Article:

Ren, S, Chen, Y, Westland, S orcid.org/0000-0003-3480-4755 et al. (1 more author) (2021) A comparative evaluation of similarity measurement algorithms within a colour palette. Color Research & Application, 46 (2). pp. 332-340. ISSN 0361-2317

https://doi.org/10.1002/col.22591

A comparative evaluation of similarity measurement algorithms within a colour palette

S Ren[1], Y Chen[2], S Westland[2] and L Yu[2]
1 University of Oxford, UK
2 University of Leeds, UK

Abstract

Recently there has been interest in the development of colour palettes from images. Colour palettes have long been used by designers to communicate colours and their relationships but increasingly palettes are being derived automatically from digital images, concepts, or from a plethora of digital design tools online. Methods to predict differences between palettes are growing in popularity. This study is concerned with the prediction of visual self-similarity for colour palettes with large numbers of patches. A psychophysical experiment was carried out to collect the human judgements of similarity and then six different algorithms were introduced and evaluated in terms of their ability to predict the psychophysical data. Two methods to quantify the agreement between the visual data and algorithm predictions were used based on regression analysis with coefficient of determination for the goodness of fit and multidimensional scaling with loss function Kruskal's stress. Of the six algorithms, the Pearson correlation coefficient method was considered to give the best performance.

1 Introduction

Colour palettes are collections of colour patches; they are ubiquitous in the design field and used as descriptions of images or collections of images (e.g. mood boards) and to represent colour themes. Currently, there are many digital tools for generating colour palettes[1-5]. For example, Adobe Colour CC (previously Adobe Kuler) is an internet application that allows a user to create 5-colour schemes based on various concepts of colour harmony according to the user's choice preferences or by uploading an image[1]. Colour palettes are a simple and visual way to capture colour relationships and are useful for representing designers' colour ideas based on their colour preferences, experience and knowledge of aesthetics[6]. Colour palettes can communicate specific topics or harmonies and can be used by designers, often based on their own colour preferences, to respond to design briefs[7].

Although colour palettes are often used by designers in a design context there is also work that has attempted to automatically generate colour palettes from images[24] or from words[25] or concepts[26]. This has led to related work to predict the visual difference between two colour palettes or, as in the case of this paper, the self-similarity of a colour palette (or the degree to which the patches in a colour palette are similar to each other). In a design project the colours in a palette are often very different from each other (for example, often consisting of complementary colours). However, in the case where colour palettes are derived from a word or concept, for example, then it is often desirable that the colours in the palette are more similar to each other and therefore methods to predict self-similarity can be useful.

The prediction of colour or visual difference between two colour patches is, of course, a topic of long interest[8]. In colour science, prediction of colour difference between single patches can be accurately measured by the distance between two points in a suitable colour space that represent those colours[9]. The CIELAB colour space is currently widely used and various metrics for computing differences in this space include CIE94, CMC, and most recently CIEDE2000[10-11]. However, the difference between two patches can be thought of as a special case (*n*=1) of difference between two *n*-colour palettes. There has recently been interest in predicting colour differences between *n*-colour palettes where $n > 1$. In one study, for example, psychophysical data were collected on the visual similarity between pairs of 25-colour palettes and three difference metrics (the single colour-difference model, the mean colour difference model and the minimum colour-difference model) were evalaued[12]. It was found that the minimum colour-difference model was most effective at predicting the visual data. The minimum colour-difference model (MICDM) works by averaging the colour differences between each patch in a palette and the closest patch in the second palette and the colour differences between each patch in the second palette and the closest patch in the first palette. Performance was evaluated using the coefficient of determination ($r^2 = 0.60$) and multidimensional scaling with loss function Kruskal's stress[13] (STRESS = 20.95). This study was extended with a second study in which 95 pairs of 5-colour palettes were evaluated[14]. The MICDM was again found to be effective in predicting visual differences. Using the CIELAB colour difference equation performance was slightly better than in the earlier study with $r^2 = 0.82$ and STRESS = 19.33. However, other colour difference formulae were also tested and best performance ($r^2 = 0.86$ and STRESS = 16.93) was obtained using the CIEDE2000(2,1,1).

In *n*-colour palettes, as $n \rightarrow 1$, the MICMD simply becomes a standard colour difference calculation and it is therefore not particularly surprising that CIEDE2000 was found to outperform CIELAB especially when *n* was small (*n*=5)[14]. As *n* becomes very large, the difference between two palettes becomes the equivalent of the visual difference between two images and it is not clear that MICDM would be effective or that CIEDE2000 would be the best colour difference equation. This study is concerned with *n*-colour palettes where *n* is relatively large (*n*=90). It is also concerned with the self-similarity of colour palettes (that is, comparing colour palettes with themselves) rather than estimating the visual difference between two palettes. In this work a high degree of self-similarity would be when all the colours in a palette are exactly the same as each other; a low degree of self-similarity would be when the colours in a palette are as different from each other as they possibly can be. The degree to which a colour palette is self-similar or coherent is an issue that is of interest in some recent research[15]. A psychophysical experiment was carried out to collect human visual judgements of the self-similarity of a total of 30 90-colour palettes and the results are presented. Various metrics to predict the visual data are described and evaluated.

2 Experimental

A psychophysical experiment was conducted to collect subjective judgements of self-similarity for each of the colour palettes. Visual judgements for palettes presented on a computer display were scaled using a slider bar based on a Likert Scale[16].

2.1 Colour Palettes Preparation

A set of 30 colour palettes were obtained from a previous study15 in which participants were presented with an adjective and each asked to select three colours that they associate with that word.

Figure 1: A representation of 30 of the 90-colour palettes used in this study

Since 30 participants took part in that study, a 90-colour palette was generated for each word. In that previous study the palettes were used to study word-colour associations; however, in this study the palettes are just used as a convenient set of palettes that have varying degrees of self similarity. In this study, for each 90-colour palette, the colour patches were displayed in a 30×3 array (see Figure 1) and presented to participants on a digital display.

2.2 Participants

In many previous psychophysical studies of colour difference the sample size was around 20-30 participants[12, 14]. Therefore, 30 participants were recruited in this study (includes 14 female and 16 male); all were 18-years old or more and with normal colour vision. The research purpose was briefly explained to each participant when they were recruited.

2.3 Experiment Design

The purpose of this experiment was to collect data on the self-similarity of colour palettes. The experiment took place in a darkened room with controlled viewing conditions, lighting conditions and display technology. Each colour palette was presented on the screen, one at a time and in random order, and below each palette the participants were asked to select the level of self-similarity for the palette using a slider bar. Before the experiment the following instructions were given: "For each palette use the slider bar to indicate the extent to which the colours in the palette are different or similar to each other." The participants were unaware of the adjectives that were used in the previous study[14] to generate the palettes.

The display part of the design required the development of a GUI and this was implemented using the MATLAB programming environment. Colour palettes were displayed on an LED computer monitor (HP DreamColor LP2480zx—a 24-in. LCD

Backlit Monitor) and viewed on a uniform grey (L*=50) background. The white point of the display was CIE x = 0.3116 and y = 0.3184 as measured by a Minolta 2000 Spectroradiometer (Konica Minolta Inc., Chiyoda, Japan). Each of the colour palettes was presented separately (and in a different random order for each participant) on the computer monitor. Note that although the order in which the palettes was presented was random, the positions of the patches in any palette was fixed. A scale bar (value from 0 to 1) beneath each colour palette allowed participants to select a specific level of self-similarity. The viewing distance was about 100cm from observers' eyes to the testing screen samples. The width and height of each palette was about 14 deg and 1.4 deg visual angle respectively; we note that each patch was quite small (<0.5 deg visual angle). Some simultaneous contrast or chromatic induction was likely and was unavoidable; however, in the majority of cases each patch was surrounded by multiple different colours so that the net effect would be small in many cases.

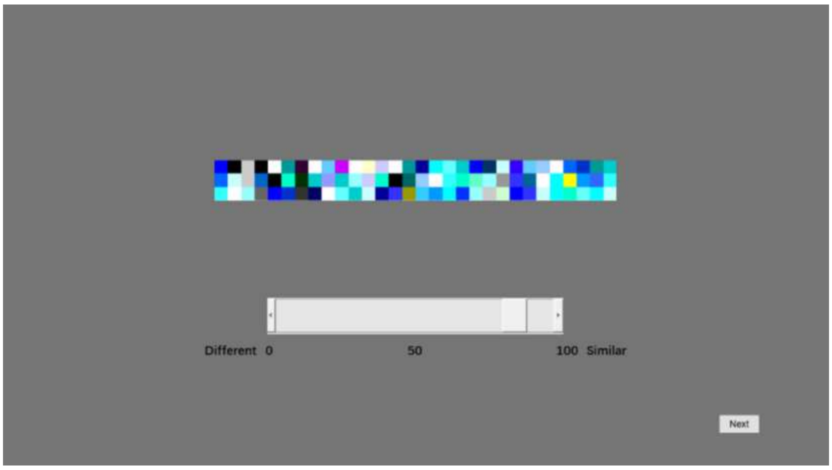

Figure 2: Screenshot of the experimental design

All observers passed a Colour Blindness Test before the experiment started. Each experimental session lasted approximately 40 minutes. The two ends of the scale were labelled similar and different. We assume that as the colour in the palettes become identical the participants will select the end of the scale labelled 'similar'; however, there will undoubtedly have been some variance in participant's perceptual understanding of the 'different' end of the scale. This is common with many magnitude estimation tasks. To ensure more consistency between the participants' use of the scale, before the experiment the participants were briefly shown all of the palettes on screen together from the experiment so that they had an appreciation of the range of palettes that would be presented. All participants also evaluated three practice palettes before starting the main experiment (these were not included in the results). Part of the reason for including the practice palettes was to allow the participants to adapt to the dark conditions.

2.4 Experimental Results

The colours were displayed in the experiment according to their RGB values. However, CIELAB values are required for subsequent colour-difference calculations[8-9]. The 2700 (30 palettes × 90 colours) RGB values were displayed one-at-a-time in the centre of the monitor and the spectral radiance was measured using a Konica Minolta CS-2000 spectroradiometer.

Table 1 The mean of visual colour difference (ΔV) of each colour palette

Colour Palettes	ΔV	Colour Palettes	ΔV	Colour Palettes	ΔV
1	0.3785	**11**	0.3560	**21**	0.4312
2	0.4102	**12**	0.5812	**22**	0.3721
3	0.6013	**13**	0.6011	**23**	0.3409
4	0.5281	**14**	0.5019	**24**	0.5173
5	0.3809	**15**	0.3742	**25**	0.4985
6	0.4771	**16**	0.3913	**26**	0.4517
7	0.5006	**17**	0.3576	**27**	0.4321
8	0.5332	**18**	0.3982	**28**	0.3981
9	0.6184	**19**	0.5119	**29**	0.3639
10	0.3827	**20**	0.4723	**30**	0.5010

The spectral radiance was converted to CIELAB values using the CIE Standard Observer (1964) and the display white point (CIE x = 0.3116, y = 0.3184). This approach was preferred to a more traditional colour-management process. In this work is more important to know precisely which colours were displayed than to display precisely defined colours. The values from the slider bar were averaged across all 30 participants to calculate the mean degree of self-similarity (ΔV) for each colour palette as shown in Table 1. Higher values of ΔV indicate greater judgements of self similarity.

3 Self-Similarity Measurement Algorithms

Six different self-similarity measurement algorithms were used to predict the visual data and these algorithms are described in this section.

3.1 Simple Colour-Difference Approaches

A number of methods to measure colour difference have been published[9-11]. In this study, two colour-difference formula are used, the CIELAB equation ΔE^*_{ab} and the CIEDE2000 equation ΔE^*_{00} . One approach to predict self-similarity of a palette would be to calculate the colour difference between each patch and each other patch in the palette and to average these values. However, previously published related studies found that for differences between palettes it was more effective to use the MICDM method and therefore a modified form of this method is used here. The application of the MICDM method to a pair n-colour palettes requires finding the closest match in the second palette for each colour in the first palette (and vice versa) and averaging these $2n$ colour differences. The method needs to be modified for the task of self-similarity for a single palette since if we find the closest colour in a palette for each colour in the palette then this will result in colours being compared with themselves and all the colour differences will be zero. The algorithm was therefore modified so that each of the n colours was compared with the other n-1 colours in the palette and the smallest colour differences recorded; these n colour differences are then averaged to compute MICDM. MICDM was calculated using the CIELAB and CIEDE2000(2,2,1) equations and these are denoted as $MICDM_{ab}$ and $MICDM_{00}$ respectively.

3.2 Feature Scaling

In CIELAB colour space, the range and unit of 'L', 'a' and 'b' value is different[17]. From the Euclidean colour distance, CIE94, CMC, and most recently CIEDE2000 were developed to address perceptual non-uniformities, while retaining the CIELAB colour space, by the introduction of application-specific weights derived from external factor[10-11]. Feature scaling was carried out in this study, to normalize the data itself and standardize the input variables ('L', 'a' and 'b' value) prior to the computation of the CIELAB colour distance measure. Feature scaling is a method used to normalize the range of independent variables or features of data. It is also known as data normalization in data processing[18-20]. Three main methods are used in this study: Standardization, Mean normalization and rescaling and Scaling to unit length. These methods are now briefly described. In each of these methods, each patch in the palette is compared to each other patch and the differences are averaged.

3.2.1 Standardized Euclidean distance

Based on the CIELAB (ΔE^*_{ab}) colour difference formulae, a Standardized Euclidean distance was used to modify the colour difference[20]. Each of the variables L*, a* and b* are standardized according to Equation 1:

$$x' = \frac{x - \bar{x}}{\sigma} \tag{1}$$

where x is the raw data, $\bar{x}$ is the mean of x, σ is its standard deviation. Following this standardization, Equation 2 is used to calculate a colour difference.

$$\Delta E_s = \sqrt{\left(\frac{\Delta L}{\sigma_L}\right)^2 + \left(\frac{\Delta a}{\sigma_a}\right)^2 + \left(\frac{\Delta b}{\sigma_b}\right)^2} \tag{2}$$

3.2.2 Mean normalization and Rescaling

Mean normalization and Rescaling are the simplest method of Feature scaling[21]. When x is an original value, x' is the normalized value, the general formula of Mean normalization is given as, Equation 3:

$$x' = \frac{x - \bar{x}}{\max(x) - \min(x)} \tag{3}$$

Rescaling also known as min-max scaling or min-max normalization and consists of rescaling the range of features to scale the range in [0, 1], Equation 4:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{4}$$

It is noticed that, when substituted into CIELAB (ΔE^*_{ab}), these two methods are equal:

$$\Delta E_R \& \Delta E_M = \sqrt{\left(\frac{\Delta L}{\max(L) - \min(L)}\right)^2 + \left(\frac{\Delta a}{\max(a) - \min(a)}\right)^2 + \left(\frac{\Delta b}{\max(b) - \min(b)}\right)^2} \tag{5}$$

3.2.3 Scaling to unit length

Scaling to unit length is to scale the components of a feature vector such that the complete vector has length one[16]. Which means dividing each value by the Euclidean length of the vector, Equation 6:

$$x' = \frac{x}{|x|} \tag{6}$$

$$\Delta E_U = \sqrt{\left(\frac{\Delta L}{|C_1| - |C_2|}\right)^2 + \left(\frac{\Delta a}{|C_1| - |C_2|}\right)^2 + \left(\frac{\Delta b}{|C_1| - |C_2|}\right)^2} \tag{7}$$

where $C_1 = [L_1, a_1, b_1]$ and $C_2 = [L_2, a_2, b_2]$.

3.3 Pearson Correlation Coefficient(PCC)

Pearson correlation coefficient method (PCC) is used to measure the direction and the strength of the linear relationship between every two variables which is the covariance of variables (x ,y) divided by the product of the standard deviations, Equation 8 (cov: covariance; σ: standard deviation)[22-23]:

$$P_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{E[(x - \sigma_x)(y - \sigma_y)]}{\sigma_x \sigma_y} \tag{8}$$

The value of PCC is $-1 \leq P \leq +1$ and corresponds to the degree of correlation.

In each colour palette, every two colours were calculated by these six algorithms respectively and the mean of them were collected to represented the colour difference of the colour palette. Table 2 shows the similarity measures for each of the colour palettes as obtained by each of the six algorithms.

4 Results and Discussion

For each colour palette, the scores obtained from each participant were averaged and these resulted in mean scores between 34.1 (for religious) and 61.8 (for fresh). Recall that the higher the score the more self-similar the participants evaluated the palette to be. To give a feeling for the variation between participants the range for the palette corresponding to religious was 16.4 - 48.9 and for the palette corresponding to fresh was 53.2 – 78.9. For each palette the standard error of the mean was calculated and these are shown in Figure 3 for 9 of the colour palettes. Overall the grand mean of visual scores was 45.5 and the average standard error of the mean was 1.36 (just under 3%). The standard error of the mean represents the inter-observer error and indicates that participants had quite good agreement in terms of their evaluations of the palettes.

Table 2 The similarity score for each colour palette using the different algorithms.

Colour Palettes	MICDMab	MICDM00	ΔE_S	ΔE_M ΔE_R	ΔE_U	PCC
1	5.1372	2.0971	2.1494	0.6428	0.9329	0.7126
2	4.3494	2.3596	2.1774	0.4936	0.8205	0.6834
3	4.4095	1.8449	1.9851	0.4210	0.5975	0.7793
4	4.9345	2.7650	2.1001	0.5480	0.6916	0.7003
5	5.4984	2.7138	2.2235	0.6084	0.8900	0.6749
6	4.5365	2.6102	2.1581	0.5815	0.7562	0.6471
7	3.4153	1.6153	2.0367	0.5290	0.6051	0.7171
8	5.7657	2.4562	2.1819	0.5249	0.7731	0.6384
9	6.1458	3.1045	2.0840	0.5027	0.7002	0.7604
10	6.6506	3.5022	2.1968	0.5922	0.8623	0.6466
11	5.0776	2.2798	2.1536	0.6276	0.8805	0.6906
12	5.7278	2.4703	2.1156	0.5637	0.7577	0.7465
13	3.6963	1.6164	2.0004	0.4717	0.5192	0.7873
14	3.9474	1.4309	2.1733	0.6435	0.8757	0.7275
15	4.3497	2.1539	2.2082	0.6339	0.9226	0.6312
16	6.2006	2.9485	2.1859	0.5609	0.8163	0.6418
17	5.5183	2.7852	2.1777	0.5992	0.8407	0.6824
18	6.4777	3.0031	2.2027	0.5474	0.8213	0.6800
19	4.8105	2.5998	2.1298	0.5743	0.7579	0.7362
20	4.7477	2.5929	2.1933	0.4905	0.7229	0.6723
21	3.7116	1.9879	2.1840	0.5015	0.7226	0.6620
22	5.4029	2.8169	2.1790	0.5802	0.8710	0.6393
23	6.6293	3.1418	2.2207	0.6143	0.8401	0.6727
24	7.1861	3.3950	2.1663	0.5991	0.8645	0.7454
25	7.7970	3.6127	2.1531	0.5766	0.8314	0.7286
26	5.7734	2.7808	2.1900	0.5473	0.8183	0.6426
27	4.6751	2.8262	2.2308	0.6107	0.8809	0.6443
28	5.0394	2.7339	2.2221	0.5500	0.8027	0.6372
29	7.7741	3.8342	2.2177	0.5937	0.7986	0.6805
30	4.8911	2.5656	2.1468	0.5632	0.7784	0.7014

ΔE_S Standardized Euclidean distance

ΔE_M Mean normalization Euclidean distance

ΔE_R Rescaling Euclidean distance

ΔE_U Scaling to unit length Euclidean distance

PCC Pearson Correlation Coefficient

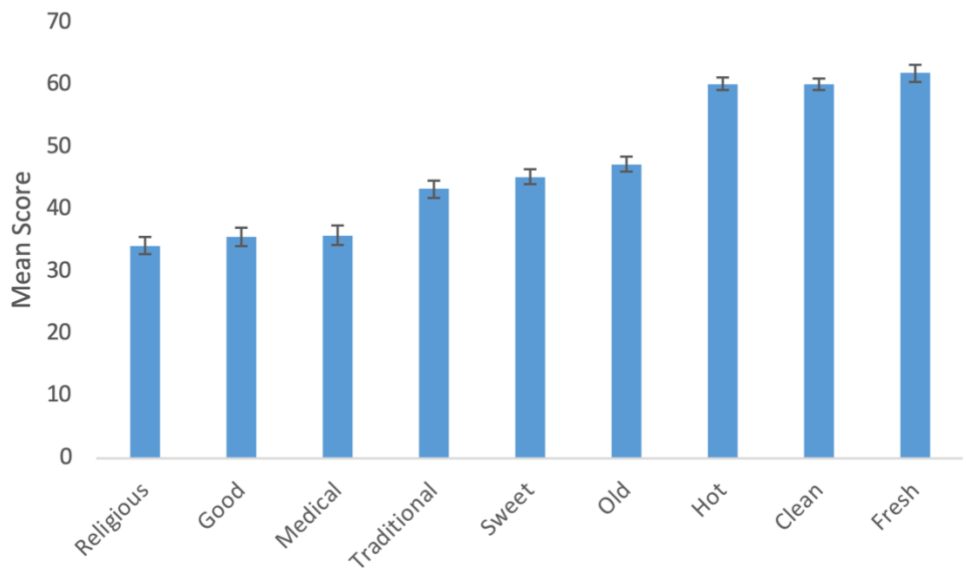

Figure 3: Mean visual scores for the three most similar palettes (fresh, clean and hot), the three least similar palettes (religious, good and medical) and three mid-range palettes (traditional, sweet and old) showing the standard error of the mean in each case.

Six different algorithms were carried out to predict self-similarity. Two methods were used to quantify performance between the visual data and the output from each of the algorithms. Firstly, the coefficient of determination (r^2) was calculated. However, there is a long-standing debate about the effectiveness of the coefficient of determination to quantify the agreement between visual and instrumental data in the field of colour difference evaluation. There is growing consensus that a method known as STRESS is more effective[6]. Therefore, the STRESS value was also calculated between the visual data and the output from each of the algorithms. Note that better agreement is indicated by higher r^2 and lower STRESS values. One advantage of the use of coefficient of determination is its association with the scatter plots that can be easily viewed and these are shown in Figure 4 for each of the six algorithms. The plot for the Standardized Euclidean difference, however, does reveal the weakness of using r^2 since data a set of x,y data where y is perfectly invariant with x can show a very high r^2 value (note how the regression line is almost horizontal in this case).

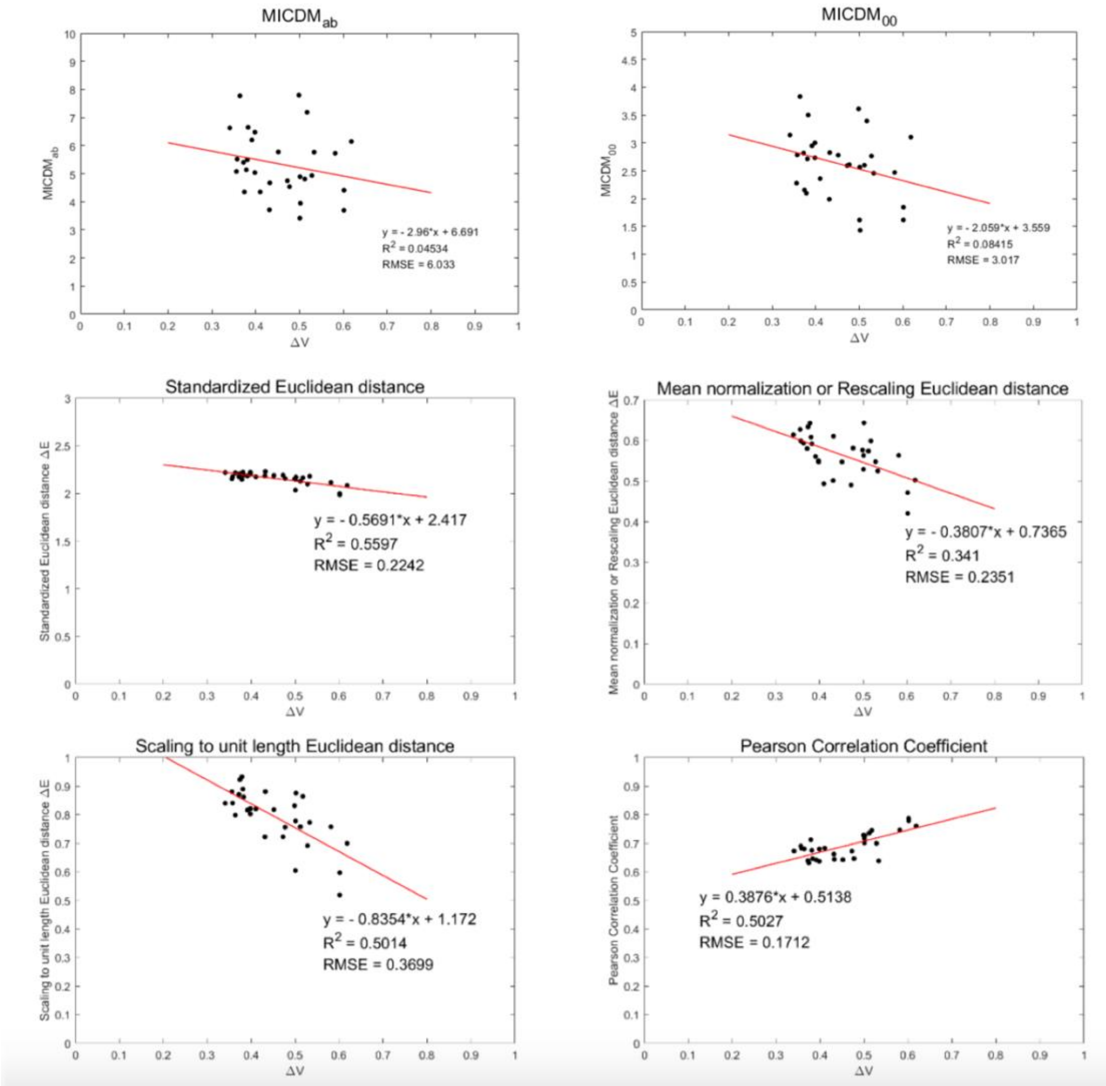

Figure 4: The coefficient of determination (r^2) of six algorithms compared with ΔV

Table 3 shows the quantitative results using r^2 and STRESS for each of the six algorithms. There is evidently some agreement between the r^2 and STRESS data. For example, note that both measures show poor agreement with the visual data (low r^2 and high STRESS) for $MICDM_{ab}$, $MICDM_{00}$ and ΔE_M ΔE_R. However, according to r^2 the ΔE_S method (r^2= 0.56) gives the best performance whereas according to STRESS the PCC method (13.78) performs best. Overall, the PCC method is considered to have the best agreement with the visual data. This is based on the observation that the regression line for ΔE_M ΔE_R is tending towards a horizontal line and the general consensus in the literature that prefers STRESS to r^2 for colour-difference evaluation. Note also that the PCC method performs best using STRESS but also gives quite a high value of r^2.

Table 3 Calculation of r^2 and STRESS values for six algorithms with ΔV

	$MICDM_{ab}$	$MICDM_{00}$	ΔE_S	ΔE_M ΔE_R	ΔE_U	PCC
r^2	0.0453	0.0842	0.5597	0.3410	0.5014	0.5027
STRESS	29.4652	31.0058	19.7154	24.0523	27.1188	13.7798

* ΔE_S *performs best compared with others in r^2.*
PCC performs best compared with others in STRESS.

5 Conclusions

In summary, in this study a psychophysical experiment was employed to obtain judgements of visual self-similarity (ΔV) for each of a set of 90-colour palettes. Six different algorithms were developed and evaluated to predict the psychophysical data. These were two methods based on MICDM ($MICDM_{ab}$ and $MICDM_{00}$) and the other methods were Standardized Euclidean distance (ΔE_S), Mean normalization Euclidean distance & Rescaling Euclidean distance (same performance) (ΔE_m and ΔE_R), Scaling to unit length Euclidean distance (ΔE_U) and Pearson Correlation Coefficient (PCC). The performance of these six algorithms has been compared two analysis methods: regression analysis (r^2) and statistical analysis (STRESS). Overall, the PCC method was considered to give the best agreement with the visual data; it gave the lowest STRESS value and one of the lowest r^2 values.

Previous research had shown that the MICDM method was effective for estimating the visual difference between pairs of palettes[8,9]. It is perhaps surprising that $MICDM_{ab}$ and $MICDM_{00}$ gave only weak agreement with the visual data in this study even though the number of patches (90) in each palette was much larger than in the previous studies. However, the task in those previous studies concerned differences between palettes whereas this work was concerned with self-similarity of a single palette. Greater self-similarity might be considered to be a measure of coherence of the palette. In one study, for example, colour palettes were automatically generated using a machine learning algorithm using a word as the input and where the colour palettes were assumed to represent the word[11]. In that study it was suggested that the degree of coherence of a palette was a likely indicator of the strength of the association between the colour palette and the word. This study suggests that the PCC might be effective method for quantifying the degree of coherence of a colour palette.

For an *n*-colour palette as *n* becomes very large there is a point when the palette is referred to as an image rather than as a palette. This suggests that the PCC methods might have interesting applications to other image-processing aims in digital imaging generally. Finally, it is recognised that this study is of course limited. Further studies are needed to corroborate the findings in this study by using colour palettes with far fewer colour patches, for example, and by using a wider source for the colour palettes themselves.

References

1. Adobe Color: https://color.adobe.com/create. Accessed January 7, 2020.

2. PaletteGenerator. https://palettegenerator.com/. Accessed January 7, 2020.

3. Colourlovers. https://www.colourlovers.com/palettes. Accessed January 7, 2020.

4. Colormind. http://colormind.io/. Accessed January 7, 2020.

5. Colour Tool. https://material.io/resources/color/. Accessed January 7, 2020.

6. Bartram L, Patra A and Stone M (2017). Affective color in visualization. Paper presented at: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 1364–1374. ACM.

7. Moretti G and Lyons P (2002). Tools for the selection of colour palettes. Paper presented at: Proceedings of the SIGCHI-NZ Symposium on Computer-Human Interaction, 13–18, ACM.

8. McDonald R (1990). European practices and philosophy in industrial colour-difference evaluation. Color Res Appl, 15(5), 249-260.

9. Witt K (1995). CIE guidelines for coordinated future work on industrial colour-difference evaluation. Color Res Appl., 20(6), 399-403.

10. Melgosa M (2000). Testing CIELAB-based color-difference formulas. Color Res Appl., 25(1), 49-55.

11. Luo MR, Cui G and Rigg B (2001). The development of the CIE 2000 colour-difference formula: CIEDE2000. Color Res Appl., 26 (5), 340-350.

12. Pan Q and Westland S (2018). Comparative evaluation of color differences between color palettes. Proceedings of the IS&T Color and Imaging Conference, 110–115.

13. García PA, Huertas R, Melgosa M and Cui G (2007). Measurement of the relationship between perceived and computed color differences. J. Opt. Soc. Am. A, 24, 1823–1829.

14. Yang J, Chen Y, Westland S. and Xiao K (2020). Predicting visual similarity between colour palettes. Color Res Appl., 45(3), 401-408.

15. Chen Y, Yang J, Pan Q and Vazirian M (2020). A method for exploring word-

colour associations. Color Res Appl., 45(1), 85-94.

16. Thornhill A, Saunders M and Lewis P (2009). Research methods for business students. Prentice Hall: London; 127-128.

17. Hunt RWG and Pointer MR (2011). Measuring colour. John Wiley & Sons.

18. Ioffe S and Szegedy C (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167.

19. Joel G (2015). Data Science from Scratch. Sebastopol, CA: O'Reilly, 99-100.

20. Juszczak P, Tax D and Duin RP (2002). Feature scaling in support vector data description. Proc. ASCI. 95-102.

21. Milligan GW and Cooper MC (1988). A study of standardization of variables in cluster analysis. Journal of Classification, 5(2), 181-204.

22. Solutions, Statistics. Correlation. Pearson, Kendall, Spearman. Accessed October 18: 2016.

23. Hauke J and Kossowski T (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. Quaestiones geographicae, 30 (2), 87-93.

24. Morse BS, Thornton D, Xia Q and Uibel J (20007). Image-based color schemes. Paper presented at: IEEE International Conference on Image Processing, IEEE, 3; III-497.

25. Lavendel L and Kohler T (2008). The story of a color advisor, Sixth Color Imaging Conference: Color Science, Systems and Applications, volume 6, pp228–229, Scottsdale (Arizona)."

26. Lindner A and Süsstrunk S (2013). Automatic color palette creation from words. Paper presented at: Color and Imaging Conference, Society for Imaging Science and Technology; 69–74.

AUTHOR BIOGRAPHIES

Yun Chen is a PhD student of colour research at the University of Leeds (UK) and currently a visiting researcher in the University of Oxford (UK). Her research focuses on colour associations and exploring word-colour associations in design. Her research interests also include colour design and data-driven colour selection.

Shuai Ren is a PhD student of tissue engineering at the University of Oxford (UK). He received Bachelor degrees in Mechanical Design from Huazhong University of Science and Technology (China), and a Master degrees in Mechanical Design from the University of Manchester (UK). His research interests include image analysis, 3D printing, Bio-sensor and Bio-reactor.

Stephen Westland is Professor of Colour Science and Technology at the University of Leeds (UK). His research interests include colour measurement, colour imaging and colour design.

Luwen Yu received Bachelor and Master degrees in Industrial Design from the School of Mechanical Science and Engineering, Huazhong University of Science and Technology, China, in 2013 and 2017. She is currently a PhD student in the School of Design, University of Leeds (UK). Her research interests include colour packaging, colour marketing and design processes.