

EBU

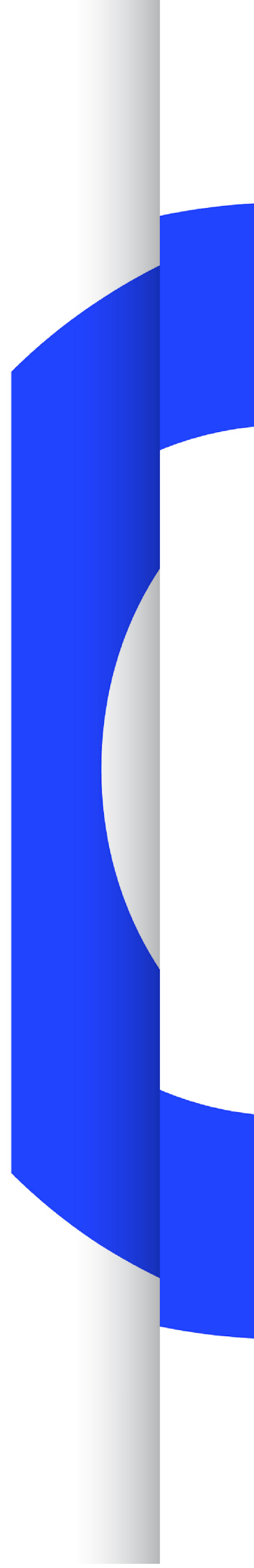
OPERATING EUROVISION AND EURORADIO

TECH 3354

**COMPARISON OF CIE COLOUR METRICS
FOR USE IN THE TELEVISION LIGHTING
CONSISTENCY INDEX (TLCI-2012)**

Source: FTV-LED

Geneva
November 2012



Contents

Introduction.....	5
1. The Principle of the Television Lighting Consistency Index (TLCI-2012)	5
2. Test setup	6
2.1 Test samples	6
2.2 Test and Reference luminaires.....	6
2.3 Camera responsivity curves.....	6
2.4 Camera matrix	7
2.5 Camera gamma-correction.....	7
2.6 Display gamma	7
2.7 Display primaries	8
3. Test procedure	8
4. Analysis of results	11
4.1 Analysis for individual subjects	11
4.2 Analysis for all subjects	14
5. Bibliography.....	16
Appendix 1: Reflectivities of Test luminaires	17
Appendix 2: CIE1964.....	19
Appendix 3: CIELUV	21
Appendix 4: CIELAB	23
Appendix 5: CMC	25
Appendix 6: CIE1994.....	27
Appendix 7: CIEDE2000.....	29

* Page intentionally left blank. This document is paginated for two sided printing

Comparison of CIE colour metrics for the assessment of the colorimetric properties of luminaires, the Television Lighting Consistency Index (TLCI-2012)

<i>EBU Committee</i>	<i>First Issued</i>	<i>Revised</i>	<i>Re-issued</i>
TC	2012		

Keywords: TLCI-2012, Television Lighting Consistency Index, Colorimetry, Luminaire.

Introduction

This document describes subjective tests conducted in order to establish the most appropriate colour metric for assessing colour differences, for use in the ‘Television Lighting Consistency Index-2012’ (TLCI-2012).

1. The Principle of the Television Lighting Consistency Index (TLCI-2012)

The purpose of the TLCI-2012 is to provide a way to assess the colour-rendering performance of luminaires when used for television lighting. The mathematical simulation mimics a complete television camera and display, using only those parameters that affect colour performance. The method calculates the difference in the colour appearance of a set of test colour samples when separately lit by a test luminaire and a reference luminaire, and shown on a reference display.

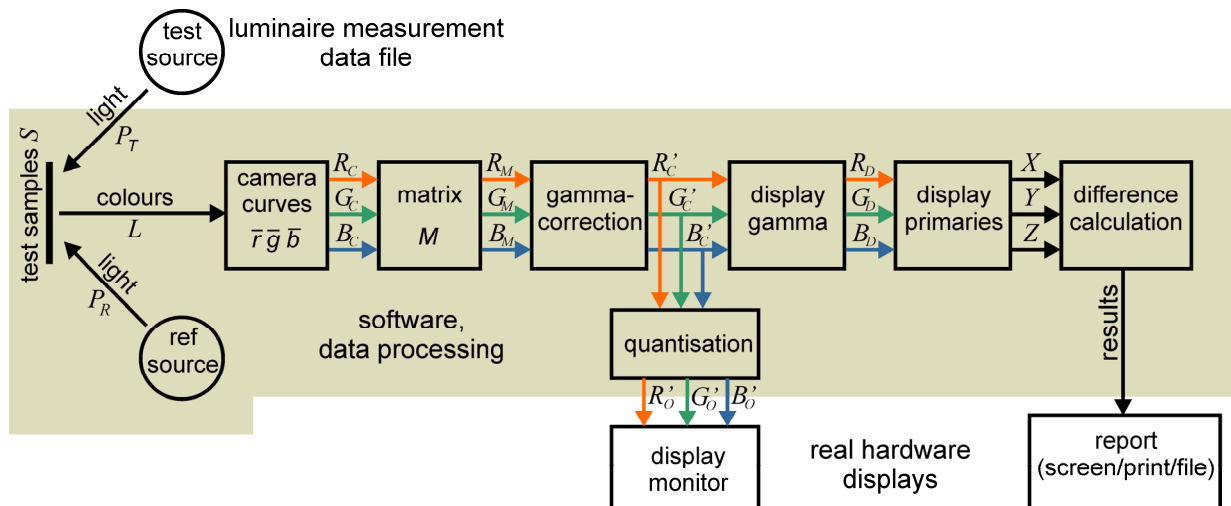


Figure 1: TLCI-2012 functional block diagram

None of the parameters of the television channel are dealt with here: the TLCI-2012 software, further described in EBU Tech 3355, was used merely to generate sets of colour differences on-screen on a computer or video display for use in subjective testing.

Six different CIE algorithms are available for calculating the visual appearance of a colour-difference; details are given in the Appendices. The oldest, CIE1964, was used in both the CRI and in the TLCI work initiated by Sproson and Taylor. But the later metrics are claimed to be better, in that they ever-more closely represent the human visual experience.

2. Test setup

2.1 Test samples

A mathematical representation of the ColorChecker® chart was used, since it is widely recognised and readily available as a test chart, even though it was not originally designed for television use. Reflectivity tables for the colour samples are given in EBU Tech 3355 document.

2.2 Test and Reference luminaires

The test samples were notionally lit by three test sources, each the measured power distribution data of a real luminaire. They were chosen to represent good, poor, and indifferent performance. The spectral power distributions are given in the Appendixes.

In each case, the Correlated Colour Temperature (CCT) of the Test luminaire was calculated and used as the specification for the Reference luminaire. The method of calculating CCT and the spectral power distribution of the standard source is given in EBU Tech 3355.

2.3 Camera responsivity curves

Rather than use data from any real camera, representative responsivity curves of Plumbicon®-type tubes were used, taken from tabulations in the Sproson/Taylor work. These were a best-estimate of the responsivities of real cameras at the time (1970 - 1990) and there was little reason to believe that data for more modern cameras would be different. Measurements on more recent cameras were being undertaken at the time, and the results can be seen in EBU Tech 3353.

Figure 2 shows that the precise choice of camera curves is not highly critical, and that even the curves from negative film stocks produce similar results. This test used the CIEDE2000 difference metric for measurements on a generic single-colour white LED luminaire.

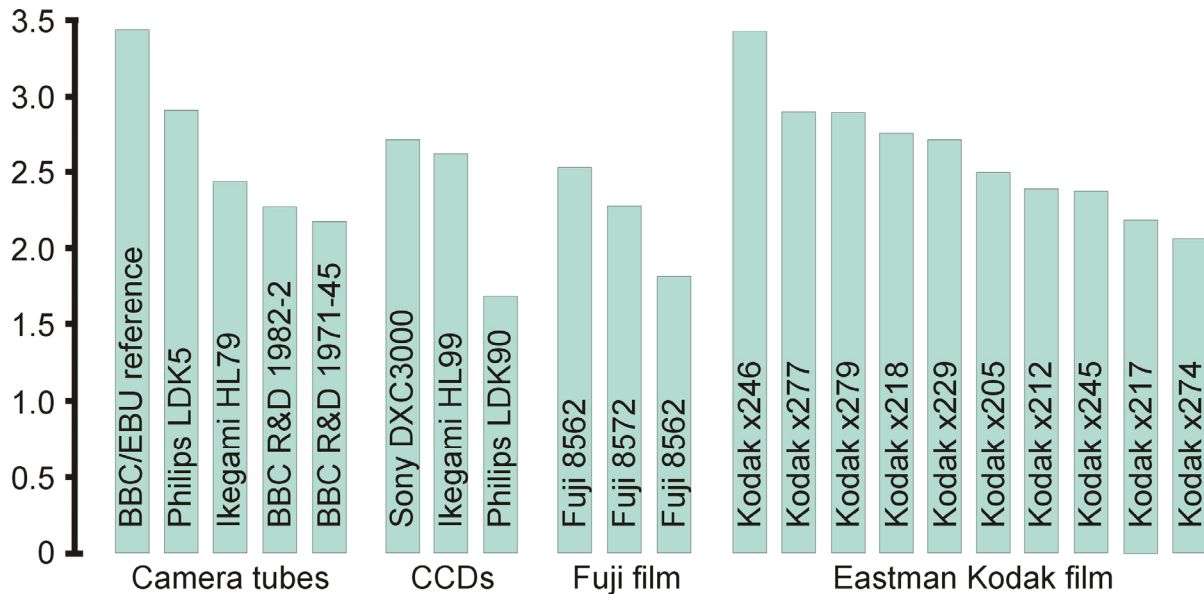


Figure 2: Average ΔE^* values for various camera-responsivity sets

The white balance level was set to 90%, which ensured that the video level for the peak white patch became 100% (since the reflectance of that patch is known to be 90.01%).

2.4 Camera matrix

The matrix values were taken from the final Taylor report (1988).

$$\begin{bmatrix} R_M \\ G_M \\ B_M \end{bmatrix} = \begin{bmatrix} 1.1380 & -0.1750 & 0.0370 \\ -0.1120 & 1.1510 & -0.0390 \\ 0.0000 & -0.0910 & 1.0910 \end{bmatrix} \bullet \begin{bmatrix} R_S \\ G_S \\ B_S \end{bmatrix}$$

Where R_S , G_S and B_S are the linear signals from the sensor(s) and R_M , G_M and B_M are the camera-output linear signals. Since this matrix had been optimised for an overall-linear television system, it produced rather over-saturated colours when used in a more realistic system with display-gamma and camera gamma-correction. Therefore, a saturation control was set to 80% which produced colour performance which is much more typical of real systems.

2.5 Camera gamma-correction

The gamma-correction was set in the equation of ITU-R BT.709, which is recommended for all HDTV or newer systems.

$$R_C' = 1.099R_C^{0.45} - 0.099 \text{ for } V > 0.018, \text{ otherwise } R_C' = 4.5R_C$$

The real-time output of the software was thus in $R' G' B'$ form, and fed from the computer via VGA, DVI, HDMI or Displayport to a television display. The software coded the video feed using ITU-R BT.709 quantisation, 8-bit, to form the R_0' , G_0' and B_0' signals.

2.6 Display gamma

For the purpose of calculating the colour-appearance, the display gamma was set to 2.4, thus:

$$V_D = V_C^{2.4}$$

In practice, the choice of display gamma is relatively unimportant since the reported colour difference values change very little between display gamma of 2.2 and 2.8. Figure 3 shows the colour differences, measured using the CIEDE2000 difference metric, for a generic single-colour white LED luminaire.

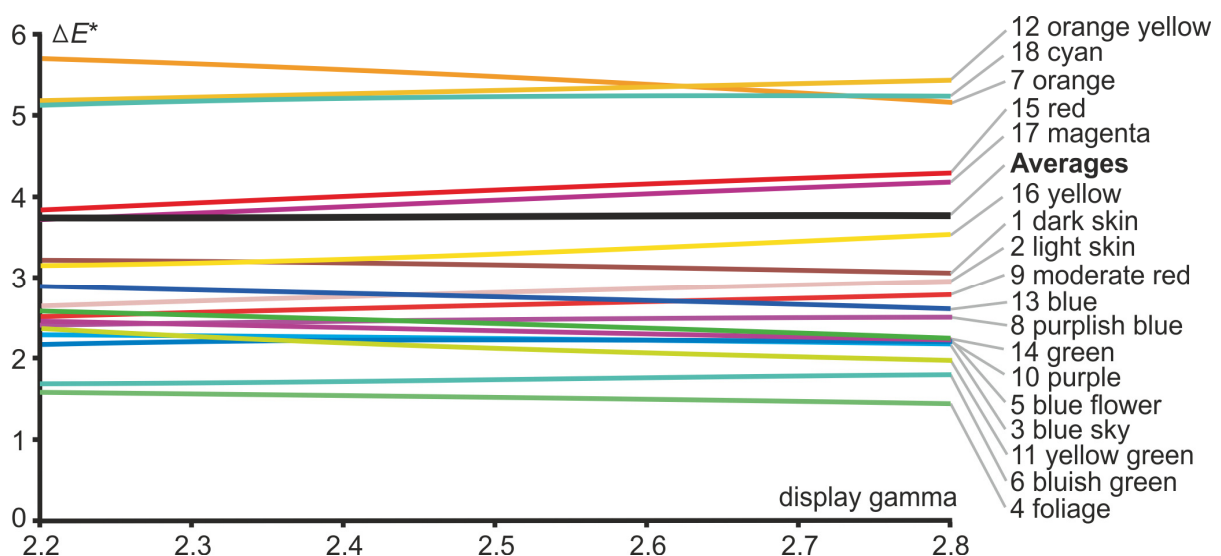


Figure 3: ColorChecker® colour differences for a generic white LED luminaire

2.7 Display primaries

ITU-R BT.709 primaries were chosen.

	x	y
R	0.64	0.33
G	0.30	0.60
B	0.15	0.06
White balance, D65	0.3127	0.3290

The display synthesis matrix, derived directly from these primaries and balance colour, is:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412391 & 0.357584 & 0.180481 \\ 0.212639 & 0.715169 & 0.072192 \\ 0.019331 & 0.119195 & 0.950532 \end{bmatrix} \bullet \begin{bmatrix} R_D \\ G_D \\ B_D \end{bmatrix}$$

and it is these tristimulus values, X_T , Y_T , Z_T and X_R , Y_R , Z_R for the Test and Reference luminaires respectively, which are used for the calculation of the difference in colour appearance and the formulation of the TLCI-2012 value Q .

3. Test procedure

The test subjects were provided with a spreadsheet that contained all the instructions needed for the performance of the test, together with marking sheets for each individual test.

Each subject was asked to display the output of the TLCI-2012 software (EBU Tech 3355) on a conventional television display or monitor, rather than on a computer display. SMPTE colour bars were made available in the software to assist in correctly setting the display brightness and colour controls.

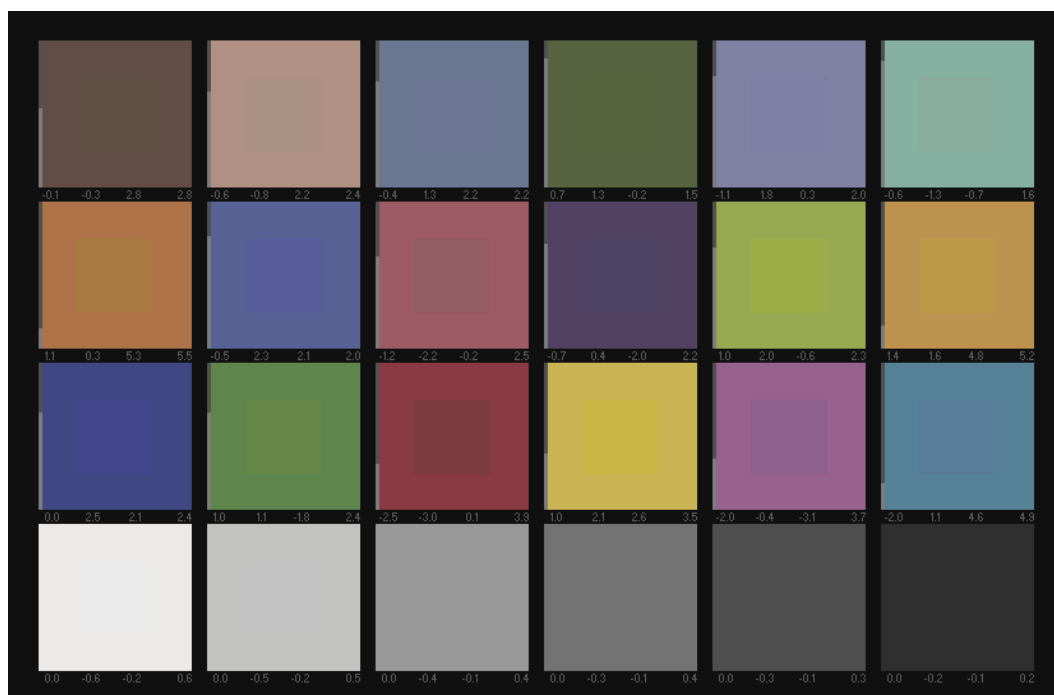


Figure 4: TLCI output screen for single-colour, white LED, CIEDE2000




























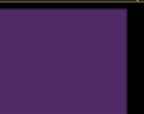








	A	B	C	D	E	F	G	H	I	J	K	L	M	
1	Test source	Fluorescent (TV white)					Reference	AUTO CCT						
2	Camera curves	BBC 1988-2 3 plumbicon												
3	Colour metric	CIEDE2000			Parameter			ΔE*						
4	Weighting	l=	1	c=	1	h=	1							
5														
6														
7	Rate the metric's scoring of horizontally adjacent pairs of colour patches	Dark skin	Light skin	Blue sky	Foliage	Blue flower	Bluish green							
8														
9		ΔE* 0.6	ΔE* 1.0	ΔE* 1.7	ΔE* 2.7	ΔE* 2.2	ΔE* 2.0							
10		<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>							
11		Orange	Purplish blue	Moderate red	Purple	Yellow green	Orange yellow							
12														
13		ΔE* 2.1	ΔE* 1.2	ΔE* 0.8	ΔE* 3.0	ΔE* 2.2	ΔE* 4.5							
14		<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>							
15		Blue	Green	Red	Yellow	Magenta	Cyan							
16														
17		ΔE* 1.2	ΔE* 3.1	ΔE* 0.8	ΔE* 2.8	ΔE* 2.2	ΔE* 3.1							
18		<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>							
19														
20		Rate the metric's scoring of vertically adjacent pairs of colour patches												
21			<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>						
22														
23			<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>						
24														
25														
26	Score	0			Mean ΔE*	2.07								
27	Minimum score	0			Minimum ΔE*	0.6								
28	Maximum score	0			Maximum ΔE*	4.5								
29	Completed	#VALUE!												

Figure 5: spreadsheet & marking sheet for Fluorescent (TV white), CIEDE2000

The TLCI-2102 software output was a representation of the ColorChecker® chart, processed through the software camera. Each patch contained the colour as analysed by the Reference luminaire, with an inset of the same colour as analysed by the Test luminaire. Below each patch, four numbers are given, the calculated colour differences ΔL^* , ΔC^* , ΔH^* and ΔE^* , from left to right, where only the ΔE^* was used for the tests.

There were 15 rounds of tests; five for each of the three Test luminaires, each using all the CIE colour-difference metrics excluding CIE1964.

Referring to Figure 5 for all spreadsheet references, cell B1 identified the Test luminaire, cell I1 identified the Reference luminaire (being a standard source of the same CCT as the Test luminaire). Cells B3 and I3 identified the relevant colour difference metric and parameter.

The subject was asked to give a verdict on the relationship between the ΔE^* values reported for two adjacent colour patches at a time, in the orange-shaded cells (L23 is highlighted as an example). Below each colour patch the ΔE^* value is given, and the subject is asked to select from a restricted table of verdicts, which is a subset of a 5-point scale:

Score	Verdict
+2	Agree with numbers, or too close to tell
+1	Agree with direction but not with numbers
(0)	(not assessed)
(-1)	(not allowed)
-2	Disagree with direction

Score “+2” was to indicate that the subject agreed that the two ΔE^* values correctly described the relationship between the two colour differences. Score “+1” was to indicate that the general sense of the numbers was correct in terms of the relationship of the difference, but that the numbers didn’t give an accurate feel for the magnitude. Score “-2” was to indicate that the sense of the numbers was wrong (e.g. for the two skin tones, that the difference for Dark Skin was more visible than that for Light Skin, in the example shown in Figures 4 and 5).

Scores “0” and “-1” were disallowed: “0” (actually a blank entry) was used to signify that a test had not yet been done, and “-1” would logically be associated with a verdict that the direction was wrong but the numbers correct, which cannot happen. It could be argued that a score of “0” could signify that it was too difficult to assign a score, but, in the event, forcing a polarised score did not prove difficult for any of the subjects.

The spreadsheet presented a total of all the scores in cell B26. Cells B27 and B28 recorded the maximum and minimum scores given, while cells H26 to H28 recorded the mean and the range of the ΔE^* values for the test. In practice, only the total-score value (cell B26) was used in the analysis. The maximum total score-value for each test was thus $2 \times (3 \times 5 + 2 \times 6) = 54$.

When all 15 combinations of Test luminaire and colour-difference metric had been tested, the subject was then asked to perform a cross-check test.

Each of the three Test luminaires was used to create a TLCI-2012 software output (see Figure 6) but this time with the colour-difference error values suppressed. The subject was then asked to assign a number to the visibility of the colour difference within each colour patch. No guidance was given; the subject was free to use any scale and any precision. All that was asked was for the numbering to be as consistent as possible. The subjects were told that this was a cross-check, a source of data which could and would be used to establish the reliability of the test results.

Throughout all the tests, which were self-administered, the subject was free to repeat any test, changing results if necessary.

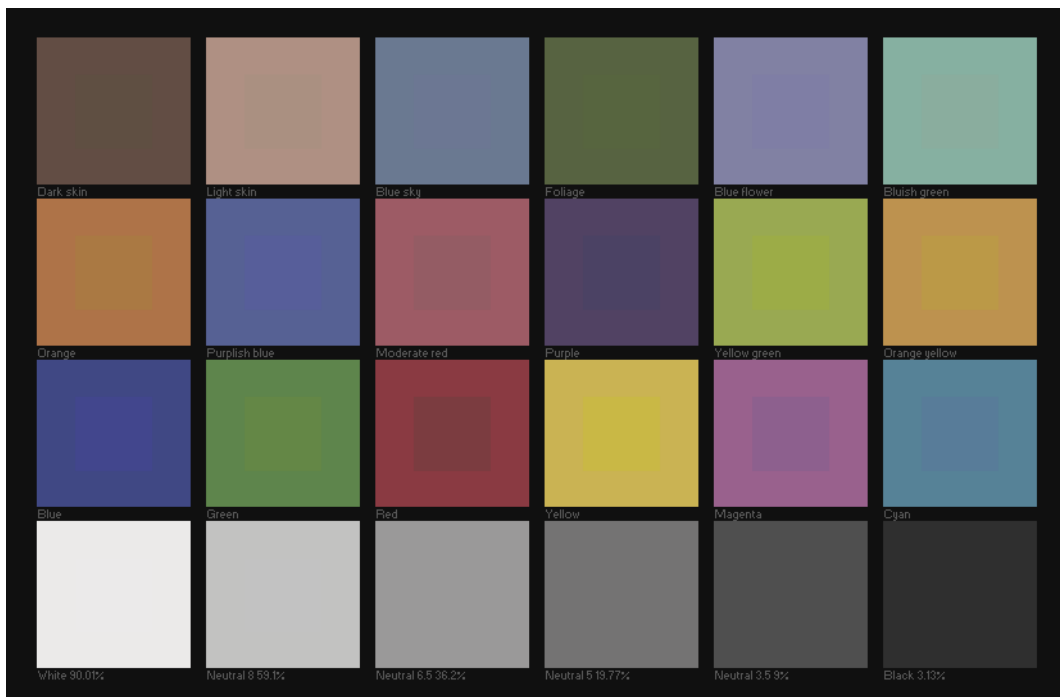


Figure 6: TLCI-2012 output screen for single-colour, white LED, CIEDE2000, with no Δ error reports

4. Analysis of results

4.1 Analysis for individual subjects

Separate analysis was made for each subject’s results, as shown in Figure 7.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
1	Subjective scoring					High quality source, LED	Fluorescent, TV white	Generic single white LEDs	Average	High quality source, LED	Fluorescent, TV white	Generic single white LEDs	High quality source, LED	Fluorescent, TV white	Generic single white LEDs	Sum	
2						Δ average			Δ PMR			Score					
3						CIELUV	2.82	5.02	7.25	5.03	1.95	1.89	2.28	-14	10	24	20
4						CIELAB	2.31	4.39	5.86	4.19	2.08	1.93	1.60	-12	5	-3	-10
5						CMC	1.41	2.46	3.88	2.58	2.13	1.87	1.88	5	15	15	35
6						CIE1994	1.12	2.11	3.04	2.09	2.32	1.89	1.81	13	19	18	50
7						CIEDE2000	1.08	2.07	2.96	2.04	2.12	2.18	1.86	11	18	31	60

Figure 7: Scoring analysis, subject ‘N’

Columns F to L record the statistics of the ΔE^* errors, which were not used in further analysis. Columns M to O copy the score values from cell B26 in each score sheet. Column P totals the inter-patch scores for each CIE colour-difference metric.

It was expected that there would be a steady increase in these levels from CIELUV to CIEDE2000, and that is generally what happened, except that CILEUV performed significantly better than anticipated. Note that the highest inter-patch score for each luminaire did not necessarily coincide with that for the sum (e.g. the green high-lighted cells M6, N6 and O7).

Figure 8 shows this result graphically.

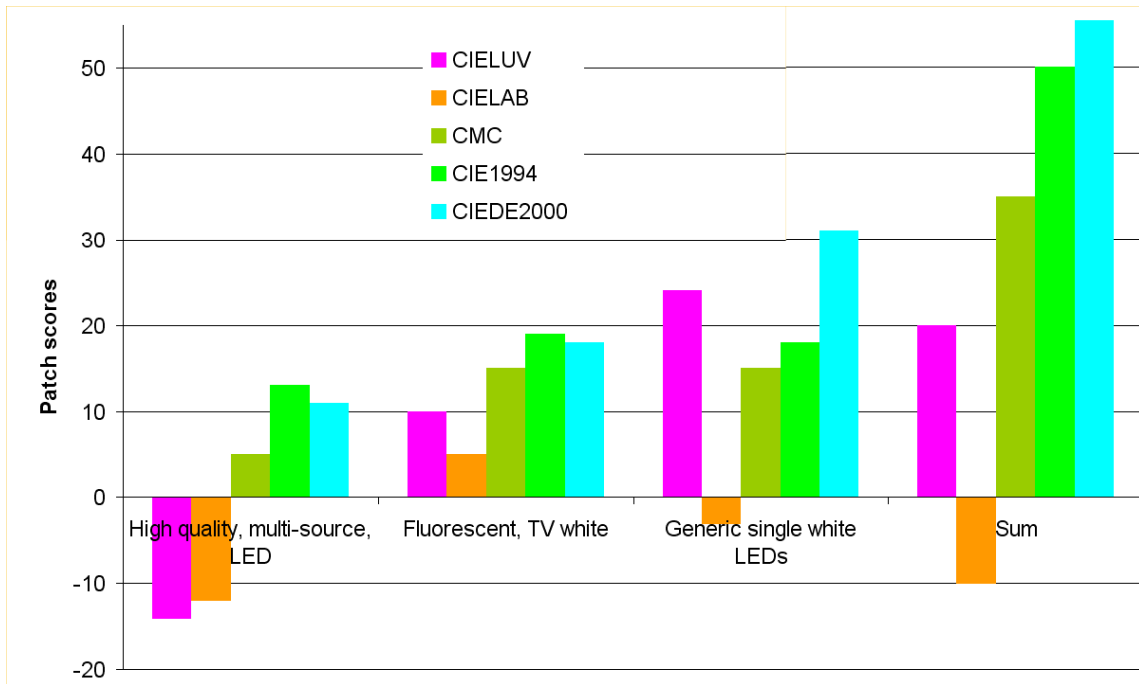


Figure 8: Inter-patch score distribution for subject 'N'

It is evident that the inter-patch scores are almost random for the high-quality luminaire. This is almost inevitable since the individual colour-differences are almost all small, and therefore difficult to assess. The most reliable scoring comes with the generic single-colour white LED luminaire, which has the largest performance errors, and therefore the most visible colour differences. Nevertheless, the Sum inter-patch score shows a clear progression towards CIEDE2000. The big surprise is that CIELAB performs consistently poorly, and that CIELUV is consistently better. None of the metrics is perfect, or even nearly so.

The subject's personal rating of the colour-differences (the cross-check test) was recorded in columns D, J and P, the pale blue highlighted cells in Figure 9. These values are treated as a separate metric, here called 'Eyes'.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
9																					
10			ΔE* values	Generic single white LEDs (112)					Fluorescent, TV white					Gekko Kelvin Light 5600K							
11			Assessments	Eyes	LUV	LAB	CMC	1994	2000	Eyes	LUV	LAB	CMC	1994	2000	Eyes	LUV	LAB	CMC	1994	2000
12		1	Dark skin	1.5	3.6	3.3	5.9	2.7	2.8	0.8	1.3	0.9	0.9	0.7	0.6	0.1	1.5	1.3	2.2	1.0	1.1
13		2	Light skin	1.4	4.4	3.1	4.6	2.3	2.4	1	2.6	1.8	1.2	1.1	1.0	0.5	3.1	2.0	1.6	1.2	1.2
14		3	Blue sky	1	4.1	3.6	3.1	2.4	2.2	0.5	3.8	3.2	2.6	2.0	1.7	0.2	1.3	0.8	0.5	0.5	0.5
15		4	Foliage	1.2	3.0	3.1	1.8	1.6	1.5	1.5	5.8	5.9	3.3	2.8	2.7	0	1.3	1.5	0.8	0.7	0.7
16		5	Blue flower	1.5	5.7	4.2	2.4	2.2	2.1	1.4	6.2	4.9	2.7	2.4	2.2	0	1.7	1.4	1.1	0.9	0.9
17		6	Bluish green	1.4	3.7	2.8	1.8	1.7	1.6	0.5	4.1	2.9	2.2	2.1	2.0	0.2	4.2	3.2	2.3	2.1	2.1
18		7	Orange	2	10.3	7.8	7.3	4.6	5.5	0.8	4.1	3.1	2.7	1.8	2.1	0	2.5	1.5	1.0	0.7	0.7
19		8	Purplish blue	1	7.9	7.2	3.4	2.9	2.0	0.5	4.1	4.0	2.0	1.7	1.2	0.3	0.3	0.3	0.1	0.1	0.1
20		9	Moderate red	2	9.7	5.6	2.9	2.6	2.5	0.5	2.6	1.6	0.9	0.8	0.8	0.5	3.2	2.0	1.0	0.9	0.9
21		10	Purple	1.8	3.8	2.8	2.5	2.0	2.2	1	6.7	5.0	2.9	2.6	3.0	1	5.1	4.8	3.0	2.4	2.3
22		11	Yellow green	1.6	6.4	7.4	2.8	2.4	2.3	1.2	7.1	8.0	2.9	2.4	2.2	0	2.5	3.0	1.1	0.9	0.9
23		12	Orange yellow	2	10.3	9.0	5.6	4.6	5.3	1.5	9.5	7.9	4.6	4.0	4.5	0	1.8	1.7	0.7	0.5	0.5
24		13	Blue	1	7.8	8.7	3.6	2.9	2.4	1.2	2.1	3.4	1.8	1.4	1.2	0.5	2.0	1.9	0.9	0.7	0.5
25		14	Green	1.6	4.8	4.9	2.5	2.5	2.4	1.5	8.1	8.5	3.6	3.1	3.1	0.2	3.0	3.2	1.3	1.1	1.1
26		15	Red	3	16.5	9.1	5.5	4.4	3.9	0	3.3	2.0	1.0	0.8	0.8	1.5	5.5	3.0	1.9	1.5	1.3
27		16	Yellow green	2	9.1	9.4	3.9	3.4	3.5	1.6	7.6	7.4	3.1	2.7	2.8	0	1.6	1.9	0.7	0.5	0.5
28		17	Magenta	2.5	10.9	6.0	4.0	4.0	3.7	0.2	7.4	4.7	2.4	2.4	2.2	1.3	5.1	4.2	2.1	1.9	1.9
29		18	Cyan	2.8	8.5	7.5	6.2	5.5	4.9	2	4.0	3.9	3.4	3.2	3.1	0	5.0	3.8	3.0	2.6	2.3
30			Correlation coefficient, Eyes with metric	0.732	0.447	0.525	0.747	0.697		0.391	0.581	0.663	0.659	0.644		0.668	0.504	0.382	0.415	0.395	
31																					

Figure 9: 'Eyes' ratings for subject 'N'

Each column of 'Eyes' scoring was compared with the ΔE* values for each metric and the correlation coefficients were recorded in row 30. A positive value indicates a general agreement; a

maximum possible value of 1 would indicate a perfect match. It is not yet clear what these results could imply, except that none of the metrics seems to correlate particularly well with the visual experience of subject ‘N’.

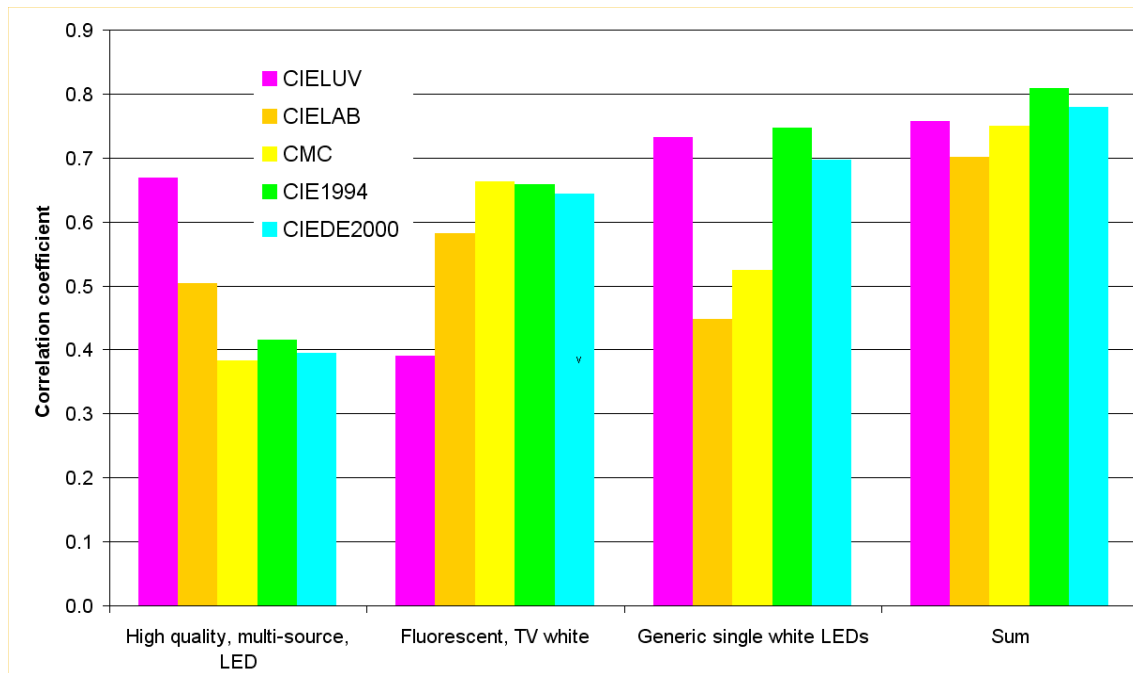


Figure 10: ‘Eyes’ correlation coefficients for subject ‘N’

Figure 11 shows further analysis, which is more revealing, however.

In Figure 11 the contents of cells R3 to T7 are copied from row 30 of each of the luminaires in Figure 9. Cells U3 to U7 are the correlation coefficients between the ‘Eyes’ values of columns D, J and P in Figure 9 and the ΔE^* values for each metric, for all three luminaires taken together. Row 8 shows the correlation coefficients, for each luminaire, between the inter-patch scores (columns M to O) and the ‘Eyes’ correlation coefficients (columns R to T). This clearly reveals the difficulty of assessing the small colour-differences for the high-quality luminaire. Cell W8 shows the correlation coefficient for all luminaires (columns P and U). This single value indicates the quality of the observations.

	C	D	E	M	N	O	P	Q	R	S	T	U	V	W
1	Subjective scoring			High quality, multi-source, LED	Fluorescent, TV white	Generic single white LEDs	Sum		High quality, multi-source, LED	Fluorescent, TV white	Generic single white LEDs	All luminaires		Correlation coefficient, all Scores with all CCs
2				Score					Correlation coefficients					
3		CIELUV		-14	10	24	20		0.668	0.581	0.732	0.757		
4		CIELAB		-12	5	-3	-10		0.504	0.581	0.447	0.701		
5		CMC		5	15	15	35		0.382	0.663	0.525	0.750		
6		CIE1994		13	19	18	50		0.415	0.659	0.747	0.809		
7		CIEDE2000		11	18	31	60		0.395	0.644	0.697	0.780		
8				Correlation coefficient, luminaire scores with CCs					-0.85	0.89	0.82		0.89	

Figure 11: ‘Eyes’ secondary correlation coefficients for subject ‘N’

However, analysis of the full group of subjects is necessary to reveal the underlying implications.

4.2 Analysis for all subjects

Figure 12 shows the first part of the overall analysis. Cells B3 to T7 are values copied from cells P3 to P7 from the individual analysis sheets. Row 8 values are copied from cell W8 in the individual sheets. An arbitrary decision was taken, to rank subjects with correlation coefficients greater than +0.75 as ‘very good’, +0.5 to +0.75 as ‘good’, +0.25 to +0.5 as ‘medium’, and less than +0.25 (including all negative values) as poor. Cells are colour coded to make this clear.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
1	Scores, all luminaires																				Average cell scores			
2	Metric	A1	A2	B	C	D	E	F	G1	G2	G3	H	I	J	K	L	M	N	O	P	Select	ALL		
3	CIELUV	75	97	95	127	58	30	38	50	41	11	77	42	49	79	59	24	20	15	16	0.98	0.98		
4	CIELAB	56	25	86	108	30	16	40	55	17	-13	27	18	29	30	47	-4	-10	-5	3	0.61	0.54		
5	CMC	92	95	72	118	51	34	65	1	49	56	95	62	56	70	63	14	35	-1	19	0.98	1.02		
6	CIE1994	101	74	99	119	74	45	71	53	63	59	94	60	59	66	71	32	50	45	26	1.22	1.23		
7	CIEDE2000	133	108	117	121	47	54	84	99	70	89	111	80	59	75	82	45	60	43	39	1.48	1.48		
8	Corr Coeff	0.455	0.004	0.659	0.593	0.158	0.590	0.685	0.542	0.136	0.256	0.803	0.567	0.656	0.503	0.563	0.829	0.891	0.475	0.527				
9	Rating	Med	Poor	Good	Good	Poor	Good	Good	Good	Poor	Poor	VG	Good	Good	Good	Good	VG	VG	Med	Good				
10	Sex	M	M	F	M	M+F	M	M	M	M	M	M	F	F	M	M	M	M	F	M				
11	Display	Plasma	LCD	Plasma	Plasma	LCD	LCD	LCD	LCD	LCD	LCD	Plasma	Plasma	LCD	Plasma	LCD	LCD	LCD	LCD	LCD				
12	Type	TV	TV	TV	TV	TV	Mon	Mon	Mon	Mon	Mon	TV	TV	Mon	TV	Mon	Mon	Mon	Mon	Mon				

Figure 12: All-subject analysis, inter-patch scores

Some of this ranking is due to the type of display used. Subjects A, B, C, H, I and K used consumer television displays; the others used professional broadcast monitors. A2 (subject A, second test) and D both score poorly on this basis, probably due to the inconsistent colouring experienced in consumer LCD displays due to their limited viewing angle. E to G, J, and L to P all used LED-backlit LCD professional broadcast monitors in controlled broadcast environments, the others were all in domestic environments.

Figure 13 shows the average inter-patch scores for all three luminaires and all subjects (cells W3 to W7) and for only those qualified as ‘medium’ to ‘very good’. The trend is obvious; CIEDE2000 outperforms the other metrics, with CIELAB worst.

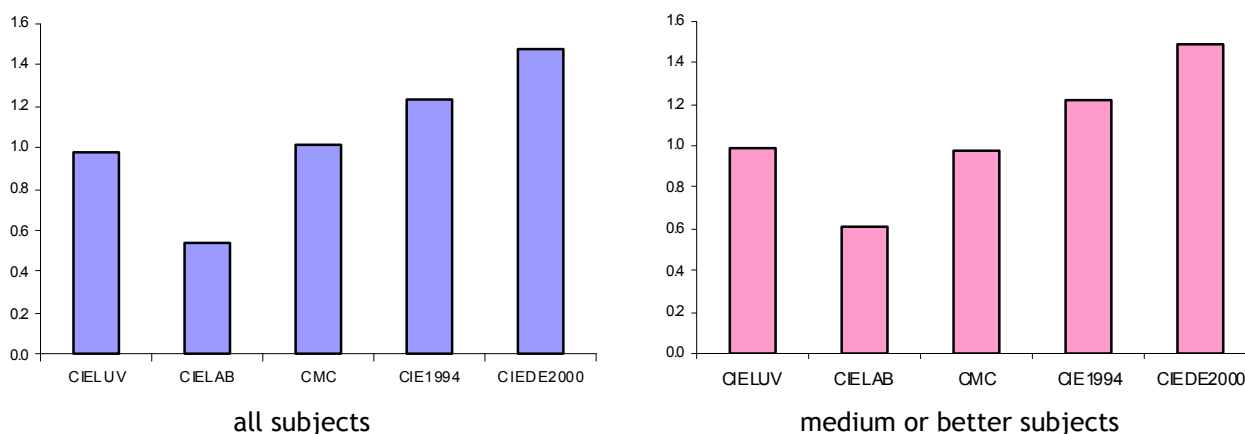


Figure 13: Inter-Patch cores for all luminaires

Since it has already been established that assessment of the small ΔE^* errors for the high-quality LED luminaire was difficult, it makes sense to exclude that luminaire from further analysis.

Figure 14 similarly shows the average inter-patch scores for the single-colour white LED and the fluorescent tube, for all subjects (cells W3 to W7) and for only those qualified as ‘medium’ to ‘very good’. Here it is very obvious that CIELAB does not perform well, although CIEDE2000 is only

marginally better when the ΔE^* errors are smaller, as for the fluorescent tube.

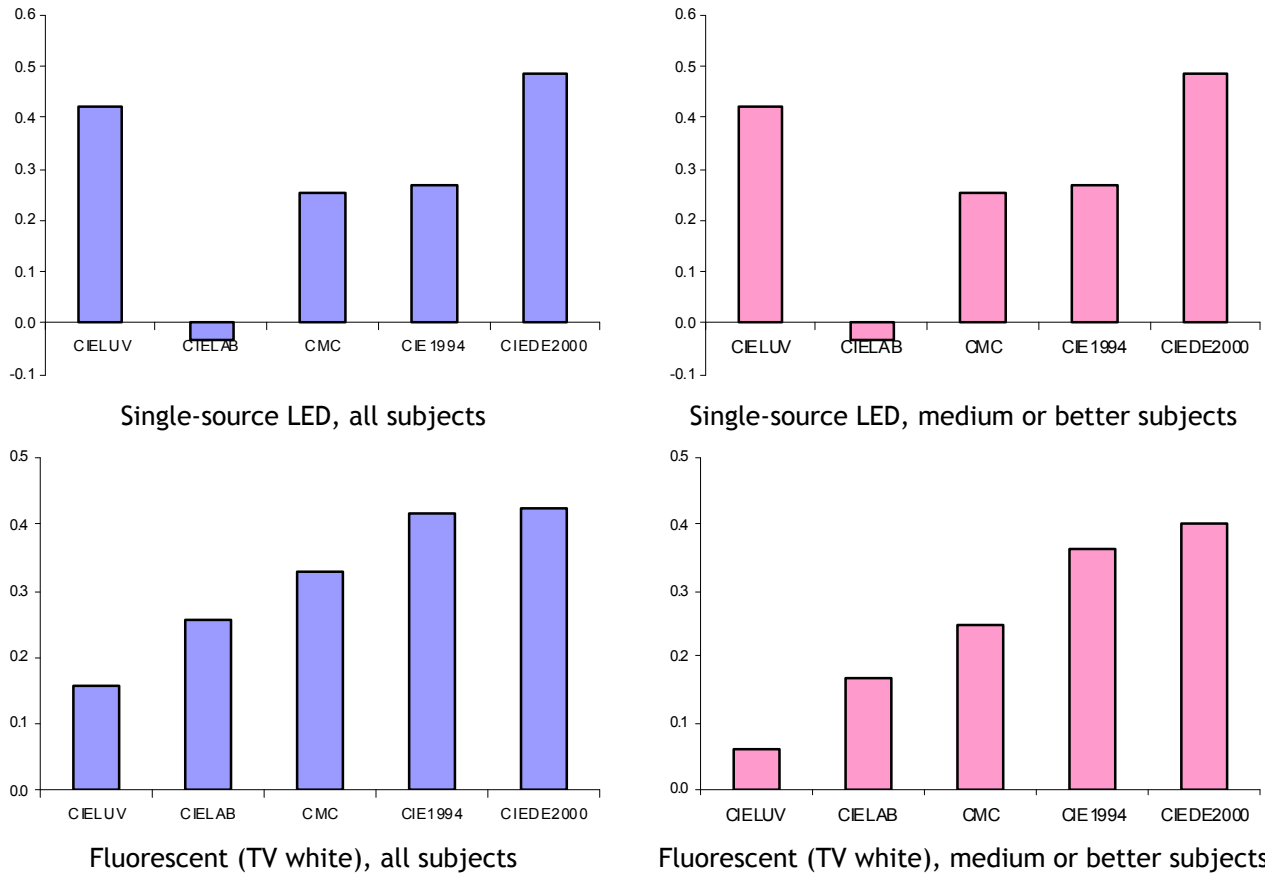


Figure 14: Inter-patch scores for (top) single-source LED and (bottom) fluorescent (TV white)

Next, the correlation coefficients (cells B8 to T8) were used as weighting parameters for the inter-patch scores for all subjects and all luminaires.

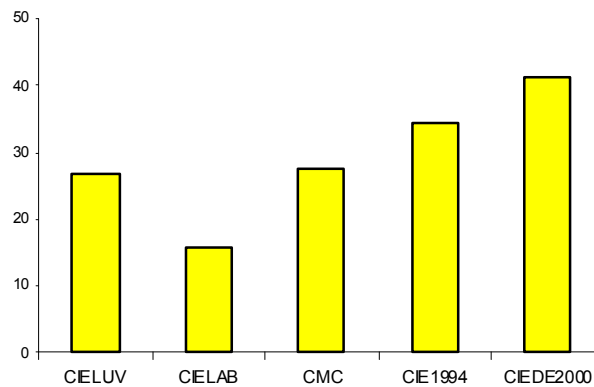


Figure 15: Inter-patch scores weighted by correlation coefficients, for all subjects and all luminaires

Again, CIEDE2000 performs best, CIELAB worst.

Next, the observations with the lowest correlation coefficient, below +0.25, were excluded, and Figure 16 shows the inter-patch scores, weighted by correlation coefficients, for the single-colour white LED and the fluorescent tube. Once more, CIEDE2000 performs best, although CIELUV performs surprisingly well for assessment of the largest ΔE^* errors, the single-colour LED.

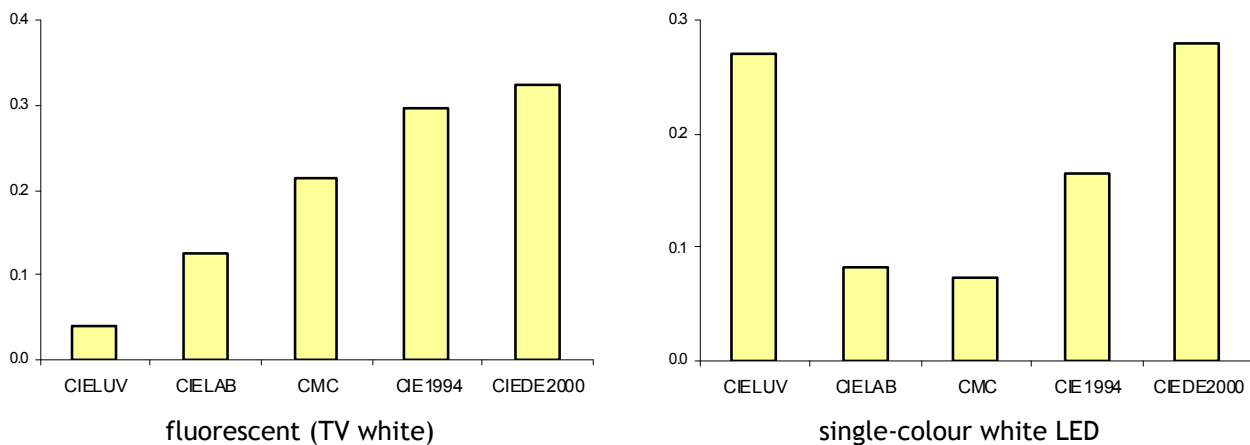


Figure 16: Inter-patch scores weighted by correlation coefficients, excluding 'poor' observations

Clearly it is difficult to displace CIEDE2000 from the top-scoring position and CIELAB from the bottom, however the data is processed. CIELUV performs better than expected when the visible errors are large, but not when they are smaller. Therefore, CIEDE2000 was chosen for use in the Television Lighting Consistency Index, TLCI-2012.

5. Bibliography

- | | | |
|--|---|--|
| Sproson, W. N. & Taylor, E. W. (1971) | A colour television illumination consistency index. | <i>BBC Research Department Report 1971-45</i> |
| Taylor, E. W. (1982) | Rank order difference analysis applied to tests of the Television Consistency Index | <i>BBC Research Department Report 1982-10</i> |
| Taylor, E. W. (1988) | The assessment of the colorimetric properties of light sources for use in television scene lighting | <i>BBC Research Department Report 1988-2</i> |
| CIE, 1995 | Method of measuring and specifying colour rendering properties of light sources (Ed. 2) | <i>Publication CIE No. 13.3 (TC-3.3), Bureau Central de la CIE, Paris, 1995.</i> |
| Roberts, A., Emmett, J. & Böhler, P (2011) | A Television Lighting Consistency Index | <i>Proceedings, IBC 2011, Amsterdam</i> |
| Roberts, A. (2011) | Why does it take so long to get a new standard adopted? | <i>The Iris, Creative Video Productions Ltd, 3rd Quarter 2011</i> |
| CIE, 2004 | Colorimetry | <i>Publication No. 15.3, Bureau Central de la CIE, Paris, 2004.</i> |
| McCamy, C. S., Marcus, H. & Davidson, J. G. | A color-rendition chart | <i>Journal of Applied Photographic Engineering, Vol. 2, #3, Summer 1976.</i> |
| CIE, 2001 | Improvement to industrial colour-difference evaluation | <i>CIE Publication No. 142-2001, Bureau Central de la CIE, Paris, 2001</i> |
| EBU, 2012 | Television Lighting Consistency Index-2012 | <i>EBU Tech 3355</i> |
| Roberts, A. | Circles of Confusion | <i>EBU, 2009</i> |
| Wyszecki, G. & Styles, W. S. (1982) | Color Science: Concepts and Methods, Quantitative Data and Formulae, (Ed. 2) | <i>John Wiley & sons, 2008</i> |
| Reinhard, E., Khan, E. A., Akyüz, A. O. & Johnson, G. M. | Color imaging, fundamentals and applications | <i>A. K.Peters, 2008</i> |

Appendix 1: Reflectivities of Test luminaires

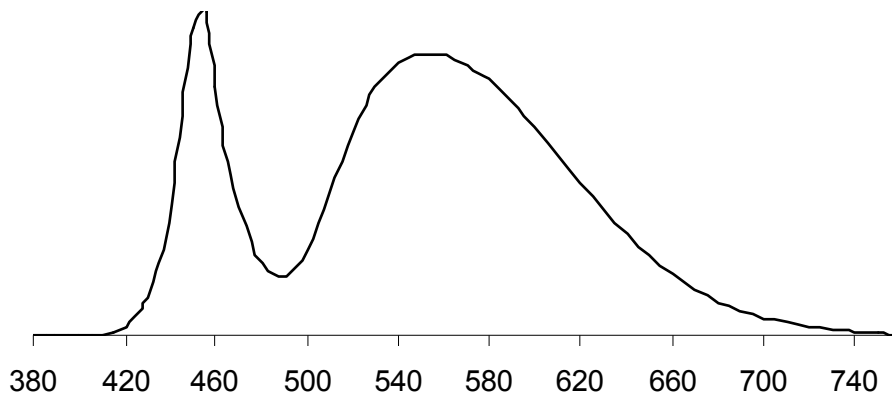


Figure A1.1: Generic, single-colour, white LED luminaire

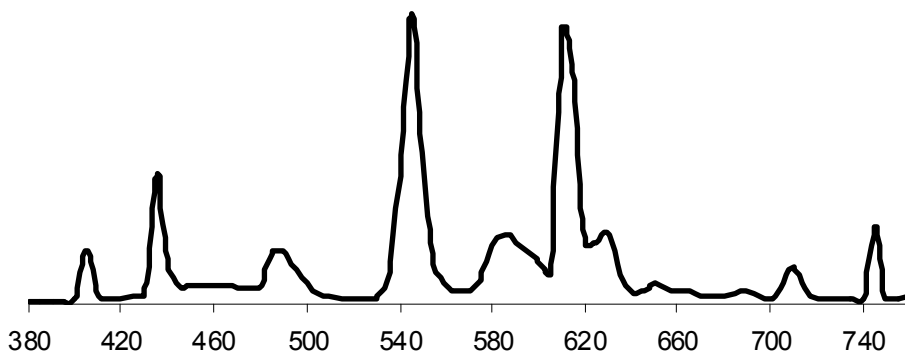


Figure A1.2: Fluorescent, 'TV white'

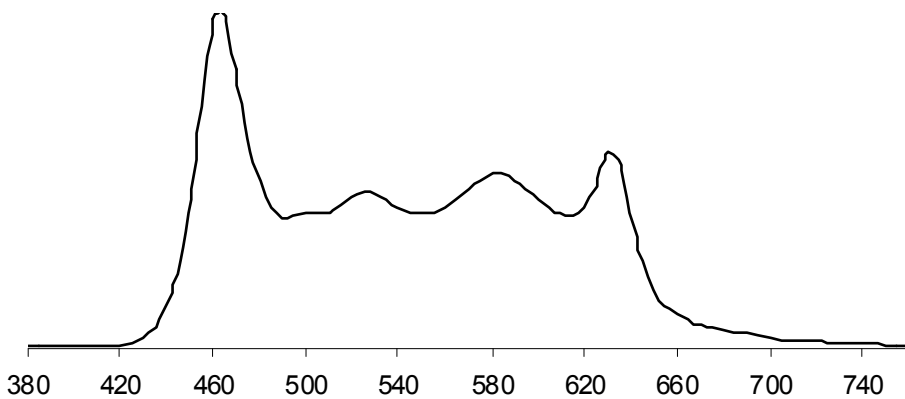


Figure A1.3 High-quality, multi-source LED luminaire

*

* Page intentionally left blank. This document is paginated for two sided printing

Appendix 2: CIE1964

This process was used in the original, 1971, BBC Report on the development of the Television Illumination Consistency Index, and is used in the Colour Rendering Index.

Note that in this method the values for XYZ are scaled to 100 (i.e. $Y = 100$ for peak white).

- First derive CIE1960 chromaticity values for a test colour for illumination by the Test source (subscript T) and the Reference source (subscript R). Subscript W refers to the white point.

$$u_T = \frac{4X_T}{X_T + 15Y_T + 3Z_T}, \quad v_T = \frac{6Y_T}{X_T + 15Y_T + 3Z_T}$$

$$u_R = \frac{4X_R}{X_R + 15Y_R + 3Z_R}, \quad v_R = \frac{6Y_R}{X_R + 15Y_R + 3Z_R}$$

- Calculate CIE1964 values for the Test- and Reference-illuminated colour samples:

$$W_T^* = 25(Y_T)^{1/3} - 17, \quad W_R^* = 25(Y_R)^{1/3} - 17$$

$$U_T^* = W_T^*(u_T - u_W), \quad U_R^* = W_R^*(u_R - u_W)$$

$$V_T^* = W_T^*(v_T - v_W), \quad V_R^* = W_R^*(v_R - v_W)$$

Where u_W and v_W are the chromaticity coordinates of the display white point.

- Calculate the overall colour difference for the test colour:

$$\text{Overall difference: } \Delta E_{64}^* = \sqrt{(13(U_T^* - U_R^*))^2 + (13(V_T^* - V_R^*))^2 + (W_T^* - W_R^*)^2}$$

Note: CIE1964 does not calculate Chroma or Hue, both of which are required for use in the TLCI-2012 calculations.

*

* Page intentionally left blank. This document is paginated for two sided printing

Appendix 3: CIELUV

Note that, in this method, the values for XYZ are scaled to 100 (i.e. $Y = 100$ for peak white).

- First derive CIE1976 chromaticity values for a test colour for illumination by the Test source (subscript T) and the Reference source (subscript R). Subscript W refers to the white point.

$$u'_T = \frac{4X_T}{X_T + 15Y_T + 3Z_T}, \quad v'_T = \frac{9Y_T}{X_T + 15Y_T + 3Z_T}$$

$$u'_R = \frac{4X_R}{X_R + 15Y_R + 3Z_R}, \quad v'_R = \frac{9Y_R}{X_R + 15Y_R + 3Z_R}$$

- Calculate CIELUV values for the Test- and Reference-illuminated colour samples:

$$L^*_T = 116f(Y_T/Y_W) - 16, \quad L^*_R = 116f(Y_R/Y_W) - 16$$

Where $f(x) = \frac{1}{3} \left(\frac{116}{24} \right) x + \frac{16}{116}$ for $x < \left(\frac{24}{116} \right)^3$, otherwise $f(x) = (x)^3$

$$u^*_T = 13L^*_T(u'_T - u'_W), \quad u^*_R = 13L^*_R(u'_R - u'_W)$$

$$v^*_T = 13L^*_T(v'_T - v'_W), \quad v^*_R = 13L^*_R(v'_R - v'_W)$$

Where Y_W , u'_W and v'_W are the luminance and chromaticity coordinates of the display white point in CIE1976 values.

$$C^*_T = \sqrt{u^{*2}_T + v^{*2}_T}, \quad C^*_R = \sqrt{u^{*2}_R + v^{*2}_R}$$

$$h_T = \arctan\left(\frac{v^*_T}{u^*_T}\right), \quad h_R = \arctan\left(\frac{v^*_R}{u^*_R}\right)$$

Note: the hue angle h is measured in degrees.

- Calculate the overall colour difference for the test colour:

$$\Delta E^*_{LUV} = \sqrt{(L^*_T - L^*_R)^2 + (u^*_T - u^*_R)^2 + (v^*_T - v^*_R)^2}$$

- Calculate individual Lightness, Chroma and Hue differences for the test colour:

$$\Delta L^*_{LUV} = L^*_T - L^*_R$$

$$\Delta C^*_{LUV} = 13 \left(L^*_T \sqrt{(u'_T - u'_W)^2 + (v'_T - v'_W)^2} - L^*_R \sqrt{(u'_R - u'_W)^2 + (v'_R - v'_W)^2} \right)$$

$$\Delta H^*_{LUV} = 2 \sin\left(\frac{h_T - h_R}{2}\right) \sqrt{C^*_T \cdot C^*_R}$$

*

* Page intentionally left blank. This document is paginated for two sided printing

Appendix 4: CIELAB

Developed and adopted by the CIE in 1976, this process provides the basis for all later developments in colour-difference measurement.

- Calculate CIELAB values for the Test- and Reference-illuminated colour samples:

$$L^*_T = 116f(Y_T/Y_W) - 16, \quad L^*_R = 116f(Y_R/Y_W) - 16$$

Where $f(x) = \frac{1}{3} \left(\frac{116}{24} \right) x + \frac{16}{116}$ for $x < \left(\frac{24}{116} \right)^3$, otherwise $f(x) = (x)^3$

$$a^*_T = 500 \left(f \left(\frac{X_T}{X_W} \right) - f \left(\frac{Y_T}{Y_W} \right) \right), \quad a^*_R = 500 \left(f \left(\frac{X_R}{X_W} \right) - f \left(\frac{Y_R}{Y_W} \right) \right)$$

$$b^*_T = 200 \left(f \left(\frac{Y_T}{Y_W} \right) - f \left(\frac{Z_T}{Z_W} \right) \right), \quad b^*_R = 200 \left(f \left(\frac{Y_R}{Y_W} \right) - f \left(\frac{Z_R}{Z_W} \right) \right)$$

Where the $X_w Y_w Z_w$ are the tristimulus values of the display white point.

$$C^*_T = \sqrt{a^{*2}_T + b^{*2}_T}, \quad C^*_R = \sqrt{a^{*2}_R + b^{*2}_R}$$

$$h_T = \arctan \left(\frac{b^*_T}{a^*_T} \right), \quad h_R = \arctan \left(\frac{b^*_R}{a^*_R} \right)$$

Note: the hue angle h is measured in degrees.

- Calculate the overall colour difference for the test colour:

$$\Delta E^*_{LAB} = \sqrt{(L^*_T - L^*_R)^2 + (a^*_T - a^*_R)^2 + (b^*_T - b^*_R)^2}$$

- Calculate individual Lightness, Chroma and Hue differences for the test colour:

$$\Delta L^*_{LAB} = L^*_T - L^*_R$$

$$\Delta C^*_{LAB} = C^*_T - C^*_R$$

$$\Delta H^*_{LAB} = 2 \sin \left(\frac{h^*_T - h^*_R}{2} \right) \sqrt{C^*_T \cdot C^*_R}$$

*

* Page intentionally left blank. This document is paginated for two sided printing

Appendix 5: CMC

Developed by the Society of Dyers and Colourists and adopted by the CIE in 1988, this process is generally thought to be an improvement on earlier methods. It subdivides the CIELAB colour space into an infinite set of ellipsoidal microspaces (one for each test colour considered), the axes of which are tilted and scaled to better match the perception of small colour differences.

- Calculate CIELAB values for the Test- and Reference-illuminated colour samples:

$$L^*_{T} = 116f(Y_T/Y_W) - 16, \quad L^*_{R} = 116f(Y_R/Y_W) - 16$$

Where $f(x) = \frac{1}{3} \left(\frac{116}{24} \right) x + \frac{16}{116}$ for $x < \left(\frac{24}{116} \right)^3$, otherwise $f(x) = (x)^3$

$$a^*_{T} = 500 \left(f \left(\frac{X_T}{X_W} \right) - f \left(\frac{Y_T}{Y_W} \right) \right), \quad a^*_{R} = 500 \left(f \left(\frac{X_R}{X_W} \right) - f \left(\frac{Y_R}{Y_W} \right) \right)$$

$$b^*_{T} = 200 \left(f \left(\frac{Y_T}{Y_W} \right) - f \left(\frac{Z_T}{Z_W} \right) \right), \quad b^*_{R} = 200 \left(f \left(\frac{Y_R}{Y_W} \right) - f \left(\frac{Z_R}{Z_W} \right) \right)$$

Where the X_w Y_w Z_w are the tristimulus values of the display white point.

$$C^*_{T} = \sqrt{a^*_{T}{}^2 + b^*_{T}{}^2}, \quad C^*_{R} = \sqrt{a^*_{R}{}^2 + b^*_{R}{}^2}$$

$$h_T = \arctan \left(\frac{b^*_{T}}{a^*_{T}} \right), \quad h_R = \arctan \left(\frac{b^*_{R}}{a^*_{R}} \right)$$

Note: the hue angle h is measured in degrees.

- Calculate individual Lightness, Chroma and Hue differences for the test colour:

$$\Delta L^*_{LAB} = L^*_{T} - L^*_{R}$$

$$\Delta C^*_{LAB} = C^*_{T} - C^*_{R}$$

$$\Delta H^*_{LAB} = 2 \sin \left(\frac{h^*_{T} - h^*_{R}}{2} \right) \sqrt{C^*_{T} \cdot C^*_{R}}$$

- Calculate scalars for the ellipsoid semi-axes for the colour under test, in each case, the values are those for the average colorimetric values:

$$S_L = \frac{0.040975 L^*}{1 + 0.01765 L^*}, \quad \text{unless } L^* < 16, \text{ when } S_L = 0.511$$

$$S_C = \frac{0.0638 C^*}{1 + 0.0131 C^*} + 0.638$$

$$F = \sqrt{\frac{C^{*4}}{C^{*4} + 1900}}$$

$$T = 0.36 + abs(0.4 \cos(35 + h^*)) \text{ unless } h^* > 164 \text{ and } h^* < 345, \\ \text{when } T = 0.56 + abs(0.2 \cos(168 + h^*))$$

$$S_H = S_C(FT + 1 - F)$$

- Calculate the resulting difference value:

$$\text{Overall difference: } \Delta E^*_{CMC} = \sqrt{\left(\frac{\Delta L^*_{LAB}}{lS_L}\right)^2 + \left(\frac{\Delta C^*_{LAB}}{cS_c}\right)^2 + \left(\frac{\Delta H^*_{LAB}}{S_H}\right)^2}$$

Where the values l and c are weighting factors, usually both unity.

They can be varied to customise the metric, changing the effects of lightness and chroma respectively.

If they are varied from unity then their values should be included in the name of the system, e.g. CMC(2:1).

For the subjective tests, $l = c = 1$.

Appendix 6: CIE1994

This is an elaboration of the process defined in CMC. It is intended for D65 viewing, and therefore is appropriate for television use. The ΔE^* value is intended to be matched to human vision, in that a value of 1 is the threshold of discriminability.

- Calculate CIELAB values for the Test- and Reference-illuminated colour samples:

$$L^*_{T} = 116f(Y_T/Y_W) - 16, \quad L^*_{R} = 116f(Y_R/Y_W) - 16$$

Where $f(x) = \frac{1}{3} \left(\frac{116}{24} \right) x + \frac{16}{116}$ for $x < \left(\frac{24}{116} \right)^3$, otherwise $f(x) = (x)^3$

$$a^*_{T} = 500 \left(f \left(\frac{X_T}{X_W} \right) - f \left(\frac{Y_T}{Y_W} \right) \right), \quad a^*_{R} = 500 \left(f \left(\frac{X_R}{X_W} \right) - f \left(\frac{Y_R}{Y_W} \right) \right)$$

$$b^*_{T} = 200 \left(f \left(\frac{Y_T}{Y_W} \right) - f \left(\frac{Z_T}{Z_W} \right) \right), \quad b^*_{R} = 200 \left(f \left(\frac{Y_R}{Y_W} \right) - f \left(\frac{Z_R}{Z_W} \right) \right)$$

Where the X_w Y_w Z_w are the tristimulus values of the display white point.

$$C^*_{T} = \sqrt{a^{*2}_{T} + b^{*2}_{T}}, \quad C^*_{R} = \sqrt{a^{*2}_{R} + b^{*2}_{R}}$$

$$h_T = \arctan \left(\frac{b^*_{T}}{a^*_{T}} \right), \quad h_R = \arctan \left(\frac{b^*_{R}}{a^*_{R}} \right)$$

$$\Delta L^*_{LAB} = L_T - L_R$$

Note: the hue angle h is measured in degrees.

- Calculate scalars for the ellipsoid semi-axes for the colour under test, in each case, the values are those for the average colorimetric values:

$$S_L = 1$$

$$S_C = 1 + 0.045 \sqrt{C^*_{T} \cdot C^*_{R}}$$

$$S_H = 1 + 0.015 \sqrt{C^*_{T} \cdot C^*_{R}}$$

- Calculate individual Lightness, Chroma and Hue differences for the test colour:

$$\Delta L^*_{94} = L^*_{T} - L^*_{R}$$

$$\Delta C^*_{94} = C^*_{T} - C^*_{R}$$

$$\Delta H^*_{94} = 2 \sin \left(\frac{h^*_{T} - h^*_{R}}{2} \right) \sqrt{C^*_{T} \cdot C^*_{R}}$$

- Calculate the resulting difference value:

$$\Delta E^*_{94} = \sqrt{\left(\frac{\Delta L^*_{94}}{k_L S_L}\right)^2 + \left(\frac{\Delta C^*_{94}}{k_C S_c}\right)^2 + \left(\frac{\Delta H^*_{94}}{k_H S_H}\right)^2}$$

Where the k values are usually unity, but can be varied to customise the calculation.

If the values of k are varied from unity, then their values should be included in the system notation, e.g. CIE1994(2:1:1).

For the subjective tests, $k_L = k_C = k_H = 1$.

Appendix 7: CIEDE2000

This is a further refinement of CIELAB, intended to improve the correlation between the metric and human vision in the saturated blue and near neutral colour regions.

- Calculate CIELAB values for the Test- and Reference-illuminated colour samples:

$$L^*_T = 116f(Y_T/Y_W) - 16, \quad L^*_R = 116f(Y_R/Y_W) - 16$$

Where $f(x) = \frac{1}{3} \left(\frac{116}{24} \right) x + \frac{16}{116}$ for $x < \left(\frac{24}{116} \right)^3$, otherwise $f(x) = (x)^3$

$$a^*_T = 500 \left(f \left(\frac{X_T}{X_W} \right) - f \left(\frac{Y_T}{Y_W} \right) \right), \quad a^*_R = 500 \left(f \left(\frac{X_R}{X_W} \right) - f \left(\frac{Y_R}{Y_W} \right) \right)$$

$$b^*_T = 200 \left(f \left(\frac{Y_T}{Y_W} \right) - f \left(\frac{Z_T}{Z_W} \right) \right), \quad b^*_R = 200 \left(f \left(\frac{Y_R}{Y_W} \right) - f \left(\frac{Z_R}{Z_W} \right) \right)$$

Where the $X_w Y_w Z_w$ are the tristimulus values of the display white point.

$$C^*_T = \sqrt{a^{*2}_T + b^{*2}_T}, \quad C^*_R = \sqrt{a^{*2}_R + b^{*2}_R}$$

$$h_T = \arctan \left(\frac{b^*_T}{a^*_T} \right), \quad h_R = \arctan \left(\frac{b^*_R}{a^*_R} \right)$$

Note: the hue angle h is measured in degrees.

- Calculate non-linear modifiers for the actual colour under test:

$$L'_T = L^*_T, \quad L'_R = L^*_R$$

$$\overline{C^*} = \frac{C^*_T + C^*_R}{2}$$

$$g = \frac{1}{2} \left(1 - \sqrt{\frac{\overline{C^*}^7}{\overline{C^*}^7 + 25^7}} \right)$$

$$C'_T = \sqrt{(1+g)a^{*2}_T + b^{*2}_T}, \quad C'_R = \sqrt{(1+g)a^{*2}_R + b^{*2}_R}$$

$$h'_T = \arctan \left(\frac{b^*_T}{(1+g)a^*_T} \right), \quad h'_R = \arctan \left(\frac{b^*_R}{(1+g)a^*_R} \right)$$

Note: h^* is measured in degrees, not radians. The signs of a^* and b^* must be observed in order to generate an angle within the range 0 - 360 degrees.

$$\bar{L}' = \frac{L'_T + L'_R}{2}, \quad \bar{C}' = \frac{C'_T + C'_R}{2}, \quad \bar{h}' = \frac{h'_T + h'_R}{2}$$

$$R_C = 2\sqrt{\frac{\bar{C}'^7}{\bar{C}'^7 + 25^7}}$$

$$\Delta\Theta = 30 \exp\left(-\left(\frac{\bar{h}' - 275}{25}\right)^2\right)$$

$$R_T = -R_C \sin(2\Delta\Theta)$$

$$T = 1 - 0.17 \cos(\bar{h}' - 30) + 0.24 \cos(2\bar{h}') + 0.32 \cos(3\bar{h}' + 6) - 0.2 \cos(4\bar{h}' - 63)$$

$$S_L = 1 + \frac{0.015(\bar{L}' - 50)^2}{\sqrt{20 + (\bar{L}' - 50)^2}}$$

$$S_C = 1 + 0.045\bar{C}'$$

$$S_H = 1 + 0.015\bar{C}'T$$

- Calculate the lightness, chroma and hue differences for each colour:

$$\Delta L^*_{2000} = L'_T - L'_R$$

$$\Delta C^*_{2000} = C'_T - C'_R$$

$$\Delta H^*_{2000} = 2 \sin\left(\frac{h'_T - h'_R}{2}\right) \sqrt{C'_T \cdot C'_R}$$

- Calculate the resulting difference value:

$$\text{Overall difference: } \Delta E^*_{2000} = \sqrt{\left(\frac{\Delta L^*_{2000}}{k_L S_L}\right)^2 + \left(\frac{\Delta C^*_{2000}}{k_C S_C}\right)^2 + \left(\frac{\Delta H^*_{2000}}{k_H S_H}\right)^2} + R_T \left(\frac{\Delta C'}{k_C S_C}\right) \left(\frac{\Delta H'}{k_H S_H}\right)$$

Where the k values are weighting factors, usually unity.

They can be varied to customise the metric, affecting lightness, saturation and hue, respectively.

If they are varied from unity, then their values should be included in the name of the system, e.g. CIEDE2000(2:1:1).

For the subjective tests, $k_L = k_C = k_H = 1$.