Color displays for the color blind

Jan Walraven & Johan W. Alferdinck
TNO Human Factors Research Institute
Soesterberg, The Netherlands

Abstract

Color is a powerful medium for coding, structuring and emphasizing visual information and, as such, used in many computer applications. However, this tool is less effective, or even counterproductive, in the case of people with impaired color vision. This problem can be remedied to a reasonable extent, provided the display designer is able to anticipate the chromatic trouble spots of a particular color palette. For that purpose, a color editor was designed that allows an image to be displayed as if viewed through the eyes of a color-deficient observer. The model used for computing the color transformations, makes use of state-of-the-art knowledge concerning the polymorphism of human cone pigment and the spectral filtering of the eye lens and macular pigment. As a result, the color editor not only enables the emulation of dichromatic color vision, but also of anomalous trichromatism, the more complex, but also more frequently occurring form of deficient color vision (75% of the colorblind population). In addition to its use as a diagnostic design tool, the editor also provides the means for adjusting the color look-up table to the individual needs of a color-deficient display user.

Introduction

A significant proportion of the population views the world through eyes that are even more colorblind than the normal trichromatic eye. The estimates vary for different parts of the world, but for the Caucasian population the numbers converge on about 8% for the male and 0.5% for the female population. The problems that may arise because of impaired color vision are most prominent when color is used for information coding, as in traffic signaling and electric wiring. Due to the widespread use of the computer, the color display is now also taking its toll. An increasing number of color-deficient people is confronted by their chromatic handicap, in particular when working with color coded man-machine interfaces, like control panels in the process industry and military applications.

It is now becoming common practice that applicants for functions that involve working with color displays, are screened for deficient color vision. This may be unavoidable in applications that heavily depend on correct color recognition (like red is ‘danger’, and green is ‘safe’). However, for the majority of color display users it is not the *identification,* but rather the *discrimination* of color that is all that matters.

In order to provide designers and users with a tool that allows them to taylor the color display to specific task requirements and/or individual needs, we developed a color editor[1], the TNO color editor, with built-in colorimetric tools, including an algorithm for transforming colors from normal to deficient color space. It is thus possible to view the color display with the same perceptual constraints as experienced by a color-deficient observer, and hence, to identify potential trouble spots in the design. The editor can be used for both the diagnosis and the optimal solution for chromatic discrimination problems.

Defective color vision

The various congenital forms of defective color vision (which is what we are addressing here) can be attributed to either a complete lack, or a modification of one of the three classes of cone pigments[2]. The deficiency of the pigment can reside in either the L-, M-, or S-cones (often still called the ‘red’, ‘green’ and ‘blue’ cones), the photoreceptors sensitive to long (L), middle (M) or short (S) wavelengths of the spectrum. The corresponding types of color vision defects are indicated by protan, deutan and tritan, respectively. These classes can be subdivided into three *dichromatic* forms, protanopia, deuteranopia and tritanopia, and three *anomalous trichromatic* forms, protanomaly , deuteranomaly and tritanomaly. The (blue-deficient) tritans are very rare; estimates vary from 0.005 to 0.1%, which should be compared to the 8% of the male and 0.5% of the female population with a protan or deutan disorder.

About 75% of the incidence of impaired color vision can be accounted for by the anomalous trichromats. These people deserve special attention, therefore, not only because of their numbers, but also because they are most likely to benefit from an improved color design. The dichromats, with their single (blue-yellow) discrimination-axis may well represent a too problematic group. Still, here too it is always possible to strife for an optimal color design, of course[3].

Traditionally, it has always been assumed that the anomalous pigments are different from the normal pigments, but there is now quite some evidence indicating that the pigments of both the color defective and the normal eye draw from the same two sets of M and L pigments[4]. According to this view, normal color vision would only differ from anomalous color vision in that one or more pigments are drawn from *two* different pigment clusters (L and M), whereas anomalous vision is subserved by two or more pigments drawn from a *single* cluster (either M or L). Due to the progress made in the molecular genetics of photo pigments, we know now that there may be quite a variety of L and M pigments[4,5,6,7]. Normal color vision seems to be subserved by a single M pigment, peaking at λ_{max} = 530 nm, and two different L pigments, peaking at about 552 or 557 nm (in vitro)[8], which are probably distributed about equally over the population. The anomalous pigments may result from gene recombinations of the “primary” L and M pigments, producing at least 9 different hybrid pigments with intermediate values of λ_{max}, maximally shifted by about 6 nm to either shorter wavelenghts (for the L pigment), or to longer wavelengths (for the M pigment)[9]. Still, most of that variability can be explained by just two major members of the L and M pigment clusters[9,10], which, following the nomenclature of DeMarco et al.[11], will here be designated as L' and M'.

Although anomalous color vision results from the loss of one class (cluster) of photo pigments, it is nevertheless trichromatic because *two* pigments are available from the single remaining M or L cluster. However, since the two pigments are drawn from the same cluster, their absorption functions will be nearly overlapping, so color discrimination will be relatively poor. In the case of dichromatic vision, there is only *one* pigment available, either from the L or M cluster, resulting in a still further detoriation of color discrimination.

Simulating deficient color vision

In as far as color deficiencies can be localized at the level of the cone pigments, it is possible to compute the attending reduction in color space and color discrimination, for any given set of cone spectral sensitivities. Making certain assumptions, one can also *visualize* the perceptual constraints of a given type of color deficiency on a color monitor, provided that the relevant display variables that determine the output of the RGB primaries are known. That is, the (calibrated) gamma functions (DAC value *vs*. radiance) and the CIE x,y coordinates, or better still, the spectral power distributions of the display primaries.

The basic set of equations, relating the physical stimulus to the visual stimulus, is given by

$$L = k\int l(\lambda)L_e(\lambda) \quad d\lambda$$
$$M = k\int m(\lambda)L_e(\lambda) \quad d\lambda$$
$$S = k\int s(\lambda)L_e(\lambda) \quad d\lambda$$

where l(λ), m(λ) and s(λ) are the absorption spectra of the L-, M- and S-cones, $L_e(\lambda)$ is the radiance of the display (in $W.m^{-2}.sr^{-1}.nm^{-1}$), and k is a proportionality constant. In analogy to the CIE XYZ system of color specification, the three quantities L, M and S, are designated here as *cone tristimulus values.*

Without going into the intricacies and pitfalls of display calibration[12], it is assumed that the radiance of the display can be described as the sum of the three separate contributions of the RGB primaries (the display tristimulus values). This implies that LMS can be related to RGB through the matrix transformation

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = k\,[A] \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

It is common practice to derive matrix A *indirectly*, that is, by employing the CIE XYZ system as the colorimetric link that connects with both RGB and LMS space[13,14]. However, the link between LMS and XYZ is not perfect, necessitating the Judd-Vos colorimetric modification[14] (through X'Y'Z'). Matters become even more complicated for transformations involving anomalous pigments[11], so we decided to circumvent the CIE system altogether and compute the matrix *directly*, thereby going through the matrix multiplications of (normalized) RGB emission and LMS absorption spectra. In this respect, our approach differs from other studies[3,16] that, being limited to dichromatic vision (which is based on ‘normal’ cone pigments), face less complications when using the XYZ system.

The inverse of matrix A, matrix A^{-1}, is used for the transformation from the LMS to the RGB domain, as is required for visualizing the effect of changes in the LMS tristimulus values. Here too, the transformation does not involve the use of the XYZ system.

The anomalous cone tristimulus values are also computed with eq.(1), simply by replacing the normal cone absorption spectra by their deviant counterparts l'(λ), m'(λ) or s'(λ). Assuming that the ‘wiring’ of the cone systems is the same for normal and deficient color vision, the resultant cone tristimulus values, denoted as L' M' and S', can be represented in normal (cone) color space and, going through the inverse matrix A^{-1}, also be expressed in RGB coordinates. It is this very assumption that is at the basis of our simulation.

Cone fundamentals and deficient color vision

Dichromacy

For the case of dichromatic color vision we assume, in line with general consensus, that there are only two photo pigments involved, the S-pigment in combination with either an L or M-pigment. The λ_{max} of the L- and M-pigment may vary[17], but values of 530 nm for the M-pigment of the protanope, and 560 nm for the L-pigment of the deuteranope, may be considered as reasonable averages[17,21]. In addition, we assume replacement of cone pigment, rather than the loss of an entire receptor system.

Not much is known about the tritanope, but in view of the different evolutionary origin of the blue cones, application of the replacement model may here be somewhat more speculative than for the other cone types. However, this is actually an academic question, because the blue cones are so scarce, probably in the order of a few percent of the total cone population[18], that it does not make much of a difference whether they are replaced or not.

Anomalous trichromacy

Deriving spectral sensitivity functions for anomalous trichromats is not a simple affair[11]. One has to decide first on how to decompose the cone spectral sensitivity functions, the so-called fundamentals, into a component due to the cone pigment proper, and the separate spectral modifications introduced by the lens and the macular pigment[21].

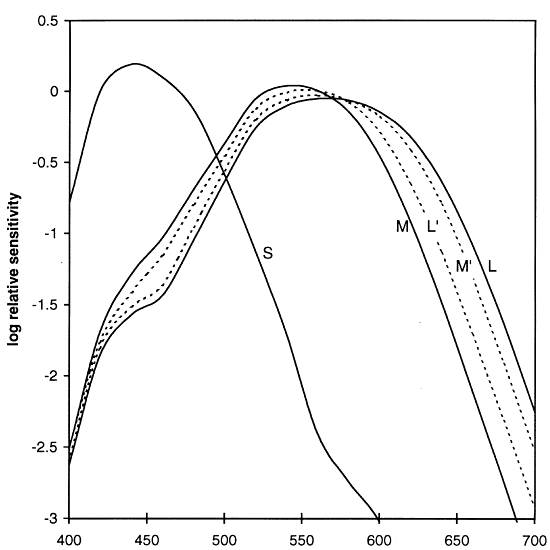

Figure 1. Smith-Pokorny cone fundamentals[22] normalized for equal sensitivity at equal-energy white[14]. The dashed curves represent the cone fundamentals derived for anomalous L'- and M'- cone pigments[10].

The general principle, which is also implemented in our model, is to employ a pigment absorbance template and assume shape invariance[1,19] when the template is shifted over a normalized frequency scale[20] ($\nu/\nu_{max} = \lambda_{max}/\lambda$). The pigment template will ultimately be based upon the recent standardization efforts by the technical committee TC 1-36 of the CIE[21], but this work is still in the drafting stage.

In the mean time we have opted for the Smith-Pokorny fundamentals[22], normalized for equal-energy white[14]. As for the spectral transmission of the eye media, we use the relative optical density spectra recommended by TC 1-36 of the CIE[21]. Althoughthe latter are also still in the drafting stage, the estimates for the lens[23] and macular pigment[24] may nevertheless be considered as representing the present state-of-the-art.

Following DeMarco et al.[11], the λ_{max} of the M-pigment was shifted to 553 (for the protanomalous pigment) and the λ_{max} of the L-pigment to 560 nm (for the deuteranomalous pigment). The simulation algorithm uses this information as a variable, so other estimates of the L-pigment shifts can be used as well. The results for the Smith-Pokorny fundamentals are shown in Figure 1.

Mapping color-deficient gamuts

We found it convenient to analyze our results in a fundamental chromaticity diagram based on LMS tristimulus values[25]. This diagram, in which the axes plot the relative values l = L/(L+M+S) and m = M/(L+M+S) —analogous to x and y in the CIE diagram—is shown in Figure 2.

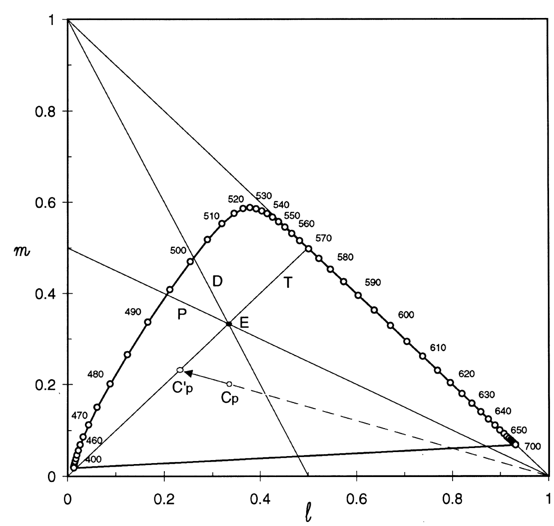

Figure 2. Chromaticity diagram based on l,m cone coordinates[25], showing protan (P), deutan (D) and tritan (T) confusion lines for equal-energy white. The tritan confusion line also represents the locus onto which all chromaticities for both protanopes and deuteranopes project, as is illustrated for the protanope ($c_P \rightarrow c_p'$).

Figure 2 shows that, for the *dichromatic* eye (protan, deutan and tritan), the chromaticity locus for white (E) is not represented by a point, but a line. All points on such a confusion line are perceived as having the same chromaticity (in this case white).

Note that the tritan confusion line (T) conforms to l=m, and since the same holds true for any color seen by protanopes(L'=M) and deuteranopes (M'=L), that line also represents the locus of all possible chromaticities that can be seen by the protan and deutan group. So, simulating a particular chromaticity for the red-green dichromat, simply amounts to projecting it onto the tritan confusion line, along trajectories through the protan (l=1, m=0) or deutan (l=0, m=1) confusion centers. In Figure 2 this is illustrated for the protan (transformation of c_p to c_p'.)
The tritan chromaticity locus is also a line, but its location in the l,m diagram is not as clear-cut as it is for the other dichromats. If one assumes that there is no 'refilling' of the S-cones with L- and/or M-pigment, then the chromaticity locus is represented by the hypothenuse (l+m=1) of the l,m diagram. Alternatively, in the case of a fifty-fifty replacement with L- and M-pigment, one obtains a line parallel through the hypothenuse, and passing through the white-point (l+m=2/3). Speculating that dichromats have the same code for white as normals (l=m=s)[14], we simulate the tritan's color world according to the latter option, but this is based on intuition rather than facts.

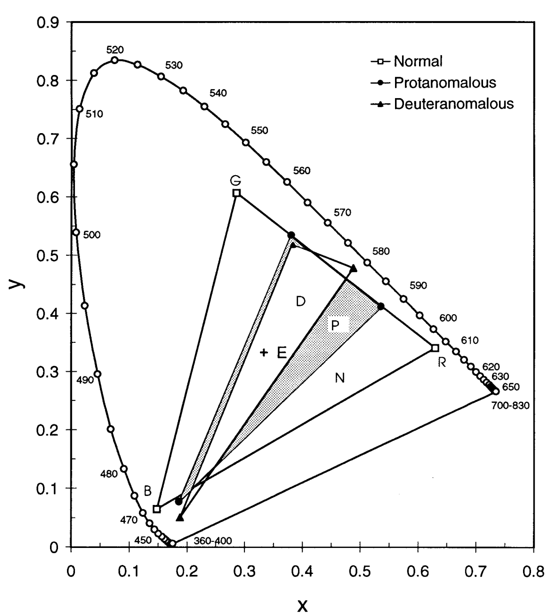

Figure 3. The RGB chromaticity space of a monitor, for a normal (N), a protanomalous (P) and a deuteranomalous (D) observer.

Results from simulations of the perceptual constraints of anomalous color vision are shown in Fig. 3, which illustrates how the RGB gamut shrinks after transformation from normal to anomalous color vision.
We chose here for a representation in the CIE x,y chromaticity diagram, because this provides a more familiar frame of reference. The CIE x,y coordinates were derived by applying a matrix transformation that connects our RGB (monitor) space to XYZ space. The RGB inputs are obtained from the LMS→RGB transformation (through the inverse matrix A^{-1}).

It is of interest that, in spite of the fact that protanomalous vision is characterized by a reduction in senstivity for red light, the available chromaticity space in the red region of the x,y diagram compares favorably to that of the deuteranomalous observer. This can be explained by the relatively large overlap of the deuteranomalous fundamentals (M' and L) in the red region of the spectrum (see Figure 1). Of course, the protanomalous observer is still faced with the penalty of insensitivity for red light, an aspect that is lost in the x,y representation.

In Figure 3 it can also be seen that, after the deuteranomalous transformation, some colors fall slightly outside the boundaries of the RGB chromaticity space (saturated yellows and bluish purples). For the purpose of simulation this poses only a small problem that can be dealt with by using a gamut mapping procedure[26]. Alternatively, one can limit the color gamut to colors that, after the transformation, stay within the RGB chromaticity space.

Spotting the chromatic discrimination problems of the color-deficient observer

To evaluate the limitations of defective color vision, the most straightforward way is to compute color differences (_E) in a perceptually uniform color space. The ability to do so is incorporated in the 'coldef' (color deficiency) menu of our color editor. It computes, in whatever color space, the _E values of all possible color combinations within a given set of colors. An example, for the protanomalous transformation, is given in Table 1.

Table 1. Color distances, _E$^{*}_{uv}$, for a set of equi-luminant (12 cd m^{-2}) colors, respectively for normal (shaded cells) and protanomalous color vision. Bold printed numbers indicate values for which _E$^{*}_{uv}$ ≤ 30. The chromaticity coordinates (x,y) of the colors are shown in the first row.

Colors	White x = 0.333 y = 0.333	Red x = 0.371 y = 0.363	Green x = 0.352 y = 0.404	Blue x = 0.279 y = 0.231	Yellow x = 0.402 y = 0.469	Magenta x = 0.304 y = 0.260	Cyan x = 0.304 y = 0.283
White	x	56.7	42.8	49.4	53.9	41.4	44.0
Red	***15.7***	x	83.1	85.5	59.7	57.5	100.2
Green	*34.8*	*37.1*	x	86.8	40.1	84.2	55.7
Blue	*50.0*	*58.7*	*83.2*	x	103.1	30.1	47.4
Yellow	*49.9*	*46.5*	***19.6***	*99.8*	x	88.2	87.8
Magenta	*39.9*	*46.9*	*74.3*	***12.7***	*34.1*	x	64.5
Cyan	***22.8***	*37.4*	*50.7*	*34.1*	*68.8*	***28.9***	x

Table 1 shows the color differences for a set of 7 equi-luminant colors (12 cdm^{-2}), that are perceptually equi-distant from the white point (Munsell 5/6 locus). The _E values (in CIE LUV space) for this set were computed before and after transformation to anomalous color space, respectively. Taking a certain criterion value, it is thus possible to detect unacceptable reductions in _E$^{*}_{uv}$. This is indicated by the bold entries in Table 1, which relate to a color difference of _E$^{*}_{uv}$ ≤ 30. This is not an exessive criterion; it is about in the middle of the range of _E values recommended in human factors standards on the use of color in displays[30].

On the basis of the type of information shown in Table 1, the color table of a display design can be adjusted in order to accomdate color-deficient users. To aid the designer in that task, the TNO color editor[1] is equipped with an algorithm for calculating a default palette, that yields _E values that all meet the desired criterion. These calculations can be based on either the actual gamma functions and emission spectra (or chromaticities) of the RGB primaries of the display in question, or on assumptions regarding a standard monitor profile.

Prospects

That color displays could be better adapted to the needs of users with impaired color vision has already been recognized before, in particular in relation to the design of (electronic) topographic maps[1,27]. The TNO color editor[1], which has been first demonstrated at the CeBIT exposition in Hannover (March, 1995), provides the tool for attacking this problem, but still awaits commercial exploitation.

Actually, we do not expect that designers will be the main user group, simply because there are not too many. Moreover, the majority seems to lack the necessary background in color perception, and hence, is unaware of how the indiscriminate use of color can completely wreck an otherwise good display design[28]. Anybody who is familiar with the color displays used in industrial applications will have noted the highly saturated colors on black backgrounds, resulting in displays that have little resistance against the effects of ambient illumination, contain illegible colored text/background combinations, and are prone to perceptual artefacts[29].

It is more likely, therefore, that the color-deficient user rather then the display designer will become the main target group for our color editor. With that in mind, we have also developed an expert system that administers an on-screen color vision (self)test, and then uses the test results for the automatic adjustment of the color look-up table to the individual needs of the display user.

References

1. J.Walraven and J. Varkevisser, *Report TNO-TM 1995* ***A-68***, TNO Human Factors Research Institute, Soesterberg, (1995).
2. J. Pokorny, V.C. Smith, G. Verriest and A.J.L.G. Pinckers, *Congenital and acquired color vision defects*. Grume and Straton, New York, 183 (1979).
3. G.R. Meyer and D.P. Greenberg, *IEEE Comp.Graphics & Applns* ***8***, 28 (1988).
4. M. Neitz, J. Neitz and G.H. Jacobs, *Vision Res.* ***33***, 117 (1993).
5. J. Nathans, D. Thomas and D.S. Hogness, *Science* ***232***, 193 (1986).
6. J. Nathans, T.P. Piantanida, R.L. Eddy, T.B. Shows and D.S. Hogness, *Science* ***232***, 203 (1986).

 Copyright 1997, IS&T

7. M. Neitz, J. Neitz and G.H. Jacobs, *Science* ***252***, 971 (1991).
8. J. Neitz and G.H. Jacobs, *Vision Res.* ***30***, 620 (1990).
9. S.L. Merbs and J. Nathans, *Science* ***258***, 464 (1992).
10. M. Neitz, J. Neitz and G.H. Jacobs, *Vision Res.* ***33***, 117 (1993).
11. P. DeMarco, J. Pokorny and V.C. Smith, *J.Opt.Soc.Am.* ***A9***, *1465* (1992).
12. D.L. Post, *Color in electronic displays* (eds. H. Widdel and D.L. Post), Plenum Press, New York, 299 (1992).
13. A. Stockman, D.I.A. MacLeod and N.E. Johnson, J.Opt. Soc.Am ***A 10***, 2491 (1993).
14. J. Walraven and J.S. Werner, *Vision Res.* ***31***, 2185 (1991).
15. J.J. Vos, *Color Res.& Appln.* ***3***, 125 (1978).
16. F. Viénot and J.D. Mollon, *Nature* ***376***, 127 (1995).
17. E. Sanocki, D.T. Lindsey, J. Winderinx and D.Y. Teller, Vision Res. ***33***, 2139 (1993).
18. J.D. Mollon and J.K. Bowmaker, *Nature* ***360***, 677 (1992).
19. J. Pokorny and V.C. Smith, J.Opt.Soc.Am ***67***, 1196 (1977).
20. R.J.W. Mansfield, *The visual system* (eds. A. Fein and J.S Levine), Liss, New York, 89 (1985).
21. CIE, *Draft 8 TC 1-36*, CIE, Vienna, Austria (1995).
22. V.C. Smith and J. Pokorny, *Vision Res.* ***15***, 161 (1975).
23. D. van Norren and J.J. Vos, *Vision Res.* ***15***, 749 (1974).
24. J.J. Vos, *Report TNO-IZF* ***1972-2***. TNO Human Factors Research Institute, Soesterberg (1972).
25. J. Walraven and T. Benzschawel, *Color in electronic displays* (eds. H. Widdel and D.L. Post), Plenum Press, New York, 87(1992).
26. L.W. MacDonald, *IS&T and SID's Color Imaging conference*: *Transforms & Transportability of Color*, 193 (1993).
27. J.M. Olson and C.A. Brewer, *Annals Ass. Amer. Geographers* ***87***,103 (1997).
28. J. Walraven, *Color in electronic displays* (eds. H. Widdel and D.L. Post), Plenum Press, New York, 3 (1992).
29. J. Walraven, *Displays* ***6***, 35 (1985).
30. J. Walraven, P.M. van Bergem-Jansen and J.W.A.M. Alfer-dinck, *Report TNO-TM 1995* ***A-68***, TNO Human Factors Research Institute, Soesterberg, (1995).