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 pseudoisochromatic 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).

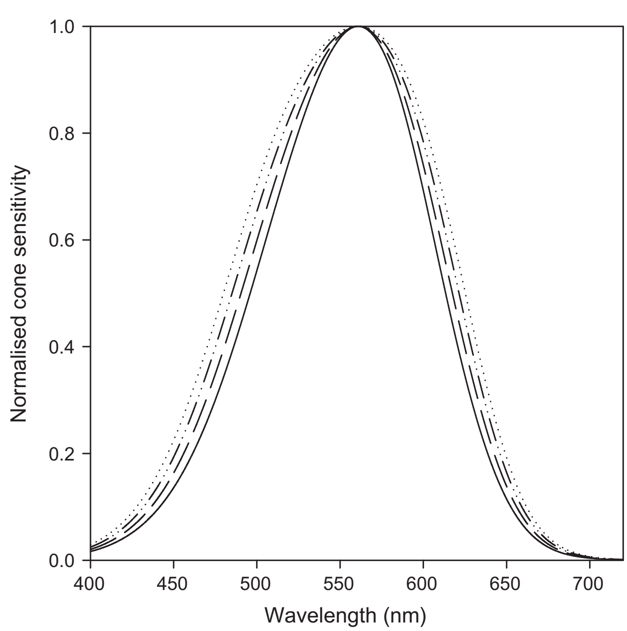

Fig. 1. The effect of photopigment optical density on the spectral sensitivity of a cone. Here, the sensitivity of a cone expressing photopigment at 561 nm is shown at OD of 0.1 (solid line), 0.3, 0.5 and 0.7 (from innermost to outermost). Note that the spectra here are normalized to their maxima. Without 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.