Inter–Relating Colour Difference Metrics

Pei–Li Sun and Ján Morovic
Colour & Imaging Institute, University of Derby
Derby, United Kingdom

Abstract

A method for predicting statistics of colour differences computed using a ΔE metric given a value of a different ΔE metric is introduced in this paper. The result of this method is the ability to determine what range of values of one metric can be expected for a given ΔE value computed in terms of another metric. This in turn allows for inter–comparison of results from different studies even if these reported their findings in terms of various colour difference formulae. This paper also illustrates the use of the inter–comparison method proposed in it and shows that the method works with very high levels of accuracy and that it can be applied to a range of colour difference metrics including even Euclidean distance in device–dependent colour spaces. The effect of applying advanced colour difference formulæ in CAM97s2 space is also illustrated and a close relationship with their original CIELAB–based versions is shown.

Introduction

Colour difference (ΔE) metrics are a powerful and frequently used tool for expressing the difference in colour appearance of a pair of colours or of sets of colour pairs. Given this function ΔE metrics can be used for a very wide range of applications. First, they can quantify the performance of a range of physical as well as computational systems like determining the repeatability of measuring instruments, the uniformity of imaging media or the accuracy of device characterisation models. Second, they can express the effect of some operation on a set of colours – e.g., the magnitude of colour change due to gamut mapping. Third, they can also measure the behaviour of the visual system for a range of tasks – e.g., the perceptibility threshold of colour differences between complex images.

Furthermore, by having uses including all the above ones they also allow for relating values from one application with those from another. For example, knowing the uniformity of a display, it can be seen whether the accuracy of the characterisation model is of similar magnitude as this would mean that errors from it would be similar to those inherent in the system. It can then also be known whether these characterisation errors would be visible, given information about perceptibility thresholds.

While the picture presented so far is one where there are no difficulties, there is one significant complication that needs to be taken into account. Namely that there are a large number of different colour difference metrics which can also be performed in a number of colour spaces. As such the above kind of synergy between different studies can only be had if they all express their findings in terms of the same colour difference metric. If, for example, perceptibility thresholds are known in terms of CIE ΔE^*_{ab} [1] and the uniformity of a display is expressed in CIE ΔE_{94} [2] then no direct comparison can be made between the two results. The simplest solution clearly is to re–compute colour differences for both sets of data (i.e. in this case perceptibility and display uniformity) and then to make the comparison. This, however, is often not possible as one might only have the ΔE values rather than the pair data for which they were computed.

The question therefore arises of whether it is possible to determine what value of colour difference metric ΔE_1 corresponds to a given value of another colour difference metric ΔE_2. In other words, given mappings from 3D into two different 1D spaces is it possible to predict one 1D mapping from the other? The answer to this is 'no', as such a task is clearly underdetermined. However, what is possible is to predict what range of ΔE_1 values corresponds to a given ΔE_2 value and vice versa whereby the smaller ΔE_2 is the narrower a range of ΔE_1 values can be expected. For $\Delta E_2=0$, ΔE_1 will also be zero and the greater ΔE_2 gets the more potential there is for two different metrics to map the 3D difference in different ways.

This paper will therefore present a framework for predicting intervals of ΔE values in one metric given a ΔE value in another metric and results of applying the method to frequently used metrics will be shown. Given this method it will be possible to compute corresponding ΔE intervals and as such compare findings from different studies even if they used different ΔE metrics.

Method

The method used here for inter–relating the results of two different ΔE metrics – ΔE_1 and ΔE_2 – is based on using a least-squares polynomial fitting technique for predicting the statistics of one from the other. The data on which the fitting will be based consists of a large number of pairs of random XYZ samples. To avoid errors introduced by extreme XYZ values and large ΔE values which might never be encountered in practice, in this study only colours from within an *Apple Studio Display* CRT's gamut will be used. Furthermore only XYZ pairs with ΔE_{97s2} values under 70 will be considered in the fitting as differences above this kind of threshold will result impractically large in ΔE ranges. Note that ΔE_{97s2} is Euclidean distance in CAM97s2[3] space, which is a modification of CIECAM97s.[4]

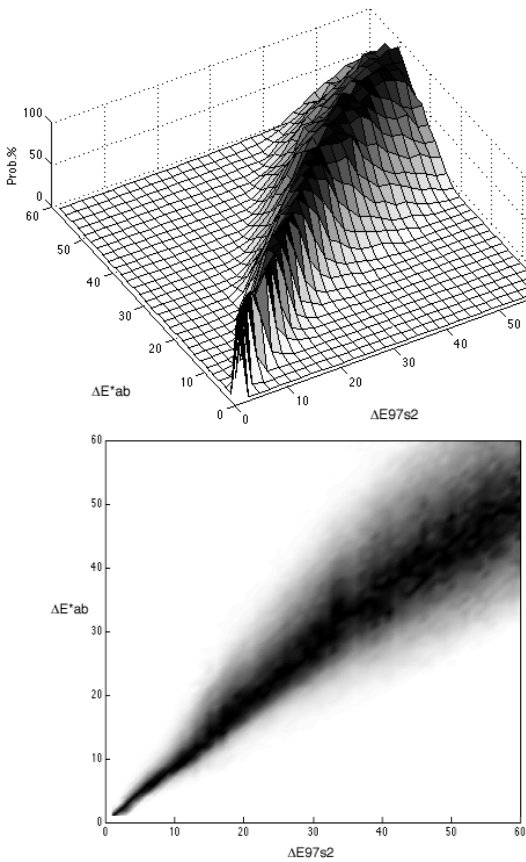

Figure 1. ΔE_2 (ΔE^*_{ab}) versus ΔE_1 (ΔE_{97s2}) LUT. White represents zero and black represents one.

The process for generating a function for predicting a statistic of ΔE_2 for a given ΔE_1 value (e.g., to predict the mean ΔE_{94} corresponding to a given ΔE_{97s2} value) is as follows:

1. Set–up a 2D histogram, which will be populated to indicate the frequency of each ΔE_1 versus ΔE_2 combination occurring. The interval for each bin in the following test was set to one ΔE unit for both ΔE_1 and ΔE_2 metrics.
2. Randomly generate monitor RGB value pairs and using a forward CRT characterisation model obtain corresponding pairs of XYZ values. Note that, as pairs with smaller ΔE_1 values will be relatively fewer, more random colour-pairs that result in smaller ΔE_1 (or ΔE_{RGB}) values need to be provided for the test.
3. For each pair of XYZ values from Step 2 compute values of the ΔE_1 and ΔE_2 metrics using the appropriate transforms.
4. Each pair of ΔE_1 and ΔE_2 values is then used to increment the frequency in the corresponding bin of the 2D histogram from Step 1.
5. After processing all test colour pairs, calculate the statistics of ΔE_2 (e.g., mean, standard deviation) for each ΔE_1 value.
6. Use a least-squares technique to fit ΔE_1 values to each given ΔE_2 statistic.

To illustrate the data dealt with in the present method, Fig. 1 shows the 3D LUT obtained after step 4 for ΔE_1 and ΔE_2 being ΔE_{97s2} and ΔE^*_{ab} respectively. To show the nature of the data more clearly from the point of view of later extracting statistics of ΔE_2 for each value of ΔE_1, it has been processed so as to make the maximum probability at each ΔE_1 value the same. Looking at the LUT directly would show very large values for low ΔE_1s and much smaller ones for larger ΔE_1 values (Fig. 2) as the sum of probabilities for each ΔE_1 is approximately the same but their ranges are proportional to the magnitude of ΔE_1.

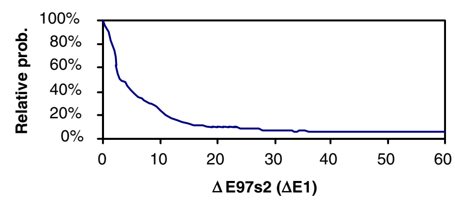

Figure 2. Relative maximum probability as a function of ΔE_1.

Note that, as the maximum ΔE_1 considered here was 70 and the bin interval was one ΔE unit, there were 70 sets of data that were used for the polynomial fitting. Furthermore, as hundreds of thousands of samples are examined in total, each set contained at least one thousand samples and the statistics reported here are therefore based on a sample of significant size.

Results

Three CIE recommended colour difference formulae, ΔE^*_{ab}, ΔE_{94} and ΔE_{2000},[5] were examined against ΔE_{97s2} in the way outlined above. The results, in terms of the coefficients for predicting mean, standard deviation, and 95% ranges (i.e. the ranges of values that for a given ΔE_1 covering 95% of the ΔE_2's distribution) between ΔE_1 and ΔE_2 statistics will be given next.

In the following discussion, the 95% ranges of the considered statistics refer to the ranges between the 2.5th and the 97.5th percentiles of the data. As the distribution of ΔE_2 statistics is not necessarily normal, the 95% range deter-mined in this way will be more accurate than the value one would get by assuming the data to be normally distributed.

ΔE_{97s2} to ΔE^*_{ab}

The mean and 95% range values for the ΔE_{97s2} to ΔE^*_{ab} mapping are shown in Fig. 3. As can be seen, the middle line has a slope close to 45 degrees and this suggests that both ΔE_{97s2} and ΔE^*_{ab} metrics have, on average, got similar units when predicting perceptual colour differences. As expected it can also be seen that the 95% ranges increase as ΔE_{97s2} increases. On the other hand, a significant trend can be seen for all the statistics and this suggests that a second-order polynomial passing through the origin be used for modelling them.

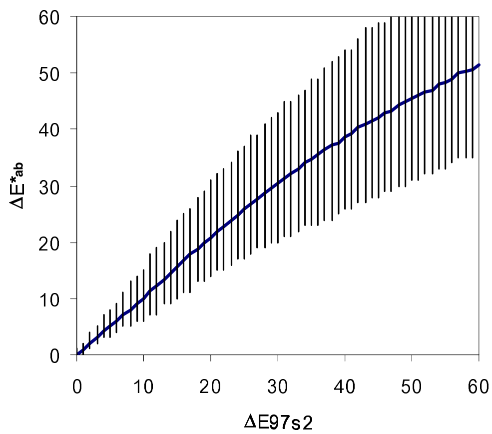

Figure 3. Relationship between ΔE_{97s2} and ΔE^*_{ab}. Vertical lines indicate 95% ranges.

The results of fitting a second-order polynomial function without offset to the relationship between ΔE_{97s2} and ΔE^*_{ab} statistics are shown in the fist row of Table 1. From there it can be seen that there are two coefficients to be optimised for predicting each ΔE^*_{ab} statistic.

For instance, if some characterisation model is found to have a mean error of 3 ΔE_{97s2} units, the predictions of corresponding ΔE^*_{ab} statistics of mean, standard deviation, 2.5th percentile and 97.5th percentile will be 3.3, 0.8, 2.1 and 4.9 respectively. In other words the 3 ΔE_{97s2} difference can be expected to correspond to differences between 2.1 and 4.9 ΔE^*_{ab} 95% of the time and on average to be 3.3 ΔE^*_{ab}.

The accuracy of these predictions can be evaluated by looking at the determination coefficient R^2 between the original data and its predictions. It can be seen from these values in Table 1 that each of them is very close to 1 and this suggests that all the predictions are highly reliable.

Next, two advanced colour difference formulae – ΔE_{94} and ΔE_{2000} – will be compared against ΔE_{97s2}. The mean and the 95% range values for the mapping from ΔE_{97s2} to ΔE_{94} and ΔE_{2000} statistics are shown in Fig. 4. As can be seen, both ΔE_{94} and ΔE_{2000} have similar characteristics of their relationship to ΔE_{97s2}. The slopes of the fits are less than 1, and even the 97.5th percentiles of the fits do not cross the 45-degree lines (shown as dashed lines in Fig. 4).

Table 1. Predictions of ΔE^*_{ab} statistics for given ΔE_{97s2} values (Note, the coefficients apply only when ΔE_{97s2} is below 70).

Function	$\Delta E^*_{ab} = c_1 (\Delta E_{97s2}) + c_2 (\Delta E_{97s2})^2$		
Coefficients	c_1	c_2	R^2
Mean	1.12	-0.0044	0.999
Stdev	0.27	-0.0020	0.997
2.5th percentile	0.70	-0.0016	0.998
97.5th percentile	1.67	-0.0085	0.999
Sample size	376,714 random RGB pairs		

ΔE_{97s2} to ΔE_{94} and ΔE_{2000}

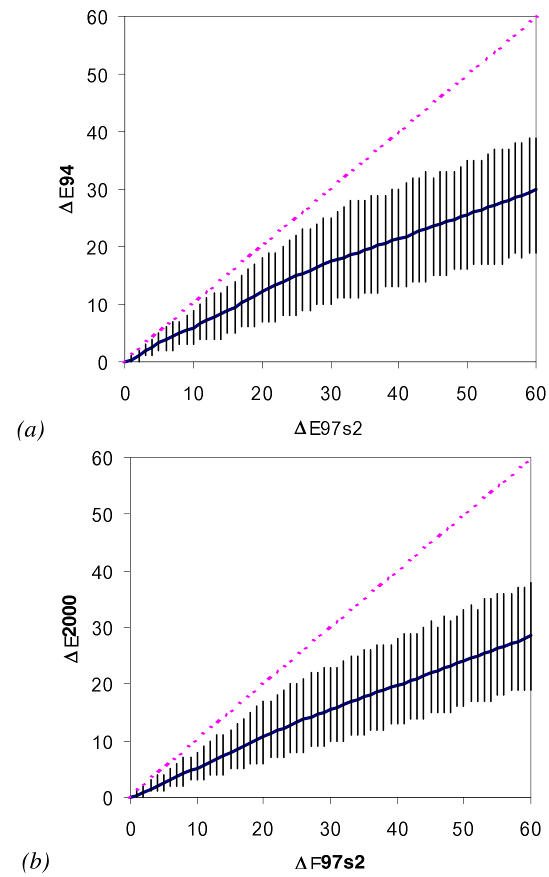

Figure 4. Relationship between (a) ΔE_{97s2} and ΔE_{94} and (b) ΔE_{97s2} and ΔE_{2000}. Vertical lines indicate 95% confidence intervals.

Therefore, one can be confident that at least 95% of colour differences computed in terms of ΔE_{94} and ΔE_{2000} will be smaller than corresponding ΔE_{97s2} values. For example, if there is a characterisation error of 3 ΔE_{97s2} units, one can expect corresponding ΔE_{94} or ΔE_{2000} values to be smaller than 3. The coefficients and R^2 values of the curve fits are shown in Table 2 where the high R^2s again suggest that the predictions are reliable.

Table 2. Predictions of ΔE_{94} and ΔE_{2000} statistics for given ΔE_{97s2} values (Note, that the coefficients are valid only when ΔE_{97s2} is below 70).

Predictions	ΔE_{94}			ΔE_{2000}		
Functions	$\Delta E_{94} = c_1 (\Delta E_{97s2}) + c_2 (\Delta E_{97s2})^2$			$\Delta E_{2000} = c_1 (\Delta E_{97s2}) + c_2 (\Delta E_{97s2})^2$		
Coefficients	c_1	c_2	R^2	c_1	c_2	R^2
Mean	0.63	-0.0023	0.999	0.54	-0.0010	0.999
Stdev	0.17	-0.0014	0.981	0.14	-0.0010	0.989
2.5th perc.	0.34	-0.0003	0.997	0.31	0.0002	0.998
97.5th perc.	0.97	-0.0052	0.997	0.86	-0.0038	0.997
Sample size	179,535 random RGB pairs			179,535 random RGB pairs		

Looking at Fig. 4. it can be seen that the relationship between ΔE_{97s2} is and both of the other advanced colour difference formulae is similar and it is therefore also of interest to see how those two metrics compare with each other. In Fig. 5. it can be seen that they do indeed have a very close agreement and also that the 95% ranges are more narrow than for many other pairs of metrics. The parameters for predicting ΔE_{2000} statistics for ΔE_{94} values are shown in Table 3.

Applying Advanced Formulæ to CAM97s2 Dimensions

Advanced colour difference formulae have been developed as modifications of Euclidean distance in CIELAB space and at present there is no equivalent for them in colour appearance spaces like CIECAM97s or CAM97s2. As the advanced formulae have been shown to perform well in CIELAB, for which they were developed, it would be of use to see whether how similar the predictions they give when applied to CAM97s2 coordinates are to their original behaviour.

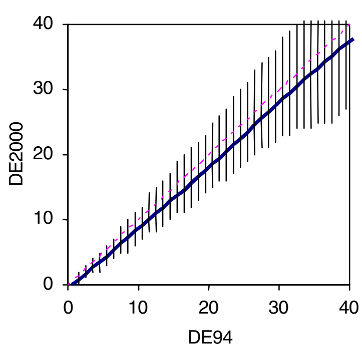

Figure 5. Relationship between ΔE_{94} and ΔE_{2000}.

Table 3. Predictions of ΔE_{2000} statistics for given ΔE_{94} values (Note, the coefficients apply only when ΔE_{94} is below 40).

Function	$\Delta E_{2000} = c_1 (\Delta E_{94}) + c_2 (\Delta E_{94})^2$		
Coefficients	c_1	c_2	R^2
Mean	0.92	0.0003	0.999
Stdev	0.12	0.0005	0.997
2.5th percentile	0.75	-0.0016	0.994
97.5th percentile	1.24	-0.0009	0.999
Sample size	208,592 random RGB pairs		

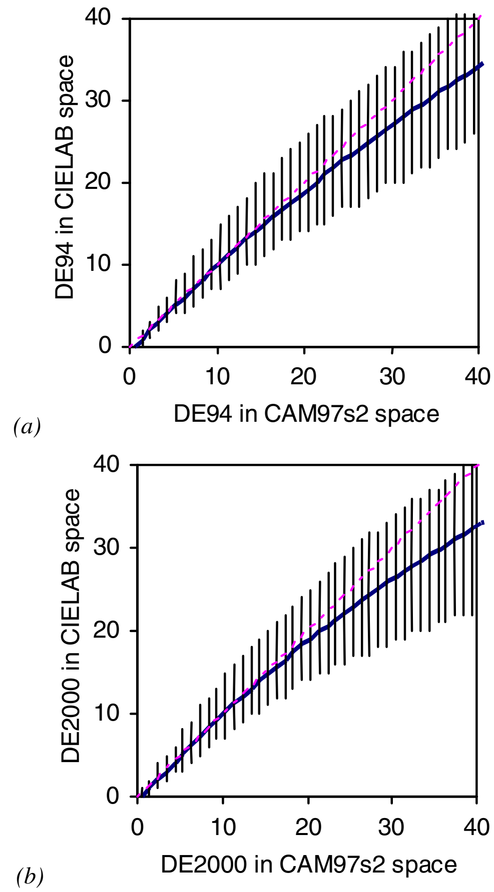

Figure 6. CIELAB versus CAM97s2 based metrics.

ΔE_{94} and ΔE_{2000} have therefore been applied in CAM97s2 space and Fig. 6 shows the relationship between their results in CIELAB and in CAM97s2. As can be seen a CAM97s2–based value has a corresponding mean CIELAB based value that is very similar to itself. Also, the 95% ranges in both these cases are narrower than for the relationship between ΔE_{97s2} and the two original formulæ. Coefficients for predicting statistics of the CIELAB based formulæ from their CAM97s2 equivalents are then shown in Table 4.

Table 4. Predictions of CIELAB–based statistics for given CAM97s2–based ΔE values.

Predictions	ΔE_{94}			ΔE_{2000}		
Functions	$\Delta E_{LAB} = c_1 (\Delta E_{CAM}) + c_2 (\Delta E_{CAM})^2$			$\Delta E_{LAB} = c_1 (\Delta E_{CAM}) + c_2 (\Delta E_{CAM})^2$		
Coefficients	c_1	c_2	R^2	c_1	c_2	R^2
Mean	1.02	-0.0038	1.000	1.03	-0.0050	1.000
Stdev	0.19	-0.0021	0.988	0.19	-0.0018	0.982
2.5^{th} perc.	0.68	-0.0003	0.996	0.76	-0.0044	0.996
97.5^{th} perc.	1.42	-0.0085	0.998	1.45	-0.0106	0.998
Sample size	228,589 random RGB pairs			226,055 random RGB pairs		

ΔE_{RGB} to ΔE_{97s2}

Even though Euclidean distance in a device–dependent RGB space (ΔE_{RGB}) is not a very reliable indicator of perceived colour difference, it is nonetheless of interest to see how it relates to ΔE_{97s2}. In particular as this is the metric in terms of which inverse device characterisation models (i.e. those predicting device RGB from colorimetric data) can most directly be evaluated.

However, to do this, the method described above has to be adapted. The ΔE_1 needs to be replaced by ΔE_{RGB} and the maximum ΔE_{RGB} for the polynomial fitting will be set as to 120 to reflect the difference between ΔE_{RGB} units and the units of colorimetric ΔE metrics.

Three media, CRT monitor (CRT), a printer using either high-resolution paper (hi-res) or plain paper (plain), were tested and the resulting plots with 95% ranges are shown in Figure 7. As can be seen, the slopes of the curves for three media decrease in the above order. The slope which transforms the overall units from ΔE_{RGB} to ΔE_{97s2} implies both the size of the gamut and the perceptual ΔE_{RGB} unit distance for a given medium. The higher slope for the CRT's RGB space is also accompanied by errors and variations of ΔE_{97s2} being larger for this medium than for the other two media in terms of their ΔE_{RGB} units. It is interesting to compare Fig. 7 with Figs. 3 and 4 to see whether device-dependent ΔE_{RGB} metrics have a better relationship with ΔE_{97s2}. However, as expected, the larger 95% ranges in Figure 7 suggest that CIELAB based colour difference formulae have better correlation with the ΔE_{97s2} metric than the ΔE_{RGB} formulae.

The coefficients of the polynomial fitting for ΔE_{97s2} against the ΔE_{RGB}s of the three media are given in Table 5 and the reliability of predictions is again shown to be very high.

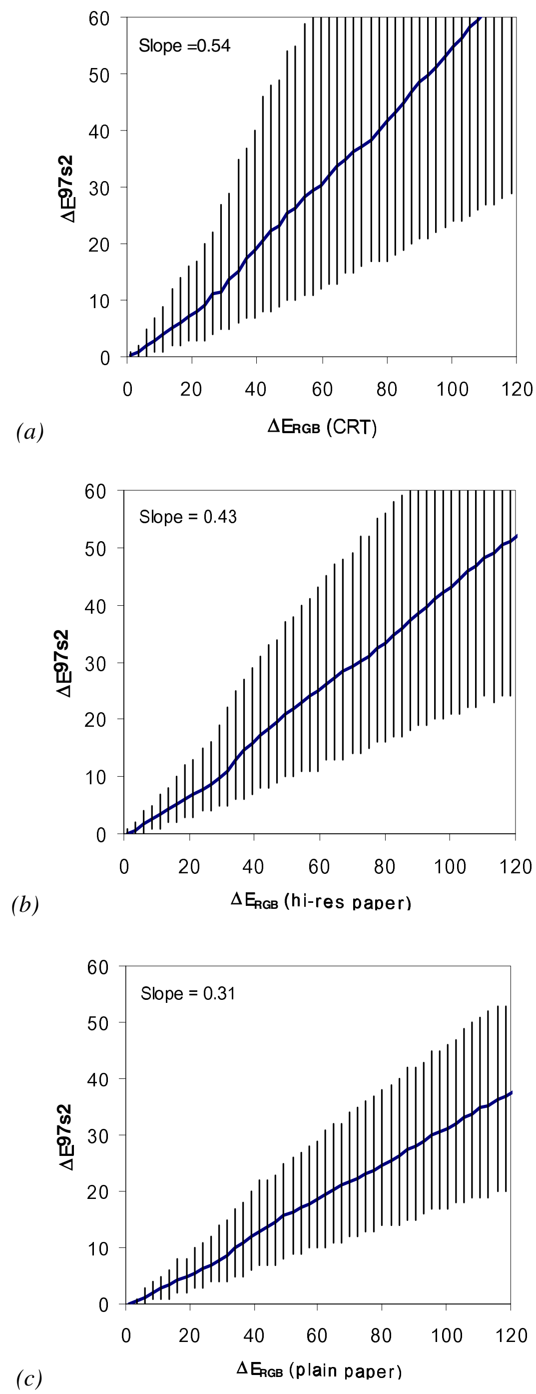

Figure 7. Relationship between (a) CRT ΔE_{RGB} and ΔE_{97s2}, (b) Hi-res ΔE_{RGB} and ΔE_{97s2} and (c) Plain ΔE_{RGB} and ΔE_{97s2}.

Table 5. Predictions of ΔE_{97s2} statistics for given ΔE_{RGB}s (Note, the coefficients are valid only when ΔE_{RGB} is below 120).

Function	$\Delta E_{97s2} = c_1 (\Delta E_{RGB}) + c_2 (\Delta E_{RGB})^2$								
Media	CRT			Hi-res paper			Plain paper		
Coefs.	c_1	c_2	R^2	c_1	c_2	R^2	c_1	c_2	R^2
Mean	0.50	0.000	0.999	0.40	0.0002	0.999	0.30	0.0001	0.999
Stdev	0.28	-0.001	0.997	0.15	-0.0002	0.999	0.10	-0.0002	0.999
2.5th perc.	0.18	0.001	0.998	0.18	0.0002	0.997	0.15	0.0002	0.997
97.5th perc.	1.24	-0.003	0.995	0.76	-0.0008	0.999	0.52	-0.0006	0.999
sample size	376,714 pairs of random RGB colours for each test								

Conclusions

The method given in the present paper can be used for predicting statistics of colour differences computed using a ΔE metric given a value of a different ΔE metric. The result of this is the ability to determine what range can be expected for values of one metric given a value computed in terms of another. This in turn allows for inter–comparison of findings from different studies even if these reported their findings in terms of the results of various colour difference formulae.

This paper also illustrated the use of the inter–comparison method proposed in it and showed that the method works with very high levels of accuracy and that it can be applied to a range of colour difference metrics including even Euclidean distance in device–dependent colour spaces. The effect of applying advanced colour difference formulæ in CAM97s2 space was also illustrated and a close relationship with their original CIELAB–based versions was shown.

References

1. CIE (1986) CIE Publication 15.2, *Colorimetry*, 2nd Edition.
2. CIE (1995) *Industrial Colour–Difference Evaluation*, CIE 116–1995.
3. C. J. Li, M. R. Luo and R. W. G. Hunt (1999) The CAM97s2 Model, *IS&T/SID 7th Color Imaging Conference: Color Science, Systems and Applications*, 262–263.
4. M. R. Luo and R. W. G. Hunt (1998) The Structure of the CIE 1997 Colour Appearance Model (CIECAM97s), *Color Research and Application*, **23**:138–146.
5. M. R. Luo, G. Cui and B. Rigg (2001) The Development of the CIE 2000 Colour-Difference Formula: CIEDE2000, *Color Research and Application*, **26**(5): 340-350.

Biographies

Mr. Pei–Li Sun received his bachelor and master degrees from the Department of Graphic Communications of the Chinese Culture University in Taipei, Taiwan, 1994 and 1996 respectively. Currently he is working at the Colour & Imaging Institute of Derby University towards his PhD entitled “The Influence of Image Characteristics on Colour Gamut Mapping.” He is about to commence work as a lecturer at the Department of Graphic Communications and Technology of the Shih Hsin University, Taiwan.

Dr. Ján Morovic received a PhD in Colour Science at the Colour & Imaging Institute (CII) where the topic of his research was “To Develop a Universal Gamut Mapping Algorithm.” He now works at the CII as lecturer in Digital Colour Reproduction and is module leader for three modules on the MSc in Imaging Science. Further he also serves as chairman of the CIE's technical committee 8–03 on Gamut Mapping.