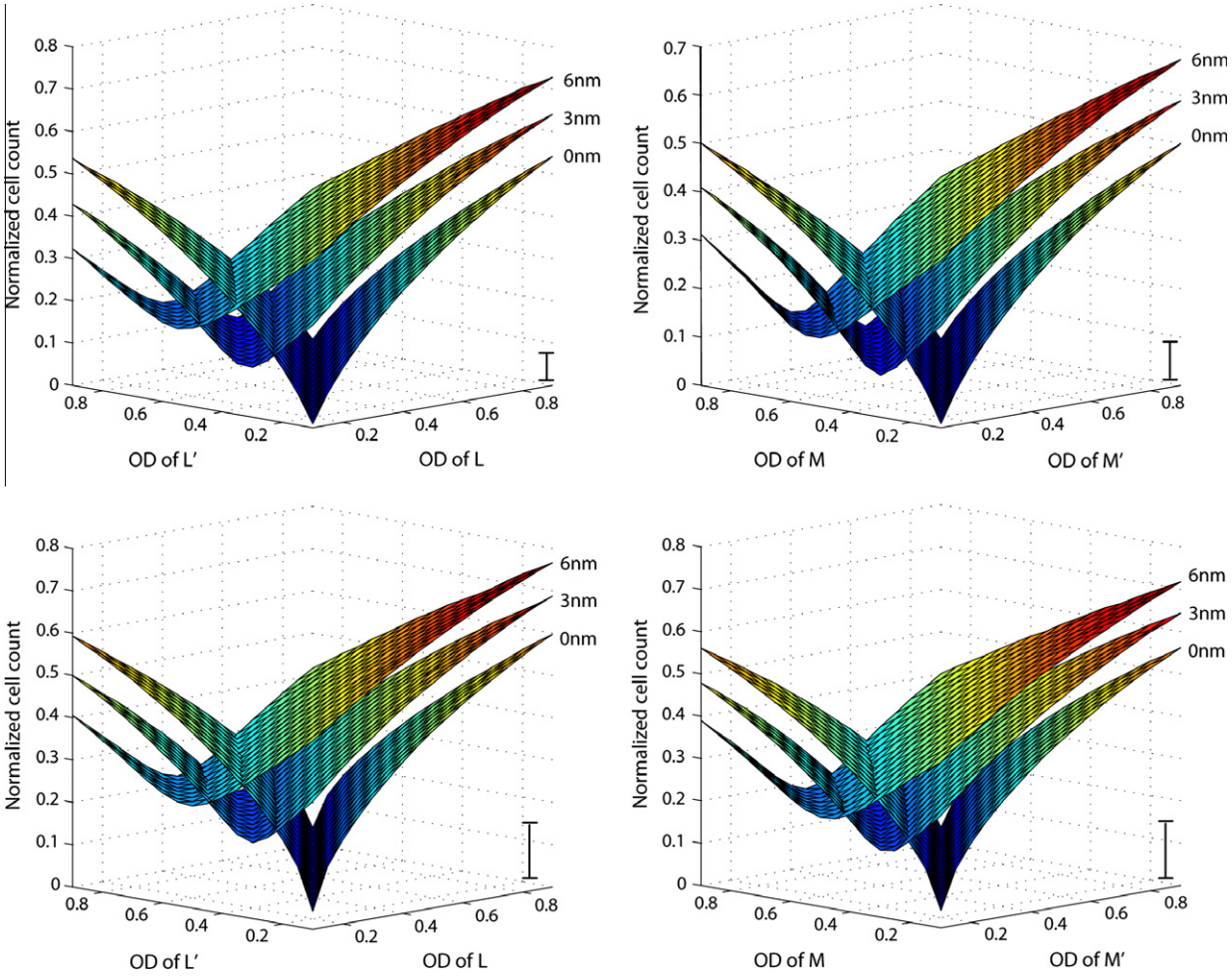

Fig. 4. The influence of photopigment OD on the normalized cell count. The upper two plots show the data for modeled deuteranomalous (left) and protanomalous (right) observers viewing the Nascimento, Ferreira, and Foster (2002) scenes. The lower two plots are for the eight Foster, Nascimento, and Amano (2004) scenes. The three surfaces in each plot correspond to peak separations of 0, 3, and 6 nm (see labels), and are produced by averaging the normalized cell count across the eight scenes in each set (see text). The error bar shown in the lower right corner of each plot represents the maximum standard deviation of the average taken across the eight scenes.

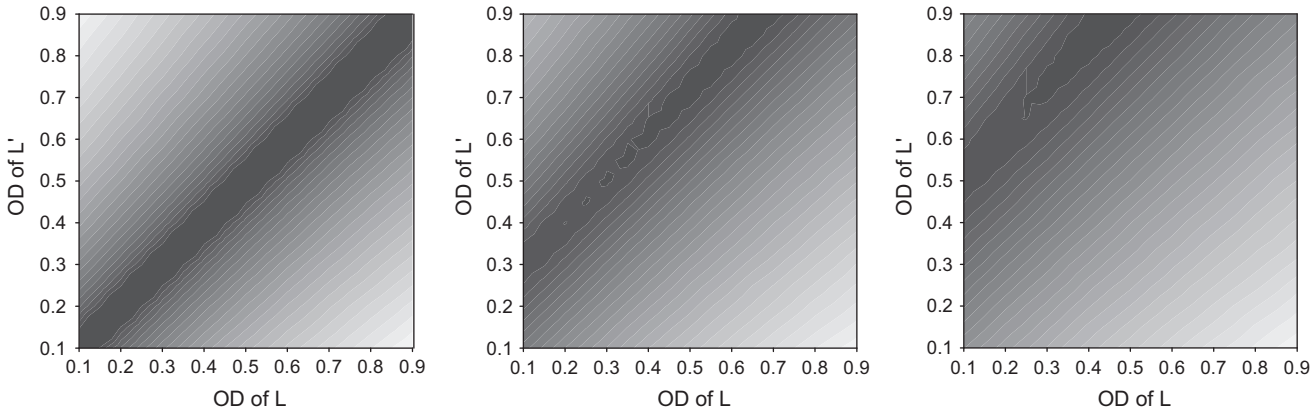

Fig. 5. The three surfaces show data for modeled deuteranomalous observers with peak separation of 0 nm (left), 3 nm (middle), and 6 nm (right) viewing the Foster, Nascimento, and Amano (2004) scenes. The shading indicates the cell count: lighter shading corresponds to higher cell count. The trough moves to more negative OD disparities as the peak separation increases.

fixes the ratio of L:L′ or M′:M in the deuteranomal and protanomal respectively. It is possible that the level of stimulation produced in the bipolar cells by cone excitation is determined purely by the number of photoisomerizations in the cone outer segment. A cone with high optical density will, then, be very much more sensitive than one with low optical density. Fig. 6 shows data for the eight Foster, Nascimento, and Amano (2004) scenes where the OD determines the scaling of the cone sensitivity spectra. This scaling does not alter the influence of OD on color vision across most of the range of OD considered. An exception must be made where one cone has a very low optical density compared to the other (i.e. large OD disparities): further increasing the disparity in this condition does not continue to increase the cell count as it does in the fixed ratio conditions.

3.4. Effect of OD disparity on color vision

Figs. 4–6 show how optical density affects color vision as measured with our cell-counting method. The troughs produced in the three-dimensional plots run along constant OD disparities for a given peak separation. For ease of visualization, then, it is valid to reduce these plots to two dimensions: OD disparity and cell count. Fig. 7 shows the averaged data for all eight of the 2004 scenes considered for a set of theoretical protanomals and deuteranomals (the OD of the longer wavelength sensitive photopigment is fixed at 0.4). Plotted in this way, the effect of OD density on cell count is easier to quantity. In general, an increase in positive OD disparity of 0.1 enhances the gamut of chromaticities by an amount equivalent to an additional 1 nm or so of peak separation. Again, this behavior is apparent when the scenes are considered individually rather than averaged.

We also considered the effect of OD disparity on the variance of signals produced in the channels analogous to L/(L + M). The averaged data show similar behavior to the cell-counting data, with a trough moving to increasingly negative OD disparity as the peak separation increases. However, an averaged plot would hide significant variation between the scenes. To illustrate this, the variance data from three of the Foster, Nascimento, and Amano (2004) scenes are shown in Fig. 8 for deuteranomalous observers. Scene 8 (rightmost plot) from the 2004 images is the only one in which

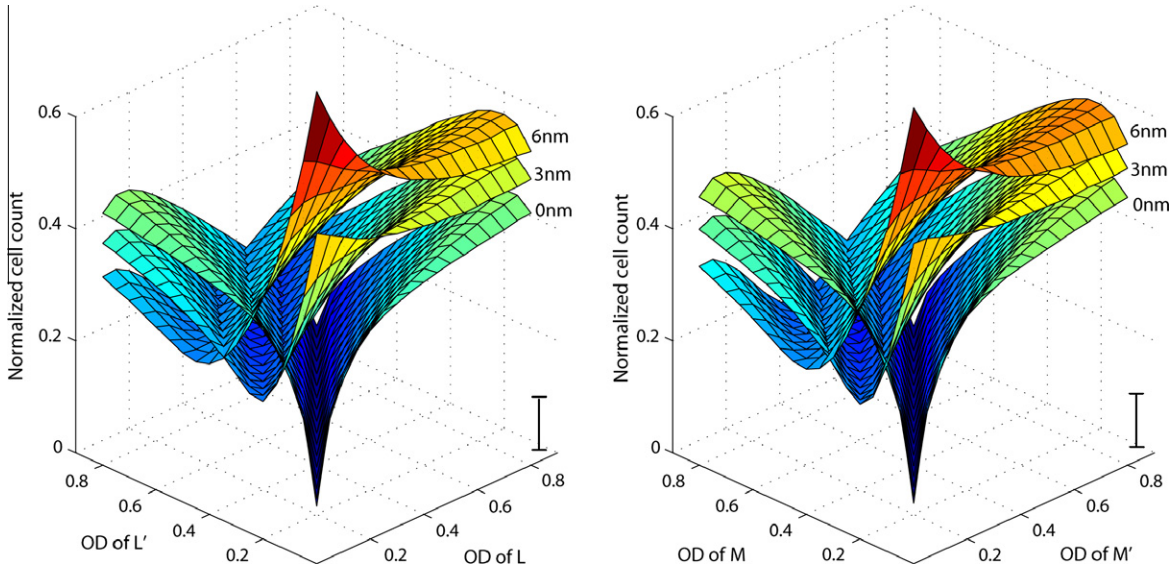

Fig. 6. Deuteranomalous (left) and protanomalous (right) observers viewing the Foster, Nascimento, and Amano (2004) scenes. Here the cone scaling is determined purely by the OD of the photopigment they contain. For each plot, the three surfaces correspond to peak separations of 0, 3, and 6 nm (labels), and the normalized cell count is averaged across the eight scenes. The error bar shown in the lower right corner of each plot represents the maximum standard deviation of the average taken across the eight scenes.

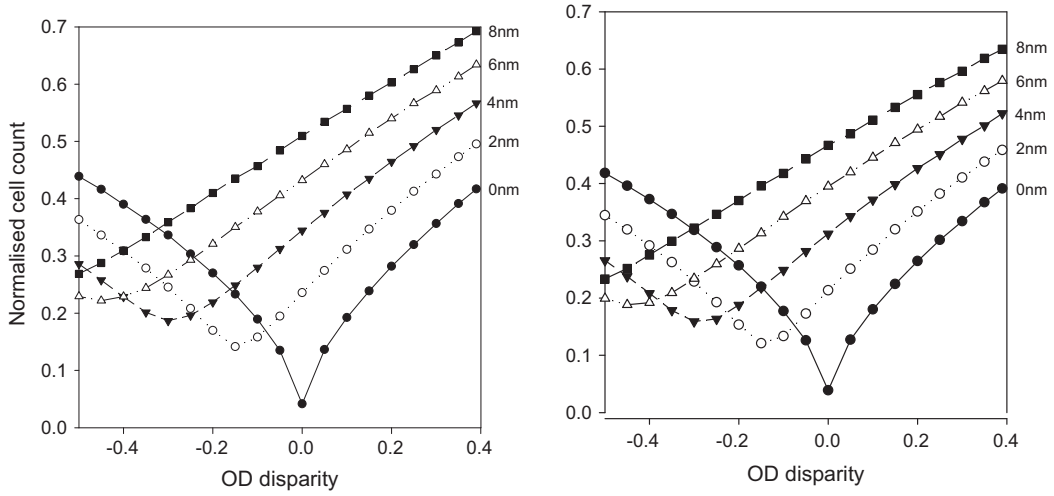

Fig. 7. The influence of OD disparity on normalized cell count across the eight Foster, Nascimento, and Amano (2004) scenes for deuteranomalous (left) and protanomalous (right) observers. Each line shows a different peak separation: 0, 2, 4, 6, and 8 nm (labeled).

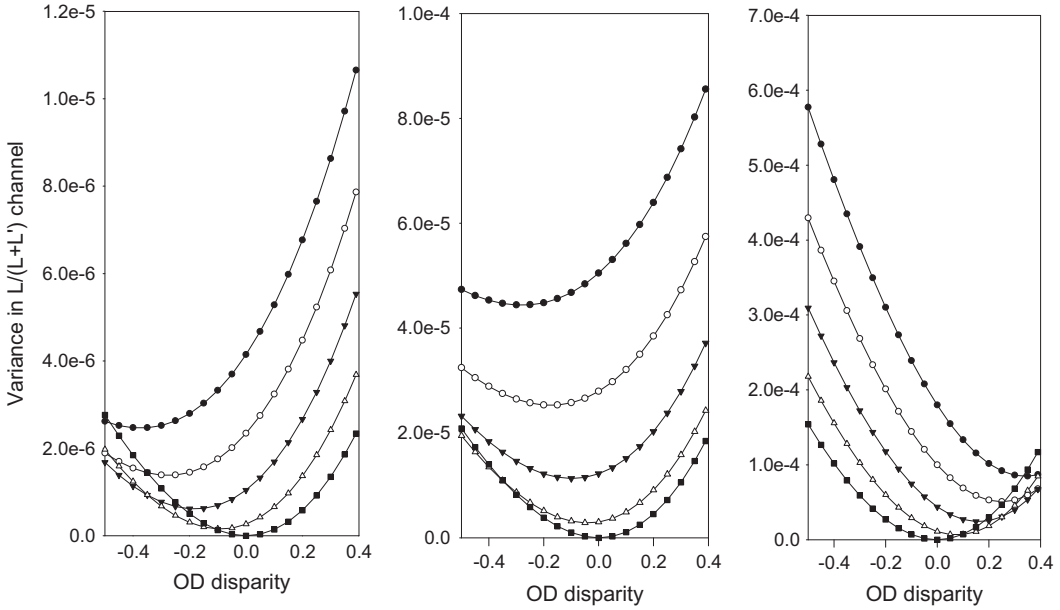

Fig. 8. Variance plots of deuteranomalous observers viewing scenes 1 (left), 4 (middle) and 8 (right) of the Foster, Nascimento, and Amano (2004) hyperspectral images. The five lines in each represent peak separations of 0 nm (solid squares), 2 nm (hollow triangles), 4 nm (solid triangles), 6 nm (hollow circles), and 8 nm (solid circles).

the trough moves to positive OD disparities as the peak separation increases (this scene shows a structure of stone and wood). The cell count metric for this scene, however, shows the same behavior as the averaged cell-counting data.

The discrepancy between cell count and variance in the anomalous equivalent of the L/(L + M) channel can be explained by the distribution of chromaticities in the modeled observer's color space. Fig. 9 shows the chromaticities of the 10,000 randomly selected pixels from scene 8 of Foster, Nascimento, and Amano (2004) in the color space of a deuteranomalous observer with negative, zero, and positive OD disparities. The majority of chromaticities in this scene fall in one of two large clusters seen in all three color spaces: the first is vertically oriented at values of S/(L + L′) greater than 1; the second is a large cluster of chromaticities at low values of S/(L + L′). This latter cluster moves along the ordinate relative to the former as OD disparity is changed. In the negative OD disparity condition, the two clusters lie at quite different values of L/(L + L′), and so the overall variance in the L/(L + L′) channel is high. In the positive OD condition, the clusters occupy similar L/(L + L′) values, so the variance in this channel is low. However, the two clusters can still be easily distinguished by their S/(L + L′) signals in the positive OD disparity condition: separation along the L/(L + L′) axis merely replicates information available in the S/(L + L′) channel. The remainder of chromaticities outside these two clusters are better spread in the positive OD disparity condition. Therefore, the cell count method favors the positive OD disparity condition since it measures spread in the two-dimensional color space, and not just along a single axis.

4. Discussion

4.1. The effect of optical density on color vision

It is established that variations in optical density must affect color vision, and that this effect is especially likely to be important

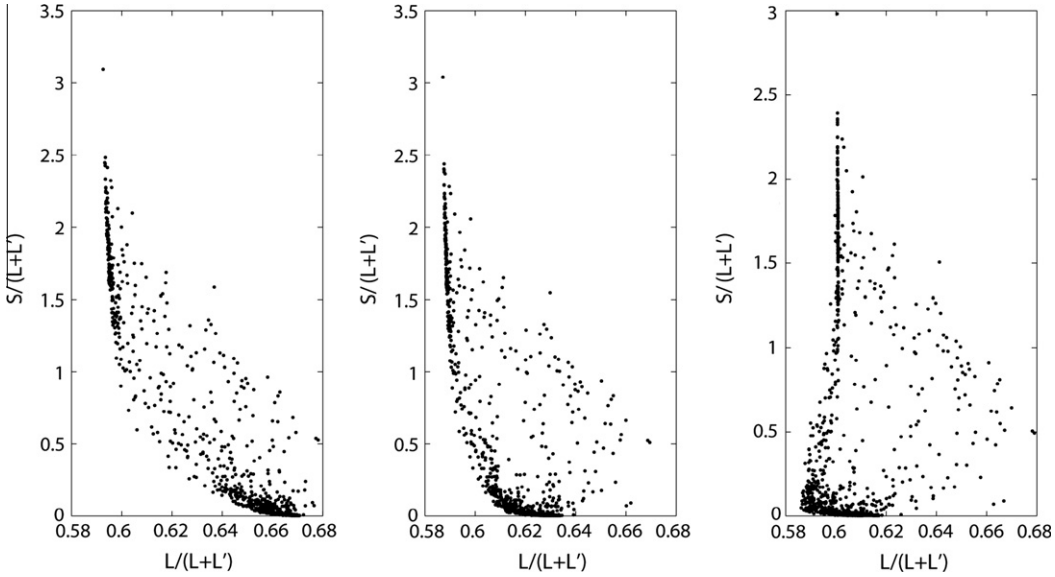

Fig. 9. Chromaticities of the 10,000 randomly selected pixels from scene 8 of Foster, Nascimento, and Amano (2004) in the color space of modeled deuteranomalous observers with 4 nm (peaks at 561 nm and 557 nm) peak separation and OD disparities of −0.2 (left), 0 (middle), and +0.2 (right).

when peak separation is low (Barbur et al., 2008; Neitz et al., 1999; Shevell & He, 1997). The present analysis quantifies the importance that OD might play in the day-to-day color vision of the anomalous trichromat. Our modeling suggests that an OD disparity of +0.1 is equivalent to an additional 1–1.5 nm peak separation in deuteranomals and protanomals, and that this is relatively unaffected by the absolute values of OD (i.e. only the disparity is important). A similar statement can be made for normal trichromats, though in this case the improvement will represent a small fractional improvement in color vision subserved by a large peak separation. We suggest, then, that optical density could play a significant role in determining the chromatic percept of the anomalous trichromat, and that predictions of the quality of color vision should not be based on peak separation alone.

4.2. Why a positive disparity is good and a negative disparity is bad

It is not initially obvious why a positive OD disparity should be so much more beneficial than a negative one when one considers only the L/(L + M), L/(L + L′), or M′/(M′ + M) channels of normals, deuteranomals and protanomals respectively. From Fig. 3 it is clear that positive and negative disparities each enhance discrimination in one part of the spectrum to the detriment of discrimination in another part (positive disparities enhance long-wavelength discrimination, negative disparities enhance short-wavelength discrimination). Part of the explanation lies in the pre-retinal filters: these preferentially filter out short-wavelength light, favoring long-wavelength contribution to color vision. Moreover, while photopigment spectra generated from the Lamb (1995) polynomial show a regular slope at long wavelengths, there is a "shoulder" at short wavelengths where they start to become more sensitive again. This behavior is even more pronounced when using the Baylor, Nunn, and Schnapf (1987) polynomial. The separation between photopigment spectra of similar wavelengths of peak sensitivity is, then, not so well maintained at short wavelengths as it is at longer wavelengths.

However, we suggest that much of the asymmetry between positive and negative disparities arises because of the S/(L + M) channel (or its anomalous equivalent). When a negative disparity is expressed, both channels, i.e. S/(L + M) and L/(L + M) (or their anomalous equivalents), favor discrimination at short wavelengths to the detriment of long wavelengths. When a positive disparity is expressed, the L/(L + M) channel (or its anomalous equivalent) is better able to extract information from long wavelengths. The difference spectra for deuteranomalous cones at different optical density disparities in Fig. 3 illustrate this behavior: the negative disparity L/(L + L′) channel supports very little chromatic discrimination above 550 nm, while the positive disparity L/(L + L′) channel supports good chromatic discrimination above 550 nm (at the expense of shorter wavelengths). A positive disparity, therefore, gives a better compromise of discrimination throughout the spectrum.

The data presented here are consistent with the findings of an earlier model (Thomas, 2005). That model used the Baylor, Nunn, and Schnapf (1987) fundamental, eight Nascimento, Ferreira, and Foster (2002) scenes, and 10 hyperspectral images generated in our labs; we validated the findings against data from 80 objects sampled with a standard spectroradiometer.

4.3. Scaling of cone sensitivities

One default assumption in our modeling is that the 2:1 scaling of L:M derived by Smith and Pokorny (1975) applies to anomalous trichromats. In the case of the normal observer, a 2:1 ratio is required to construct the average luminous efficiency function from an additive sum of the L and M cone sensitivities. The standard explanation is that average fovea exhibits a 2:1 ratio in the numbers of L and M cones. Such a ratio is confirmed by the most direct test – microspectrophotometric measurements of cone outer segments from human foveas (Bowmaker, Parry, & Mollon, 2003) – but there are believed to be large individual differences in the ratio (De Vries, 1947; Roorda et al., 2001; Rushton & Baker, 1964). The cone ratio itself may depend on the proximal promoter regions of the L and M opsin genes and on the positions of the two genes relative to the upstream locus control region, as well as on more remote upstream regulatory regions (Gunther, Neitz, & Neitz, 2008); but the normal mechanism is not well enough understood to allow clear predictions for the 'hybrid' opsin genes of anomalous observers, where the promoter region may be drawn from one of the normal genes while the exons that determine spectral sensitivity may be drawn from the other (Deeb, 2006). Prediction of the scaling factor for anomalous observers is complicated by another interesting factor: if the opsins encoded by some 'hybrid' genes

are less stable or less efficient (Williams et al., 1992), and thus are effectively present in lower concentration, the result may be reduced modulation of the synaptic signal of a cone expressing such an opsin. Thus it is reassuring for our modeling that variation in the assumed scaling of L:L′ or M:M′ has rather little effect on the predictions. Empirically, in normal observers, large variations in L:M cone ratio appear to have only small, though detectable, effects on color discrimination (Gunther & Dobkins, 2002; Hood et al., 2006).

4.4. Limitations of the model

We have used a simplistic metric (cell-counting) for the assessment of the quality of color vision. This metric does not take into account the non-uniformity of color space that results from post-receptoral compression of chromatic signals (Boynton, Nagy, & Olson, 1983; Mollon & Estévez, 1988; Tyndall, 1933). An alternative approach would be to create a perceptually uniform space (Linhares, Pinto, & Nascimento, 2008). Although this method would allow us to estimate numbers of discriminable colors, it would require a number of additional assumptions to generalize such a space to observers possessing any combination of photopigment peak sensitivities and optical densities. Moreover, our measure of the influence of photopigment optical density is an internal comparison, quantified in terms of equivalent peak separation. Clearly the number of cells filled does not tell us the absolute number of colors visible to our observers. However, assuming that the post-receptoral mechanisms of all observers are equally sensitive in comparing the outputs from different cone classes, it is a valid means for comparing the amount of information available to those mechanisms. We additionally do not take account of the non-uniformity of color discrimination across the retina (Mullen & Kingdom, 1996; Weale, 1953), a factor that we assume will equally influence all observers.

4.5. Applications and further work

Earlier models of this kind (Mollon & Regan, 2001) have been used to generate simulations of the chromatic world that can be viewed by normal observers (this requires the careful generation of an image such that each pixel produces the same triplet of cone excitations in a normal observer as does the real scene in an anomalous trichromat). Such models are easily adapted to include optical density as a variable, and can be used to render visible the benefits conferred by a positive OD disparity.

Models like this could also be used to compare the power of various tests of color vision in predicting real world color vision. Using an earlier version of this model, for instance, we compared the predicted real-world color vision of a range of anomalous trichromats to their predicted Rayleigh matching performance (Thomas, 2005). We found that the correlation between match mid-point and quality of color vision among anomalous trichromats is poor when optical density is allowed to vary, but the relationship between matching range and level of color discrimination remains robust.

Our findings could also be applicable to theories of the evolution of color vision: if trichromacy is supported by an observer expressing a single long-wavelength photopigment at two densities, could such an arrangement have allowed trichromacy in primates before the evolution of a third photopigment?

Acknowledgments

We should like to thank David Foster, Kinjiro Amano, Sérgio Nascimento, and Flavio Ferreira for making their hyperspectral images available (Foster, Nascimento, & Amano, 2004; Nascimento, Ferreira, & Foster, 2002).

Anglia Vision Research is affiliated with the Vision and Eye Research Unit.

References

Alpern, M., Fulton, A. B., & Baker, B. N. (1987). "Self-screening" of rhodopsin in rod outer segments. *Vision Research, 27*(9), 1459–1470.
Alpern, M., & Moeller, J. (1977). The red and green cone visual pigments of deuteranomalous trichromacy. *Journal of Physiology, 266*(3), 647–675.
Alpern, M., & Wake, T. (1977). Cone pigments in human deutan colour vision defects. *Journal of Physiology, 266*(3), 595–612.
Asenjo, A. B., Rim, J., & Oprian, D. D. (1994). Molecular determinants of human red/green color discrimination. *Neuron, 12*(5), 1131–1138.
Barbur, J. L., Rodriguez-Carmona, M., Harlow, J. A., Mancuso, K., Neitz, J., & Neitz, M. (2008). A study of unusual Rayleigh matches in deutan deficiency. *Visual Neuroscience, 25*(3), 507–516.
Baylor, D. A., Nunn, B. J., & Schnapf, J. L. (1987). Spectral sensitivity of cones of the monkey *Macaca fascicularis*. *Journal of Physiology, 390*, 145–160.
Beaulieu, C., Rufiange, M., Dumont, M., & Lachapelle, P. (2009). Modulation of ERG retinal sensitivity parameters with light environment and photoperiod. *Documenta Ophthalmologica, 118*(2), 89–99.
Berendschot, T. T., van de Kraats, J., & van Norren, D. (1996). Foveal cone mosaic and visual pigment density in dichromats. *Journal of Physiology, 492*(Pt 1), 307–314.
Bone, R. A., Landrum, J. T., & Cains, A. (1992). Optical density spectra of the macular pigment in vivo and in vitro. *Vision Research, 32*(1), 105–110.
Bowmaker, J. K., Parry, J. W. L., & Mollon, J. D. (2003). The arrangement of L and M cones in human and a primate retina. In J. D. Mollon, J. Pokorny, & K. Knoblauch (Eds.), *Normal and defective colour vision* (pp. 39–50). Oxford: Oxford University Press.
Boynton, R. M., Nagy, A. L., & Olson, C. X. (1983). A flaw in equations for predicting chromatic color difference. *Color Research and Application, 8*, 69–74.
Brindley, G. S. (1953). The effects on colour vision of adaptation to very bright lights. *Journal of Physiology, 122*(2), 332–350.
Burns, S. A., & Elsner, A. E. (1993). Color matching at high illuminances: Photopigment optical density and pupil entry. *Journal of the Optical Society of America A. Optics and Image Science, 10*(2), 221–230.
Dartnall, H. J. A. (1953). The interpretation of spectral sensitivity curves. *British Medical Bulletin, 9*(1), 24–30.
De Vries, H. (1947). The heredity of the relative numbers of red and green receptors in the human eye. *Genetica, 24*, 199–212.
Deeb, S. S. (2006). Genetics of variation in human color vision and the retinal cone mosaic. *Current Opinion in Genetics and Development, 16*(3), 301–307.
Ebrey, T. G., & Honig, B. (1977). New wavelength dependent visual pigment nomograms. *Vision Research, 17*(1), 147–151.
Foster, D. H., Nascimento, S. M., & Amano, K. (2004). Information limits on neural identification of colored surfaces in natural scenes. *Visual Neuroscience, 21*(3), 331–336.
Govardovskii, V. I., Fyhrquist, N., Reuter, T., Kuzmin, D. G., & Donner, K. (2000). In search of the visual pigment template. *Visual Neuroscience, 17*(4), 509–528.
Gunther, K. L., & Dobkins, K. R. (2002). Individual differences in chromatic (red/green) contrast sensitivity are constrained by the relative number of L- versus M-cones in the eye. *Vision Research, 42*(11), 1367–1378.
Gunther, K. L., Neitz, J., & Neitz, M. (2008). Nucleotide polymorphisms upstream of the X-chromosome opsin gene array tune L:M cone ratio. *Visual Neuroscience, 25*(3), 265–271.
He, J. C., & Shevell, S. K. (1995). Variation in color matching and discrimination among deuteranomalous trichromats: Theoretical implications of small differences in photopigments. *Vision Research, 35*(18), 2579–2588.
Hood, S. M., Mollon, J. D., Purves, L., & Jordan, G. (2006). Color discrimination in carriers of color deficiency. *Vision Research, 46*(18), 2894–2900.
Hurvich, L. M. (1972). Color vision deficiencies. In D. Jameson & L. M. Hurvich (Eds.). *Visual psychophysics* (Vol. 7/4, pp. 582–624). Berlin: Springer-Verlag.
Knowles, A., & Dartnall, H. J. A. (1977). The photobiology of vision. In H. Davson (Ed.). *The eye* (Vol. 2B). New York: Academic Press.
Lamb, T. D. (1995). Photoreceptor spectral sensitivities: Common shape in the long-wavelength region. *Vision Research, 35*(22), 3083–3091.
Linhares, J. M. M., Pinto, P. D., & Nascimento, S. M. C. (2008). The number of discernible colors in natural scenes. *Journal of Optical Society of America, 25*(12), 2918–2924.
MacLeod, D. I., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of Optical Society of America, 69*(8), 1183–1186.
Mansfield, R. J. W. (1985). Primate photopigments and cone mechanisms. In A. Fein & J. S. Levine (Eds.), *The visual system* (pp. 89–106). New York: Alan R. Liss.
Merbs, S. L., & Nathans, J. (1992). Absorption spectra of the hybrid pigments responsible for anomalous color vision. *Science, 258*(5081), 464–466.
Miller, S. S. (1972). Psychophysical estimates of visual pigment densities in red–green dichromats. *Journal of Physiology, 223*, 89–107.
Mollon, J. D. (1982). A taxonomy of tritanopias. *Documenta Ophthalmologica Proceedings Series, 33*, 87–101.
Mollon, J. D., & Estévez, O. (1988). Tyndall's paradox of hue discrimination. *Journal of the Optical Society of America A. Optics and Image Science, 5*(1), 151–159.
Mollon, J. D., & Regan, B. C. (2001). Simulating the perceptual world of the anomalous trichromat. *Investigative Ophthalmology and Visual Science, 42*(4), S97.

Mullen, K. T., & Kingdom, F. A. (1996). Losses in peripheral colour sensitivity predicted from "hit and miss" post-receptoral cone connections. *Vision Research, 36*(13), 1995–2000.

Nascimento, S. M., Ferreira, F. P., & Foster, D. H. (2002). Statistics of spatial cone-excitation ratios in natural scenes. *Journal of the Optical Society of America A. Optics and Image Science, 19*(8), 1484–1490.

Neitz, J., Neitz, M., He, J. C., & Shevell, S. K. (1999). Trichromatic color vision with only two spectrally distinct photopigments. *Nature Neuroscience, 2*(10), 884–888.

Penn, J. S., & Williams, T. P. (1986). Photostasis: Regulation of daily photon-catch by rat retinas in response to various cyclic illuminances. *Experimental Eye Research, 43*(6), 915–928.

Regan, B. C., Julliot, C., Simmen, B., Viénot, F., Charles-Dominique, P., & Mollon, J. D. (1998). Frugivory and colour vision in *Alouatta seniculus*, a trichromatic platyrrhine monkey. *Vision Research, 38*(21), 3321–3327.

Regan, B. C., Reffin, J. P., & Mollon, J. D. (1994). Luminance noise and the rapid determination of discrimination ellipses in colour deficiency. *Vision Research, 34*(10), 1279–1299.

Renner, A. B., Knau, H., Neitz, M., Neitz, J., & Werner, J. S. (2004). Photopigment optical density of the human foveola and a paradoxical senescent increase outside the fovea. *Visual Neuroscience, 21*(6), 827–834.

Roorda, A., Metha, A. B., Lennie, P., & Williams, D. R. (2001). Packing arrangement of the three cone classes in primate retina. *Vision Research, 41*(10–11), 1291–1306.

Rushton, W. A., & Baker, H. D. (1964). Red–green sensitivity in normal vision. *Vision Research, 4*(1), 75–85.

Sanocki, E., Teller, D. Y., & Deeb, S. S. (1997). Rayleigh match ranges of red/green color-deficient observers: Psychophysical and molecular studies. *Vision Research, 37*(14), 1897–1907.

Shevell, S. K., & He, J. C. (1997). The visual photopigments of simple deuteranomalous trichromats inferred from color matching. *Vision Research, 37*(9), 1115–1127.

Smith, V. C., & Pokorny, J. (1973). Psychophysical estimates of optical density in human cones. *Vision Research, 13*(6), 1199–1202.

Smith, V. C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research, 15*(2), 161–171.

Stockman, A., & Sharpe, L. T. (2000). Spectral sensitivities of the middle- and long-wavelength sensitive cones derived from measurements in observers of known genotype. *Vision Research, 40*, 1711–1737.

Stockman, A., Sharpe, L. T., & Fach, C. (1999). The spectral sensitivity of the human short-wavelength sensitive cones derived from thresholds and color matches. *Vision Research, 39*(17), 2901–2927.

Thomas, P. B. (2005). *Variation in colour vision: The role of cone spectral sensitivity*. Ph.D. Cambridge: University of Cambridge.

Thomas, P. B., & Mollon, J. D. (2004). Modelling the Rayleigh match. *Visual Neuroscience, 21*(3), 477–482.

Tyndall, E. P. T. (1933). Chromaticity sensibility to wave-length difference as a function of purity. *Journal of Optical Society of America, 23*(1), 15–23.

Vierling, F. (1935). *Die Farbensinnprüfung bei der Deutschen Reichsbahn*. Melsungen: Verlag Bernecker.

Weale, R. A. (1953). Spectral sensitivity and wave-length discrimination of the peripheral retina. *Journal of Physiology, 119*(2–3), 170–190.

Williams, A. J., Hunt, D. M., Bowmaker, J. K., & Mollon, J. D. (1992). The polymorphic photopigments of the marmoset: Spectral tuning and genetic basis. *EMBO Journal, 11*(6), 2039–2045.

Wyszecki, G., & Stiles, W. S. (1982). *Color science*. New York: Wiley.