Human Assignment of Grey Levels Depends on Scene Complexity

David Connah, Romain Bachy
Centre for Visual Computing
University of Bradford
Bradford, UK +44 (0)1274 236258
d.connah@bradford.ac.uk,
romain.bachy@gmail.com

Graham D. Finlayson
Colour Group University of East Anglia Norwich, UK
+44 (0)1603 593347
G.Finlayson@uea.ac.uk

Marina Bloj
Centre for Visual Computing
University of Bradford Bradford, UK
+44 (0)1274 236258
m.bloj@bradford.ac.uk

ABSTRACT
We use a direct greyscale matching paradigm to examine the role of hue, brightness and luminance in the construction of greyscale images. Observers matched iso-luminant and iso-saturated, coloured patches by adjusting the grey-level of achromatic patches; they did this in three experimental set-ups: single-patch, two-patch, and multiple-patches. The latter were presented as cartoon-like images and abstract equivalent. Most observers assign similar grey level values to all colours in the single-patch case, but show a marked hue-dependency in two-patch and multiple-patch conditions (cartoon and abstract images). This indicates both that matches for our complex images could be predicted from two-patch settings, and that scene content does not play a role in determining grey settings. The hue-dependency of the Helmholtz Kohlrausch effect is not a strong predictor of performance for patches, but it might play a role in observer settings in more complex spatial configurations.

Categories and Subject Descriptors
I.3.3 [**Computer Graphics**]: Picture/Image generation

General Terms
Experimentation, Human Factors

Keywords
Colour-to-greyscale, brightness, luminance, Helmholtz-Kohlrausch.

1. INTRODUCTION
Each year millions of greyscale reproductions of colour images are made. The majority of these are produced by removing the chromatic information, which leaves a greyscale made by encoding the achromatic colour variable based on luminance (Y) or brightness (B).

While luminance is a photometric value (with standardized visual units - cd/m^2), brightness is defined as "[an] attribute of a visual sensation according to which an area appears to emit more or less light" [1]. At constant luminance, the brightness of a colour increases with its saturation; this phenomenon is known as the Helmholtz-Kohlrausch (H-K) effect, and has been reported to have a strong hue-dependency [1]. Although this effect is not seen in all observers [2], it has been incorporated into recent colour appearance models [3], as well as colour-to-greyscale algorithms [4], and is assumed to occur independently of the spatial layout of colours.

Predicting Perceptions: The 3rd International Conference on Appearance, 17-19 April, 2012, Edinburgh, UK. Conference Proceedings Publication ISBN: 978-1-4716-6869-2, Pages: 74-78

While encoding colour as luminance is the method used in most printers and photocopiers, one problem with this approach is how to make greyscales for images that contain iso-luminant edges and/or borders in a way that preserves the image content? In this paper we investigate this problem directly. Rather than looking for computational solutions (e.g. contrast-encoding methods such as [5]) we ask how human observers solve the problem.

To do this we instruct observers to match iso-luminant, iso-saturated coloured patches by adjusting grey patches in different stimulus configurations: single-patch images (see Figure 1), two-patch images (see Figure 2), complex cartoon images with more than 5 colours (Figure 3) and abstract versions of the latter (see Figure 4). We evaluate the raw luminance settings of observers and analyse the results by comparing them to predictions from luminance encoding, and other encodings that incorporate the hue-dependency of the H-K effect (shown in Figure 1).

2. MATERIALS AND METHODS
All experiments were performed in a darkened room with the observers seated 1 metre from the monitor generating the stimuli, with a fixed head position. All observers were screened for colour deficiencies and had normal, or corrected to normal, visual acuity. For the single- and two-patch studies we used 5 and 7 observers respectively. All the 5 observers that took part in the single-patch study also took part in the two-patch study. For the complex images we used a pool of 5 different observers.

In each condition a complex checkerboard background with randomly assigned luminance values (in the range 5 to 55 cd/m^2) surrounded the stimulus; this was done to avoid contrast or crispening effects. The observers were asked to adjust the greyscale patch (see Figure 2) to match the appearance of the coloured patch as closely as possible by only adjusting the grey level. In the experiments involving more than one patch (see Figures 3, 4 and 5), the observers were additionally asked to reproduce in grey levels the observed colour-differences between the patches. These instructions preclude observers from making a luminance match (or brightness match) when more than one patch is present. The grey patches were initially assigned a random luminance from the full displayable range of the monitor (approx. 0 to 80 cd/m^2). The observers then used a response box to adjust the grey-level with increments of 5 cd/m^2 (coarse adjustment) and 0.5 cd/m^2 (fine adjustment).

2.1 Preliminary experiments
A total of 16 different colours were used in the experiment. These were initially sampled from an iso-luminant, iso-saturation, hue-circle in the CIELuv colour-space (saturation = 0.05), then transformed to monitor-dependent RGB values via a calibration matrix (all resulting colours had a luminance of 20 cd/m^2). To realise a sensation luminance match [8] the observers then adjusted each colour to match a 20 cd/m^2 achromatic patch

in a flicker photometry (results were averaged over 3 settings). The patches were then displayed in a 4x4 grid and observers were allowed to adjust the saturation of each patch until no individual patch stood-out (results were averaged over 5 trials). As a result each observer had an individual dictionary of 16 colours that were iso-saturated and iso-luminant uniquely for them.

In the one-patch experiment observers made 6 adjustments for each of the 16 colours. In the two-patch experiment a subset of 8 colours were used so that each colour could be displayed next to each of the others without requiring observers to make too many comparisons; each comparison was repeated 6 times with an equal left-right balance. For the complex stimuli, and their abstract equivalents, we firstly coloured the cartoon images in colours that roughly corresponded to the natural appearance of the scene elements (e.g. the sky was blue, the sand yellow, and the grass green). For each colour we then found the closest match (in terms of hue-angle) from within the observers' colour dictionary, and replaced the original colour with its iso-luminant, iso-saturated equivalent. Each of the 5 observers matched 12 images (6 original, and 6 abstract).

3. ANALYSIS AND RESULTS

3.1 Data correction

In each experiment we record the luminance of the grey level set to match a given colour patch. These luminance settings are then corrected by subtracting the luminance of the colour patch that has been matched; this is done to take into account the varying sensation-luminance settings of different observers. In these shifted-luminance units a value of zero indicates that the luminance (cd/m^2) of the grey patch matches that of the colour patch.

3.2 Model fitting

We test the null hypothesis that observers assign grey levels equally to all hues, versus the hypothesis that different hues are assigned different grey-levels. To do this we test the fit of a hue-dependent sinusoidal model:

$$L = a_1 + a_2 sin(x) + a_3 cos(x), \quad \textbf{(Equation 1)}$$

against the null hypothesis that $L = a_1$, where L is the observer luminance setting, x is the hue angle and parameters a_1, a_2 and a_3 are determined by linear regression [7]. Note that we choose a sinusoid since the resulting function has to be periodic (as hue is angular) and we found that adding higher frequency terms did not provide any improvement in data-fitting.

Using regression diagnostics we can test whether including hue-varying terms gives a significant improvement to the model. In addition to linear regression we also use an ANOVA analysis to test the null-hypothesis that all means are equal (which is equivalent to testing whether $L = a_1$ provides a good fit).

For the complex stimuli, as well as analysing the raw luminance values, we recast the experiment as a pair comparison. In this framework we record each time a given patch is assigned a higher luminance value than every other patch in the same image. Thus we obtain a 16x16 matrix which records in entry (i,j) how often colour i is assigned a higher luminance than colour j. From this matrix we estimate the underlying Thurstone [6] scale values.

3.3 Results

In the one-patch experiment 4 out of 5 observers made grey level settings that were not significantly different from a constant value ($p>0.1$), i.e. settings for different hues were not significantly different from one another. Similarly, 3 of those 4 observers showed no improvement in the fit of a sinusoidal model over a simple linear shift model ($p>0.1$); see Figure 6 for the settings of Observer 2 (other observers show similar settings). Only one observer showed a significant hue-dependency in their settings.

In the two-patch experiment all 7 observers showed a significant hue-dependency. This was seen both in the results of the ANOVA ($p<0.001$) and in the significantly improved fit of a sinusoidal over a simple model ($p < 0.001$); see Figure 7 for the settings of Observer 2 (other observers show similar settings).

In the complex images observers showed a high variability in their luminance settings, and as a result no hue dependency was apparent. When we applied the Thurstone scaling analysis, we found that the derived scaling factors provide a significantly better fit to the pair-comparison data than the null hypothesis that all the colours are equal ($p<0.001$). This holds true for both the original, and abstract forms of the images (i.e. there is no effect of context). Figure 8 shows the scaling factors derived over all observers.

A comparison of Figures 6, 7 and 8 shows that while variability is high in the one-patch experiment, it is much lower for the two-patch experiment and the scaling factors derived from the complex images. The shapes of the curves are also very similar, suggesting similar underlying mechanisms driving the results.

4. DISCUSSION

The results of the one-patch experiment are consistent with the findings of Yaguchi et al. [3], who find three different types of observers: those who show no H-K effect, those who show a complex effect, and those who show a simple effect. We find two types of observer, although our analysis is likely to group the complex and simple effects together as both are deviations from the model that there is no difference between mean settings.

Unlike earlier studies, we find that observer behaviour changes significantly when we increase the number of patches in the scene. For 3 of the 4 observers whose ANOVA results suggest no hue-dependency for single patches, a significant hue-dependency appears in the two-patch experiment. This suggests that future models of the H-K effect might need to take into account the spatial arrangement of image patches.

The results of the complex images show that observers find it hard to repeatedly assign similar grey when there are many colours present in the scene. The analysis of the pair comparison data, however, suggests that the H-K effect still has a strong influence over which colours tend to be made brighter or darker than other colours.

5. ACKNOWLEDGMENTS

This research was partly supported by EPSRC grants EP/E012248/1 and EP/E12159/1. We also acknowledge the help of Glen Harding for useful discussions on the data analysis.

6. REFERENCES

[1] Fairchild, M. (2005). Color appearance models. Wiley, Chichester, England, 2nd edition.

[2] Yaguchi, H., Kawada, A., Shioiri, S. and Miyake, Y. (1993). Individual differences of the contribution of chromatic channels to brightness. JOSA A, 10(6), 1373-1379.

[3] Y. Nayatani, (1997). Simple estimation methods for the Helmholtz–Kohlrausch effect. Color Res. Appl. 22, 385–401.

[4] Smith, K., Landes, P. E., Thollot, J., and Myszkowski, K. (2008). Apparent greyscale: A simple and fast conversion to perceptually accurate images and video. Computer Graphics Forum, 27(2), 193–200.

[5] Gooch, A. A., Olsen, S. C., Tumblin, J. and Gooch, B. (2005). Color2gray: salience-preserving color removal. ACM SIGGRAPH, 634–639.

[6] Thurstone, L. L. (1927b). A law of comparative judgment.

[7] Draper, N. and Smith, H. (1981). Applied Regression Analysis, 2nd Edition. New York: John Wiley & Sons, Inc.Psychological Review, 34 (4), 273-286.

[8] Kaiser, P.K. (1988). Sensation luminance: a new name to distinguish CIE luminance from luminance dependent on an individual's spectral sensitivity.Vision Res. 28, 455-456.

Predicting Perceptions: Proceedings of the 3rd International Conference on Appearance, pp. 74-78, Edinburgh, UK, ISBN 978-1-4716-6869-2, April 2012.

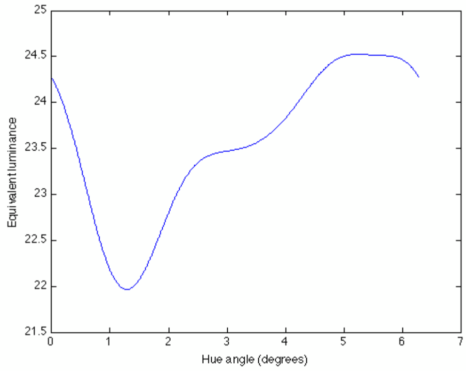

Figure 1: Estimate of the hue-dependency of the Helmholtz Kohlraush effect. We used Nayatani's Helmholtz-Kohlrausch model to predict the luminance that would match an iso- luminant, iso-saturated hue-circle (20 cd/m^2, saturation = 1 in CIELuv space). The x-axis shows the hue angle of a colour, while the y-axis shows the predicted luminance match.

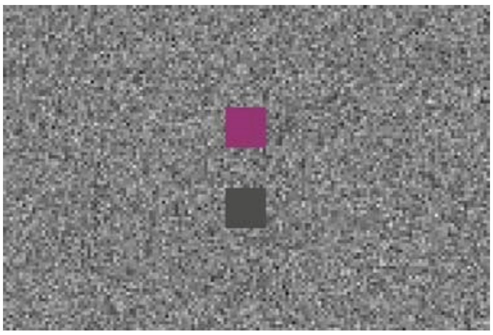

Figure 2: Example stimulus for the one-patch conditions. Individual patches subtended 1.5 degrees of visual angle, with a vertical spacing of 2 degrees visual angle between target (upper) and matching (lower) stimuli. Observers were asked to match the coloured stimulus as closely as possible.

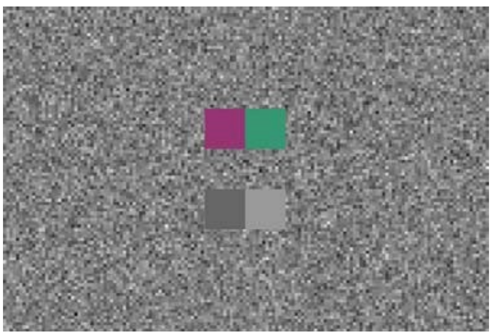

Figure 3: Example stimulus for the two-patch condition (see Figure 1 for true dimensions). Observers were asked to match the coloured stimuli and the difference between them.

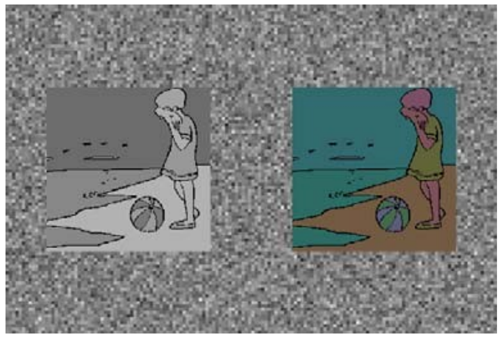

Figure 4: Example display for the complex stimulus configuration. Each separate greyscale region was adjusted to match the corresponding coloured region. Stimuli were balanced for left-right presentation.

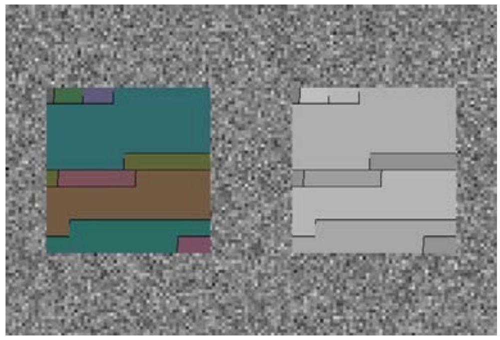

Figure 5: Example stimulus for the complex stimulus abstract-configuration. The image to the left of this figure has identical colour content to the right of Figure 3, with the spatial configuration being altered.

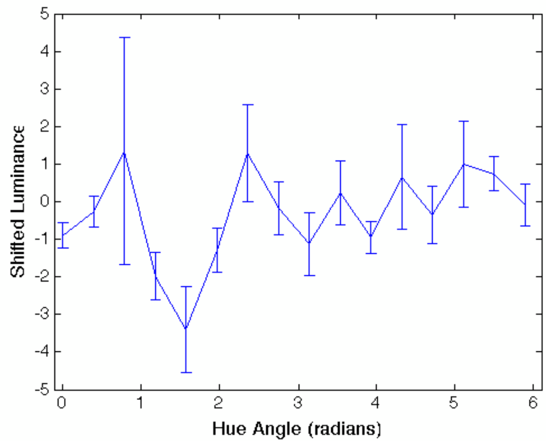

Figure 6: Shifted luminance settings (see text) plotted against hue angle for the 2nd Observer. Error-bars indicate the standard error.

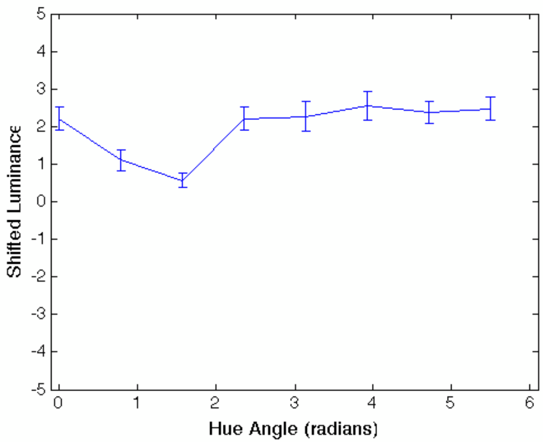

Figure 7: Shifted luminance settings plotted against hue-angle for the 2nd Observer in the 2-patch experiment. Error-bars indicate the standard error.

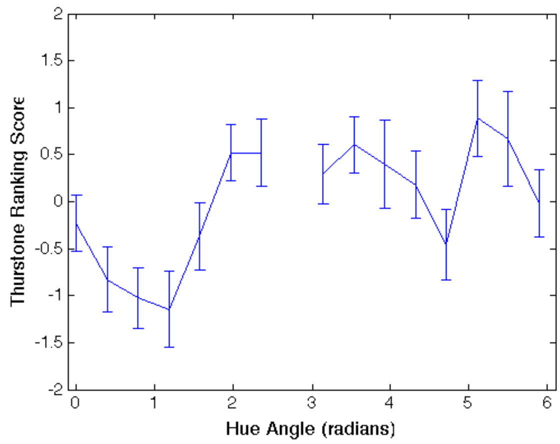

Figure 8: Thurstone ranking, derived from all observer settings in the complex-image configuration, plotted as a function of hue angle. Error bars indicate standard errors for the scale values.