Judgments about the intensity of the illumination are influenced by the association between colour and luminance in the scene

Eli Brenner; Faculty of Human Movement Sciences, VU University; Amsterdam, The Netherlands

Sérgio Nascimento; Centre of Physics, Gualtar Campus, University of Minho; 4710-057 Braga, Portugal

Abstract

In order to judge whether a surface that one is looking at is white or grey, one needs to consider the intensity of the illumination. We here show that people do not simply use the maximal luminance in the light from the scene as a measure for the intensity of the illumination but also consider how luminance and chromaticity are associated. We suggest that they take into account that there are physical limitations to the luminance that reflecting surfaces can achieve at high chromatic saturation. These limitations arise because chromaticity is the result of surfaces selectively reflecting light of different wavelengths, so that the luminance of the illumination must be higher than that of the brightest patch in the scene if that patch is not white.

Introduction

In daily life we are normally more interested in judging surface reflectance than in judging the intensity of the illumination. For judging chromaticity and saturation an emphasis on the ratio of stimulation of different kinds of cones helps remove influences of the intensity of the illumination [1][2]. Contrasts between adjacent surfaces are also insensitive to changes in the level of illumination, but for judging a surface's lightness we need to somehow consider the intensity of the illumination. How could we otherwise tell whether the walls of our room are white and the illumination is dim, or the walls grey and the illumination bright? In simple displays we cannot: the brightest surface is simply perceived as white (e.g. [3]). However, in complex scenes it is not always that simple [4][5].

We here examine one potential factor in retrieving the intensity of the illumination: the chromaticity of the brightest surfaces. Previous studies have shown that the correlation between luminance and chromaticity is used to account for the colour of the illumination [6][7], the reasoning being that surfaces that reflect the colour of the illumination well are likely to reflect a lot of light. At the extreme, white objects, if present, would reflect all of the light falling on them, and would therefore not only have the highest luminance but the light they reflect would provide a direct estimate of the illumination at that position. We here examine a similar issue in judging the intensity of the illumination.

A coloured surface is one that selectively reflects certain parts of the spectrum whereas a white surface is one that reflects as much light as possible. Since reflection can only reduce the intensity of the light at each wavelength, very bright surfaces are unlikely to have very saturated colours, because for any given illumination there is a physical limit to the combinations of luminance and chromaticity that can arise by diffuse reflection alone. The colours on this limit form a closed surface in colour space and are normally referred to as the optimal colours [8].

The constraints on the possible relationships between luminance and chromaticity should influence our judgments. If the highest luminance in the light reflected from a scene is from a surface that is clearly white, the luminance of the light reflected from that surface provides a reasonable estimate of the intensity of the illumination. However, if the highest luminance from the scene is from a surface that is, for instance, clearly blue, the intensity of the illumination must be higher than the luminance of the light reflected from that surface, because the surface does not reflect all the light falling on it at longer wavelengths (which is why it is blue). In other words, the maximal perceived saturation for purely reflecting surfaces can only be obtained for middle lightness, so the luminance of saturated coloured surfaces in a complex scene can be used to judge the level of illumination even when such surfaces do not have the highest luminance in the scene, and reliable estimates can be obtained even when there is no truly white surface available.

We test whether the relationship between luminance and chromaticity influences our judgments of surface lightness by determining the luminance at which subjects report a transition between grey and white for scenes with identical distributions of luminance and chromaticity, but different relationships between luminance and chromaticity. We find differences, showing that subjects do consider this factor.

Methods

Observers sat at a distance of 183 cm from a 48 by 31 cm screen (Sony GDM-FW900 Trinitron CRT; 1280 by 1024 pixels; 100 Hz; 8 bits per gun; calibrated with a Minolta CS-100A Chroma Meter). The screen was filled with a regularly tiled background of 12 by 8 squares, each with sides of about 1 deg. On each trial, one of three sets of specially selected patterns of colour and luminance was assigned to the squares. We will refer to the selected set as the pattern, but note that the images varied across instances of the same pattern because the colours and luminances were assigned to the individual squares at random on each trial.

A target square that was perceived to be achromatic (CIExy = [0.291, 0.328]; variable luminance) and that had the same dimensions as the tiles in the background was superimposed on the tiles 2s after the new colours and luminances were assigned to them. Subjects had to indicate whether the target square was grey or white by pressing the 'g' or 'w' key of a computer keyboard. A separate staircase procedure was used to estimate the luminance at which subjects' judgments switched from grey to white for each of three kinds of patterns (the patterns are described in the next
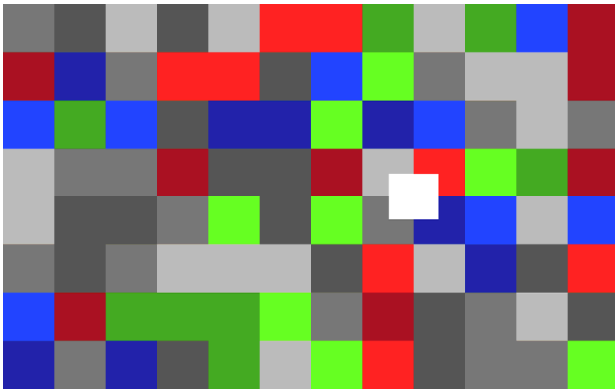
Figure 1. A possible image shown to a subject in the test condition. The target is superimposed on a pattern with both bright and dark red, green and blue squares (8 each) and squares with three grey levels (16 each).

paragraph). The room was dark except for the light from the screen.

In order to understand how the luminances and chromaticities of our three patterns were designed it is easiest to start with the pattern in which there were coloured squares of maximal luminance (test condition; Figure 1). For the 24 squares of maximal luminance, one gun was set to maximum and the other two to zero (8 squares for each of the three possible combinations). The CIExyY coordinates of the three colours were [0.622, 0.347, 18.24 cd/m^2], [0.287, 0.595, 63.99 cd/m^2] and [0.144, 0.069, 8.68 cd/m^2], for red, green and blue squares, respectively. Another 24 squares had the same colours, but 1/3 of the luminance (again 8 per colour). The remaining 48 squares were grey (CIExy = [0.291, 0.328]) and had a luminance of 2/3 of the maximal luminance of one of the three colours (16 each).

The second pattern (baseline condition) was matched to the first pattern both in luminance and chromaticity. This was achieved by switching between coloured and grey squares without changing the luminance. Thus there were squares with six grey levels (matching the coloured squares with maximum and 1/3 of the maximum luminance) and three kinds of coloured squares (each with 2/3 of the maximal luminance of the gun in question). In this way, both patterns had exactly the same distribution of luminance values and both had 16 red, 16 green, 16 blue and 48 grey squares. The third pattern (darker condition) was identical to the second, but the whole pattern was 5% darker (the luminance of each square was 95% of that in the baseline condition).

Thirteen subjects (including one of the authors) each took part in one session that was divided into 30 blocks of 10 trials: 10 blocks for each condition. The blocks of the three conditions were presented in random order. Each block started with a 10 second presentation of the pattern without the target, but with the text "short break' printed in black letters across the centre of the screen instead.

The target was centred at one of the central grid of 5 by 3 intersections of four squares. Separate staircases were used to sample the appropriate target luminance for each condition. Each staircase started at a luminance of 63.6 cd/m^2. If the subject pressed the 'w' key the target's luminance on the next presentation within the block (or on the first presentation of the next block for that condition) was decreased by 7%. If the subject pressed the 'g' key the luminance was increased accordingly.

The main question was whether subjects would assume that the illumination was more intense in the test condition than in the baseline condition. In order to make sure that our method works, which would be especially relevant if we were to find no difference in the previous comparison, we also compared the baseline condition with the darker condition.

For each subject and condition we determined the border between what he or she considered grey and what he or she considered white. To do so, we took all the data of the subject and condition in question and fit a cumulative normal distribution to the proportion of 'white' responses as a function of luminance. The fit was done more or less as proposed by Wichmann and Hill [9], but we removed the 2% least likely responses from each set to account for stimulus-independent errors, rather than fitting an additional parameter. The mean of the fit distribution indicates at what target luminance the transition from grey to white occurs.

We evaluated whether subjects assumed that the illumination was more intense in the test condition than in the baseline condition by testing whether the transition was at a higher value in the former. We evaluated whether our method works by testing whether the transition was at a higher value for the baseline condition than for the darker condition. Both comparisons were done with paired, one-tailed t-tests.

Results

Figure 2 shows an example of one subject's data with the fit cumulative normal distributions. The curve for the test condition is clearly shifted to the right with respect to the curve for the baseline condition (by about 4.5 cd/m^2), indicating that for this subject a higher target luminance was required for the target to appear to be white in the test condition. This is consistent with assuming that the illumination is more intense in the test condition. Similarly, the curve for the baseline
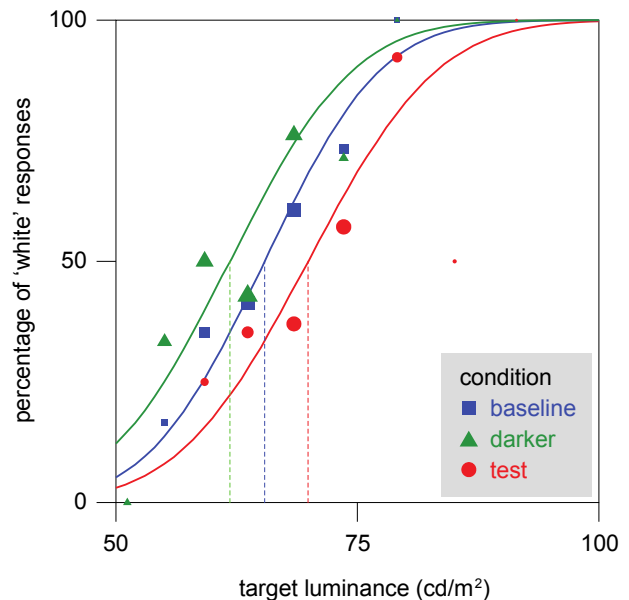

Figure 2. One subject's data. The symbols show the fraction of presentations at each target luminance for which the subject pressed the 'w' key. The size of the symbols is proportional to the square root of the number of presentations. The curves are fit cumulative normal distributions. The dashed vertical lines indicate this subject's transition points between grey and white. This is the luminance at which she is just as likely to press the 'g' key as to press the 'w' key.

condition is clearly to the right of that for the darker condition (by about 3.6 cd/m2), indicating (not too surprisingly) that a higher target luminance was required for the target to appear to be white when the pattern was brighter. The difference was approximately equivalent to the difference in luminance between the patterns (5%).

We used the means of the fit cumulative normal distributions (the target luminance values indicated by the dashed lines in Figure 2) for our further analysis. Differences between the means of the fit distributions indicate how much higher the luminance had to be for the target to look white to the subject, rather than grey, in some conditions than in others. We ignored one subject's data because although he showed the same trend as the others, the fit cumulative normal distributions were much shallower than those of all the other subjects (i.e. the fit standard deviations were larger) and he only considered targets to be white at a much higher luminance.

The left bar in Figure 3 shows how much higher the luminance had to be for subjects to consider the target to be white for the test condition than for the baseline condition. From these results it is evident that people consider the combinations of luminance and chromaticity rather than only each separately. The right bar in Figure 3 shows how much higher the luminance had to be for subjects to consider the target to be white for the baseline condition than for the darker condition. The results validate our method in that we found a mean difference that is not significantly different from the expected value of 5%.

Discussion

Our results appear to point at a high-level influence on perceived brightness, because the luminance distributions were identical for the test and baseline conditions, as were the chromatic distributions and the average stimulation of each cone type. However, the maximal stimulation of each kind of cone was not identical for the test and baseline conditions: it was 4.5% lower for the l-cone, 8.8% higher for the m-cone and 18.1% higher for the s-cone in the test condition than in the baseline condition (see [10] for the conversion from RGB-gun values to cone stimulation). The reason for this difference is that the highest l- and m-cone stimulations in the test condition are found when maximally and exclusively stimulating the green gun and the highest s-cone stimulation when maximally and exclusively stimulating the blue gun, whereas the highest stimulation for all three cones in the baseline condition is found for the grey squares with the same luminance as the maximum-intensity green ones in the test condition. Thus, if a target would only look white if it stimulated each kind of cone maximally, the luminance would have to be 18.1% higher in the test condition.

The effect that we see is smaller, but perhaps only the directly adjacent surfaces matter (see [7]; in the present study we did not consider the differently coloured squares' positions on individual trials). Another possibility is that the level is not only determined by the current pattern, but also to some extent by the pattern on previous trials, so that the effects have not yet stabilized. Moreover, this may not be the sole criterion. Of course, for the control comparison the maximal stimulation of each kind of cone is reduced by 5% when the overall luminance is reduced by 5%, and yet the average effect that we find is smaller, though not significantly so (see 95% confidence intervals in Figure 3).

Considering cone stimulation in relation to the maximal stimulation of the kind of cone in question when judging brightness is somewhat analogous to the use of the MaxRGB method for illuminant estimation [11]. The kind of mechanism that we are proposing might also explain why white surfaces look grey when the illumination is designed to artificially enhance the perceived chromaticity of a scene [12]; under such narrowband illumination white surfaces are not necessarily as much brighter than surfaces with very saturated colours as would normally be the case.

Whatever the mechanism, our study confirms that some physical regularity of natural scenes is considered in the way we interpret the light reaching our eyes, as indeed it should be if we want to be able to rely on vision to recognize objects by their surface properties despite varying illumination. In the above we only consider diffusely reflecting surfaces. Of course specular reflectance, such as the reflection of a lamp on a white mug, provides a much higher local luminance than that of the white mug, and yet the mug looks white. Thus, the above reasoning and findings are limited to diffusely reflecting surfaces. However, our findings reveal a new aspect of lightness perception, because we propose a new contribution to the way in which the lightness component of surface reflectance is scaled (or anchored [3][4]). The importance of the relationship between luminance and colour implies that it may be worthwhile to study brightness judgments in colourful
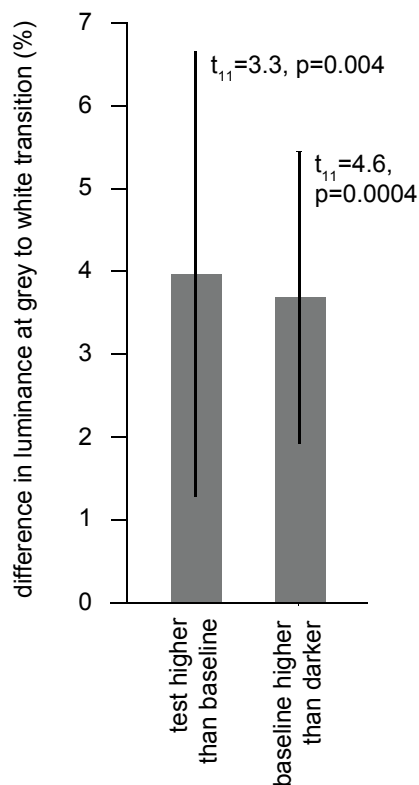

Figure 3. The two comparisons between conditions. Each bar shows how much higher the luminance had to be for the target to appear to be white in the conditions in which we expect to need a higher luminance for the target to appear to be white. The difference in luminance is expressed as a percentage to account for different overall values for different subjects and for direct comparison with the 5% difference in overall luminance between the darker and the baseline condition. For both comparisons the differences are significant. Error bars are 95% confidence intervals.

scenes, rather than isolating brightness from colour by studying the former in black and white images.

References

[1] Land EH, McCann JJ (1971) Lightness and retinex theory. Journal of the Optical Society of America 61, 1–11.

[2] Brenner E, Granzier JJ, Smeets JBJ (2007) Combining local and global contributions to perceived colour: an analysis of the variability in symmetric and asymmetric colour matching. Vision Research 47, 114-125.

[3] Gilchrist AL, Radonjiç A (2009) Anchoring of lightness values by relative luminance and relative area. J Vis 9(9):13.1-10.

[4] Gilchrist A, Kossyfidis C, Bonato F, Agostini T, Cataliotti J, Li X, Spehar B, Annan V, Economou E (1999) An anchoring theory of lightness perception. Psychol Rev 106, 795-834.

[5] Kingdom FA (2011) Lightness, brightness and transparency: a quarter century of new ideas, captivating demonstrations and unrelenting controversy. Vision Res 51, 652-673.

[6] Golz J, MacLeod DI (2002) Influence of scene statistics on colour constancy. Nature 415, 637-640.

[7] Granzier JJ, Brenner E, Cornelissen FW, Smeets JBJ (2005) Luminance-color correlation is not used to estimate the color of the illumination. J Vis 5(1):20-7.

[8] MacAdam DL (1935) Maximum Visual Efficiency of Colored Materials. Journal of the Optical Society of America 25, 361-367.

[9] Wichmann FA, Hill NJ (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys 63, 1293-1313.

[10] Granzier JJM, Brenner E, Smeets JBJ (2009) Can illumination estimates provide the basis for color constancy? Journal of Vision 9(3):18, 1-11.

[11] Funt B, Shi L (2010) The rehabilitation of maxrgb. Eighteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications, 256-259.

[12] Thornton WA (1974) Colour discrimination enhancement by the illuminant. Colour Vision Deficiencies II, Modern Problems in Ophthalmology 13, 312-313.

Author Biography

Eli Brenner graduated in Biology at Utrecht University and got a PhD from the University of Amsterdam. He was a Post-doc at Utrecht University, a visiting scientist at the Max-Planck-Institut für Biologische Kybernetik (Tübingen), and a computer-Vision Programmer at Ingenieursburo Beltech (Eindhoven), before becoming a lecturer in Physiology and Neuroscience at the Erasmus University (Rotterdam). Since 2006 he is at the Faculty of Human Movement Sciences of the VU in Amsterdam.

Sérgio Nascimento graduated in Physics by Porto University and did a PhD in Color Science in the Department of Communication and Neurosciences at Keele University, England. He is Associate Professor in Physics at Minho University where he teaches several courses in Physics, Optics and Vision Sciences. His research interests are colorimetry and colour vision, in particular, applications of multi and hiperspectral imaging, colour constancy and models of colour vision.