Variability in colour matches between displays

Phil GREEN,[1] Srikrishna NUDURUMATI,[2] Ivar FARUP[1]
[1] Colour and Visual Computing Laboratory, Gjøvik University College, Norway
[2] Global Graphics Software Ltd., Cambridge, UK

ABSTRACT

The variability between colour matches made by different observers on displays is a concern which has been addressed in several previous studies. Inter-observer variability in perceived colour matches was investigated in an experiment in which 21 observers matched a series of test colours, with the reference stimulus on a CRT and the test stimulus on an LCD display. Reference and test colour patches with a 2 degree angular subtense were presented on adjacent displays with a separation between reference and test stimuli of 7cm, and with an opaque black mask covering the remainder of the display screens. The first reference colour was a mid-tone neutral gray, followed by nine chromatic colours. Observers adjusted the test colour to produce a perceived match to the reference, and the resulting colours were measured with a Konica-Minolta CS-1000 telespectroradiometer. The results showed considerable variability in the matches in luminance. In *u'v'* chromaticity, observer variability was found to be less than in luminance, but much higher in blues than in greens and neutrals, suggesting a possible observer metamerism effect.

1. INTRODUCTION

The CIE 1931 Standard Colorimetric Observer was derived by pooling the colour matching functions of multiple observers. These original experiments, and subsequent ones in the following decades, recorded significant variability between matching functions of individual observers. While the typically smooth reflectance spectra of surface colours do not tend to give rise to significantly large inter-observer differences in colour matches, the nature of the spectral emission of self-luminous displays, often characterised by narrow peaks, is more likely to interact with differences in retinal spectral sensitivity to generate inter-observer variation in cross-media colour matches in cases where different colorant technologies are employed. Such inter-observer differences have been reported in previous studies (e.g. Oicherman 2007; Shaw, 2010; Sarkar 2010; Parab 2010; Sarkar 2011).

Modern displays use a variety of light-emitting technologies with very different spectral characteristics, and the goal of the present experiment was to provide experimental data on colour matches made by observers between two different types of displays.

2. EXPERIMENTAL

Following a small pilot study with 6 observers, 21 observers took part in the main phase of the experiment, in which 10 reference colours were matched between a CRT display (employing phosphors as the light-emitting technology) and an LCD display (with LED backlight).[1]

---

[1] The experiment was undertaken by the second author during his employment as an Early-stage Researcher in the Marie Curie Initial Training Networks (ITN) CP7.0 project.

2.1 Displays

The displays were a CRT (Lacie Softproofing Display) and an LED-backlit IPS-LCD (Dell U2412M). The R, G, B primaries of the two displays were measured with a Konica-Minolta CS1000 telespectroradiometer, and the spectral radiances are shown in Figure 1 below. The colour gamut of the two displays can be seen in CIE u',v' coordinates in Figure 2.

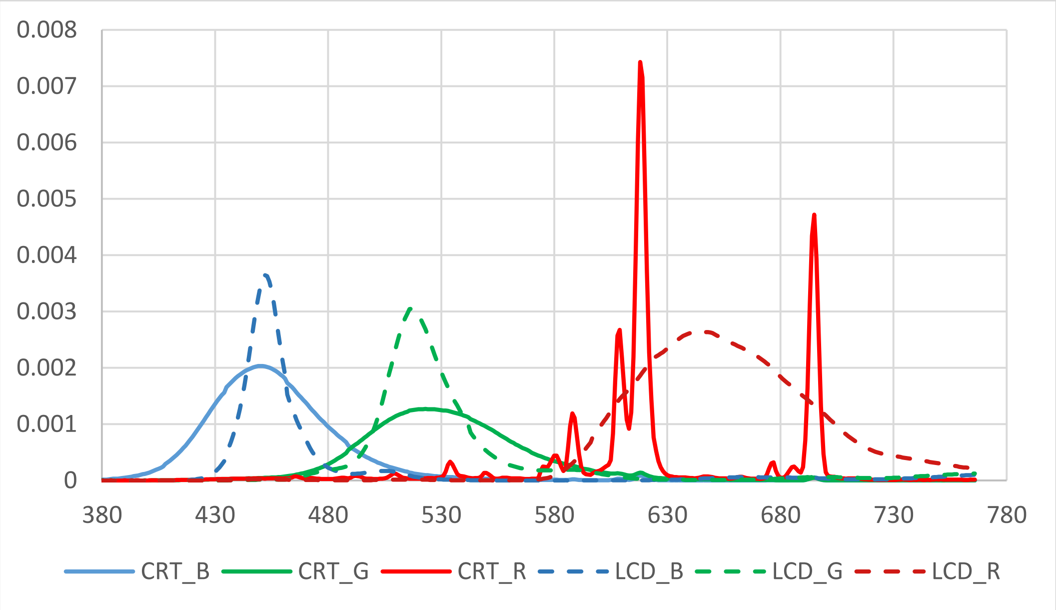

Figure 1: Spectral radiances of CRT and LCD display primaries in the experiment.

Since the dynamic range and colour gamut of the CRT was smaller than that of the LCD, the CRT was used to present the reference colours which observers then matched on the LCD. The neutral gray was presented first to provide an initial reference for brightness matching by observers. Display variability was evaluated by measuring a series of colours over a period of 12 hours after initial warm-up, using the K-M CS1000 TSR. The mean differences from the mean white point luminance over the period were 0.42 and 0.12 cd m^{-2} for the CRT and LCD respectively.

The observer visual matches and the instrumental measurement were performed at the same location on the displays, so the matches were unaffected by any spatial non-uniformity of the displays. There was a warm-up period of at least 15 minutes for the displays and the TSR prior to measurements and observations.

2.2 Observers

Following a pilot study with six observers, 21 observers aged 24-55 participated in the experiment. An Isihara test was conducted prior to the experiment for those observers had not previously performed such a test, and one was found to be colour deficient. Non-expert observers were given training in colour mixing. Observer repeatability was evaluated by observers performing matches between similar patches during each run of the experiment.

2.2 Reference Colours

10 colours were selected in RGB coordinates and displayed on the reference (CRT) display). The resulting colours were measured with the TSR and are shown in CIE u',v' coordinates

in Figures 2, where the reference colours are labelled 1-10. Some of the reference colours lie close to or on the boundary of the CRT gamut, but are within the gamut of the LCD display.

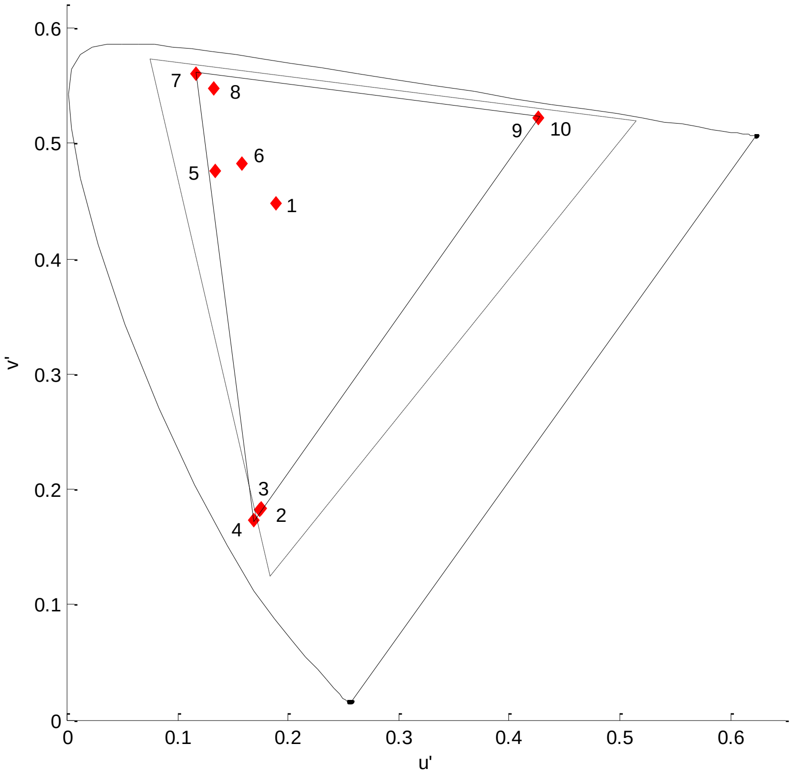

Figure 2: Reference colours in CIE u',v' with gamut of CRT (solid) and LCD (dashed) displays

2.3 Experimental Setup

The experiment was conducted in a completely dark room with the display colour patches being the only stimuli. Reference and test patches were circular and 3.5 cm in diameter. Observers were located one metre from the display faceplate, so that the colour stimulus gave an angular subtense of 2°. In order to exclude any effect due to the display backlight within the faceplate, the patches were observed through circular openings made within an opaque black sheet placed in front of the displays. The displays were placed adjacent to each other so that the reference and the test targets were located at approximately 7 cm apart. This is essentially an aperture mode rather than object mode viewing condition.

2.4 Experiment

In the experiment the reference colour was shown on the CRT display and the test colour shown simultaneously on the LCD display. Observers were asked to match the appearance of the test to the reference colour. To do this they were provided with three computer mice with two-way spherical scroll controlling the amounts of individual primaries (R, G & B) of the test colour, together with a keyboard control for adjustment of the brightness. Once the observer was satisfied with the match, the corresponding R, G, B values were recorded and subsequently measured with the TSR positioned at the observer location in the same viewing conditions.

3. RESULTS

CIE XYZ values for the reference colours and the observers matches were calculated from the measured radiances using the CIE 1931 Standard Colorimetric Observer, as shown in eqn. 1 below.

$$X = 683 \sum_{767}^{380} L(\lambda)\bar{x}(\lambda)$$

where $L(\lambda)$ is the measured spectral radiance (in W sr^{-1} m^{-2}) at 1nm intervals and $\bar{x}(\lambda)$ is the vector of matching function values for X; and Y and Z are computed analogously. The constant 683 results in Y tristimulus values in units of cd m^{-2}.

Since the experiment was conducted in a dark room with no adapting white and colours were judged in aperture mode, there is no reference white or illuminant and hence the data cannot be converted to a colour space such as CIELAB. For this reason the variability between the reference colour and the observer matches is shown in terms of luminance ΔY and chromaticity $\Delta u',v'$ (CIE 2014). The results are shown in Table 1. [We note that if the data is converted to CIELAB, using the display peak white as the reference white as is commonly done in colour management, the average CIELAB 1976 ΔE^*_{ab} colour difference from the reference is 19.5 and the inter-observer variability (expressed as the mean colour difference from the mean, or MCDM) is 10.4, which gives a misleading impression of the accuracy of the matches.]

Table 1.Differences in luminance ΔY and chromaticity $\Delta u',v'$ between reference colours and observer matches

	ΔY			$\Delta u'v'$		
Colour	Median	Max	95th percentile	Median	Max	95th percentile
1	6.57	9.6	9.51	0.0107	0.0244	0.0201
2	0.65	1.27	1.15	0.0207	0.0543	0.0537
3	0.91	2.91	2.32	0.0373	0.1023	0.0819
4	2.46	5.7	5.5	0.012	0.0526	0.0491
5	5.31	14.33	12.85	0.0128	0.0198	0.0197
6	4.01	14.9	13.62	0.0178	0.0256	0.0235
7	9.1	36.43	30.99	0.0134	0.0572	0.0391
8	3.54	32.98	23.94	0.0267	0.1424	0.1421
9	2.32	9.04	8.52	0.0253	0.047	0.0469
10	3.17	13.32	13.21	0.0238	0.0506	0.0492
Mean	3.8	14.05	12.16	0.0201	0.0576	0.0525

The variability in luminance matches is very high for some colours, shown by 95th percentile ΔY values of up to 30.99. The matches in u',v' chromaticity have smaller variability, as seen in Figures 3-5, where the reference colour is shown in red and the observer matches in black.

It should be noted that where the reference colours are close to the gamut boundary, the direction of possible matches in chromaticity space by observers was constrained.

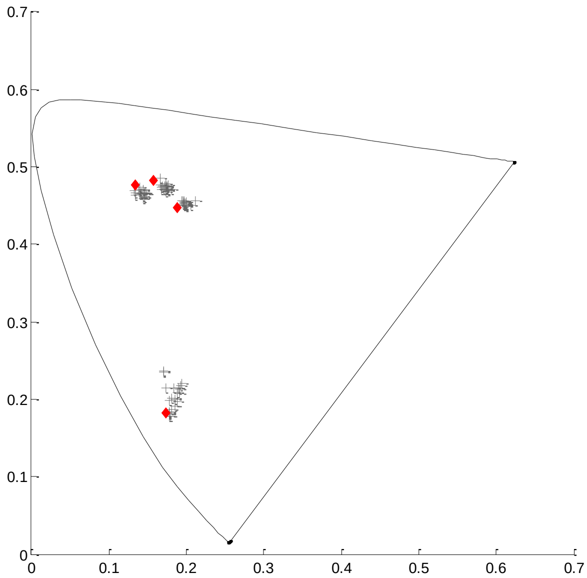

Figure 3. CIE u',v' coordinates of reference colours (red marker) and corresponding observer matches for colours 1, 2, 5 and 6

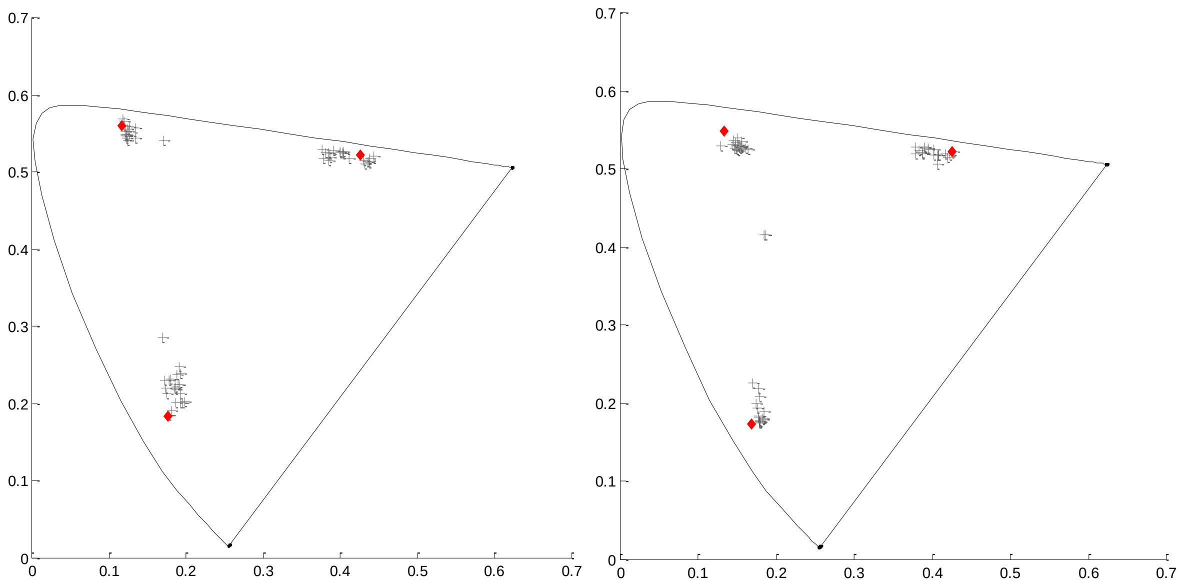

Figure 4-5. CIE u',v' coordinates of reference colours (red marker) and corresponding observer matches for colours 3, 7 & 10 and 4, 8 & 9 respectively

It can be seen from Figures 3-5 that observer variability in greenish and neutral colours is relatively small, while in red (colours 9-10) and blue (colours 2-4) regions it is considerably higher. The three blue reference colours are located at near-identical chromaticity coordinates, and the pattern of observer matches is very similar between the three, suggesting a possibly systematic difference in the way individual observers made the matches. A similar pattern can be seen in the two red colours, although less pronounced.

4. CONCLUSIONS

Observers matched a set of 10 reference colours presented on a CRT display, using an LED-backlit LCD display. Significant variability was seen in the luminance of the matches, suggesting that luminance is of less importance than chromaticity in an aperture-mode colour matching task. Variability in matching of chromaticity was small in some colours, but relatively high in certain colours, notably in blue. The similar trend seen in the matches for colours of similar chromaticity suggests there is a possible observer metamerism effect arising from different cone sensitivities of the observers. However, further analysis is needed to confirm this possibility. If present, it would be consistent with results obtained in other studies.

The results also indicate that aperture mode colour matches should be evaluated in luminance and chromaticity, with higher tolerances in luminance.

ACKNOWLEDGEMENTS

This work was supported by Marie Curie Initial Training Networks (ITN) CP7.0 N- 290154.

REFERENCES

Shaw, M. and Fairchild, M. (2002) *Evaluating the 1931 CIE Color Matching Functions*, Col. Res. Appl. 27(5) 316–329

Oicherman, B. (2007) *Effects of colorimetric additivity failure and of observer metamerism on cross-media colour matching*, (PhD thesis) University of Leeds, UK

Parab, N. and Green, P. (2011) *Soft proofing of printed colours on substrates with optical brightening agents*, Proc. SPIE 7866 doi:10.1117/12.872219

Sarkar, A., L. Blondé, P. Le Callet, F. Autrusseau, J. Stauder, P. Morvan, (2010) *Toward Reducing Observer Metamerism in Industrial Applications: Colorimetric Observer Categories and Observer Classification*, Proc IS&T 18th Color Imaging Conf. 307-313

Sarkar, A., F. Autrusseau, F. Viénot, P. Le Callet, L. Blondé, (2011) *From CIE 2006 Physiological Model to Improved Age-Dependent and Average Colorimetric Observers*, J. Opt. Soc. Am. (JOSA A), 28(10) 2033-48

CIE Technical Note 001:2014 *Chromaticity Difference Specification for Light Sources*, Vienna: CIE

Address: Prof. Phil GREEN, Colour and Visual Computing Laboratory,
Gjøvik University College, Teknologivn. 22, 2815 Gjøvik, Norway
E-mails: philip.green@hig.no, srisrikrishna9@gmail.com, ivar.farup@hig.no