An Algorithm for Categorising Colours into Universal Colour Names

Zhaohui Wang and M. Ronnier Luo
Colour & Imaging Group, Department of Colour and Polymer Chemistry, University of Leeds, Leeds, UK

Byoung-Ho Kang, Heui-Keun Choh and Chang-Yeong Kim
Samsung Advanced Institute of Technology, Suwon, South Korea

Abstract

Colour categorisation has been studied extensively on many different areas such as linguistic, psychology, physics and cognition. In colour science, colour categorisation has been used to solve the practical problems in the field of colour image reproduction and image analysis. A method to automatically categorise colours into the 11 universal colour names found by Berlin and Kay was developed here. Colour category experiment was carried out using 729 colour patches printed on four types of paper substrates. Models for categorising universal colours were developed in CIELAB and CIECAM02 colour spaces. The results show that the CIECAM02 based model outperformed the CIELAB based model and gave the same degree of accuracy as a group of observers.

Introduction

In 1969, Berlin and Kay[4] found that 11 universal basic colour terms exists in 20 languages – these colours are black, white, red, green, yellow, blue, brown, pink, orange, purple and grey. From then on, there was a flurry of critiques, responses and further studies.[4,19] Recent survey[16] on semantic aspects of colour categorisation mentioned that "colour categorisation can be understood as the reduction of the infinite differences among objects in the perceived world to cognitively usable proportions by means of which non-identical stimuli can be treated as equivalent."

Although some linguists argued that the colour categories are not universal,[7] the 11 colour terms have been shown to be very useful in colour and imaging applications such as gamut mapping,[9] colour quantisation,[15] image compression,[15] segmentation[1] and retrieval systems.[18]

A cross cultural colour naming study carried out by Lin *et al.*[8] resulted in modelling the colour categorisation process. In their study, a model was derived to categorise the 11 universal colour names in CIELAB[6] colour space. The original experiment was based on the colour chips in the Natural Color System (NCS). Hence, any colour described in terms of CIELAB values, Lightness (L^*), Chroma (C^*_{ab}), and hue angle (h_{ab}), can be named unambiguously. Each colour region is not overlapped with each other and also there have no gap in the colour space.

In 2002, a new colour appearance model, CIECAM02,[5] was adopted by the CIE, which is capable of accurately predicting the colour appearance under a wide range of viewing conditions. It is an improved version of the earlier model, CIECAM97s. It was found that CIECAM02 outperformed CIECAM97s in terms of simplicity and accuracy. Recently, significant progress has been made on the extension of CIECAM02 shows a reasonable performance for evaluating colour differences.[12]

The objective of this work is to categorise colours based on CIELAB and CIECAM02 spaces. Boundaries for each universal colour are determined by minimising wrong decision (see later) from a set of experimental data. The performances are compared between the results in CIELAB and CIECAM02. Furthermore, the results from four different paper substrates were also compared.

Colour Category Experiment

Many studies on colour categorisation have been performed based on human perception. Different colour samples were categorised by panels of observers in perceptual, cognitive, linguistic and colour science experiments.

In this study, a psychophysical experiment was carried out to ask observers to name colour patches according to 11 universal colour names. The TC9.18 RGB chart[3] released with GretagMacbeth ProfileMaker V5.0.1 was used as shown in Fig.1. Only the colours on the left marked within black frame were used, including a total of 729 colour patches.

This target was printed in the size of $27 \times 21.5cm$ by an HP 20ps Designjet printer on four different substrates: A3 plain paper (plain), HP Design Heavyweight paper (design), HP Proofing Gloss paper (gloss) and HP Bright White Inkjet paper (bright). The size of each colour patch was 0.7×0.7 c*m*. The HP Designjet Postscript RIP software was used to produce the prints without any colour correction. The printer resolution was set to $600 \times 1200dpi$. The mid-term stability of ink was investigated during a period of 12 days. Colour differences, in CIELAB units, between the same print measured in Day 1 and Day 12 were analysed. The mean ΔE^*_{ab} value was found to be 1.64 units. The printed samples were kept in dark environment for 12 days before the real experiment.

Ten normal colour vision observers according to the Ishihara test took part in the experiment. A VeriVide viewing cabinet was used to view the target under a CIE D50 simulator (fluorescent lamp). The background was a mid-grey having L^* of 50. The target was observed at a distance of 50 cm with a fixed 0/45 viewing geometry. Observers reported one of the 11 universal colour names for each colour patch according to the order from top to bottom, left to right in the target. Finally, 2916 ($= 4 \times 729$) colour names were collected. For each patch, the results were summarised in terms of percentage of component colours, e.g. 0.8 red, representing 80% of observers say that the colour has a 'red' colour name.

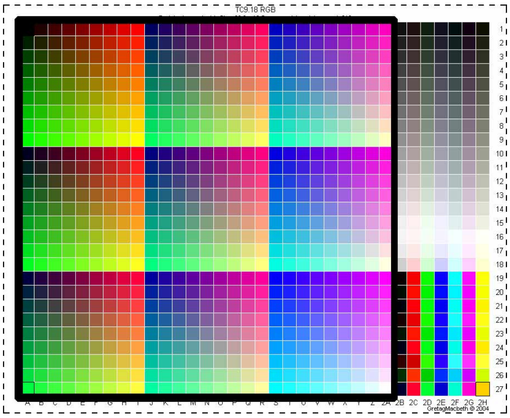


***Figure 1.** GretagMacbeth TC9.18 Chart.*

The target was measured by a 45/0 spectrophotometer, GretagMacbeth Spectrolino. The measured reflectance was converted to CIE tristimulus values under CIE1931 standard colorimetric observer according to ASTM E308 table.[2] The tristimulus values were then transformed to CIELAB specifications, L^*, C^*_{ab}, h_{ab} under the CIE D50 illuminant. For transforming the tristimulus values to CIECAM02 lightness *(J),* Chroma *(C), and* hue angle *(h*), the input parameters were 'average' surround under the D50 reference white with the luminance factor of the background (Y_b) of 20 and an adapting luminance (L_A) of 78 cd/m^2.

Based on the experiment data, an improved colour naming model was implemented to categorise all colours in CIELAB and CIECAM02 colour spaces into 11 basic colour names.

Modelling Colour Categorisation

Experimental data were first plotted in the CIELAB *a*b** and CIECAM02 *ab* diagrams as shown in Fig.2 (using CIECAM02 as example). Following the method developed by Lin *et al.,*[8] these data were divided into one achromatic (including white, grey and black) and eight chromatic colour categories according to the total number of colours named by over 50% observers. It can be clearly seen that these regions are well distributed according to hue angle. However, it appears that there are some overlaps between neighbouring colours. The most obvious example is the colour regions between red, brown and orange colours.

To develop the colour category model, three stages were involved in CIECAM02 *JCh* or CIELAB L^*C^*h space. CIECAM02 is used here to illustrate each individual stage.

Stage 1: To Determine Achromatic Colours

Achromatic colours, by definition,[14] are colours devoid of hue. The boundaries of achromatic colours, white, grey and black, are constrained by the lightness and chroma boundaries, which are defined firstly. The model assumes that these colours are hue independent. As shown in Fig.3, the experimental achromatic colours are plotted in lightness (J) vs. chroma (C). By drawing boundaries at J values of 90 and 25, the white, grey and black regions are determined. Note that it is unavoidable that some colours are included into wrong regions. By avoiding inclusion of these colours, the chroma boundaries for white and black are set at 10, and 5 for grey.

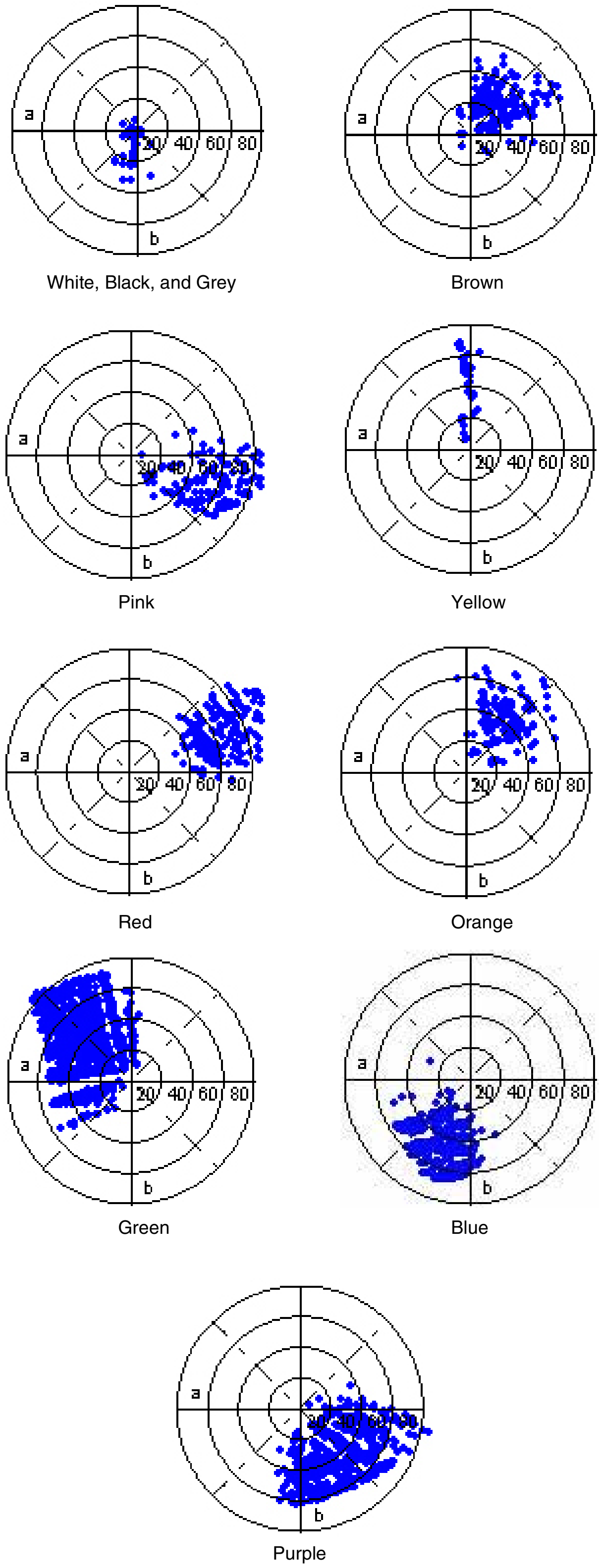


***Figure 2.** 2096 colour samples plotted according to CIECAM02 Cartesian Coordinates.*

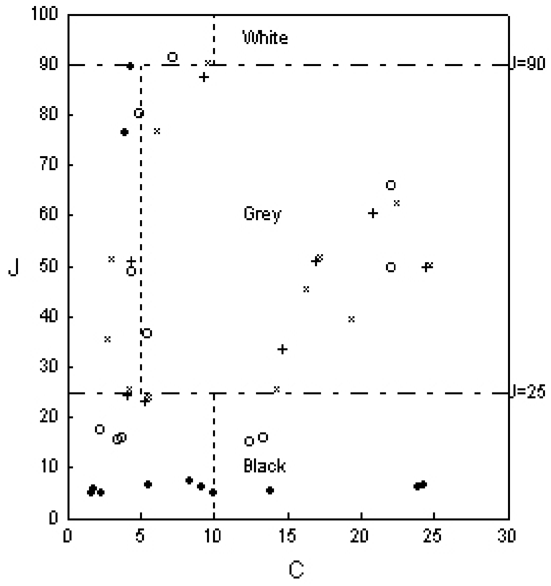


Figure 3. *Plotting of CIECAM02 lightness and chroma for White Black, and Grey in four substrates: (point) gloss, (o)design, (+)plain, (x)bright.*

Stage 2: To Determine Pink, Brown and Yellow

Boundaries for pink, brown and yellow are defined in terms of all three attributes, *J*, *C* and *h*. Yellow colours was found that only occupy a space with a high level of lightness within hue angles close to 90°. When lightness decreases, colours having the same chroma and hue angle change from yellow to green.[8] The chroma boundaries are ignored here for yellow colour.

Stage 3: To Determine Red, Orange, Green, Blue and Purple

Finally, the remaining colours, red, orange, green, blue and purple were only defined by hue angle following an anti-clockwise direction as shown in Fig.2. However, considering the colour regions of brown and orange, of pink and purple, of yellow and green in Fig.2, some data overlapped with each other. Fig.4 plots the data of brown and orange regions in CIECAM02 JC plane. It can be seen that most colours can be separated. The main overlapped data are circled in the figure. By minimising the wrong decision (see later), the hue angle for orange was determined, which results in separation of most overlapped data.

The identical procedure was used for the pink and purple regions, and yellow and green regions. After ruling out brown, pink, and yellow by using three parameters, J, C, and h, other five colours can be simply separated by hue angle only.

Tables 1 and 2 show the results for the boundaries of the 11 basic colours for CIELAB and CIECAM02 colour space separately. It can be seen that there are some differences, especially in the boundaries of blue and purple.

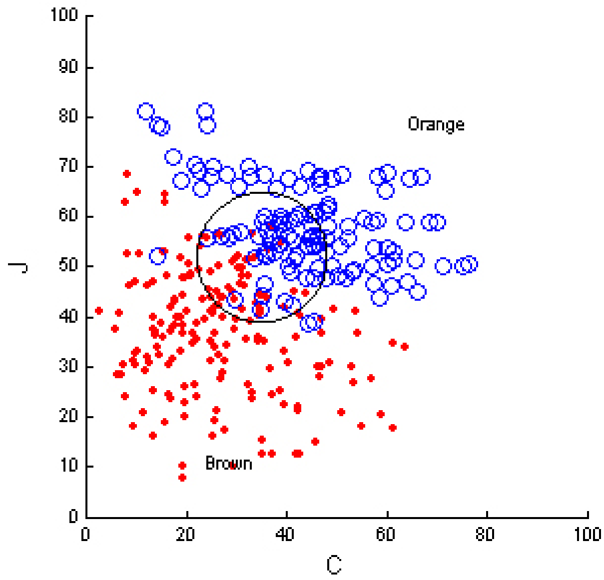


Figure 4. *Plotting of CIECAM02 Lightness and Chroma values for Brown and Orange.*

Table 1 Boundaries of colour categories in CIELAB colour spaces

Stages	Basic colours	L^*		C^*_{ab}		h_{ab}	
		<	≥	≤	>	≤	>
1	White	100	92	5	0	-	-
	Grey	92	35	5	0	-	-
	Black	35	0	5	0	-	-
2	Brown	60	10	45	5	20	80
	Pink	100	50	85	5	320	5
	Yellow	100	83	-	-	85	105
3	Red	-	-	-	-	5	45
	Orange	-	-	-	-	45	80
	Green	-	-	-	-	80	210
	Blue	-	-	-	-	210	285
	Purple	-	-	-	-	285	5

Table 2 Boundaries of colour categories in CIECAM02 colour spaces

Stages	Basic colours	J		C		h	
		<	≥	≤	>	<	≥
1	White	100	90	10	0	-	-
	Grey	90	25	5	0	-	-
	Black	25	0	10	0	-	-
2	Brown	65	15	45	5	15	80
	Pink	100	43	85	5	300	5
	Yellow	100	75	-	-	75	100
3	Red	-	-	-	-	5	40
	Orange	-	-	-	-	40	80
	Green	-	-	-	-	80	210
	Blue	-	-	-	-	210	269
	Purple	-	-	-	-	269	5

Discussions

A measure of "wrong decision" (WD) was used to optimise the boundaries for each colour region. If a colour is categorised in a wrong colour region against the visual results, it will be counted as a WD, which can be further divided into two parts: wrongly included (WI) and wrongly excluded (WE). WI indicates the number of colours wrongly included. WE indicates the number of colours wrongly excluded. The performances of the model in each colour space are given in Table 3. (Note that, the number of basic colours is the total number of basic colours named by over 50% of subjects which means that colour holds a highest rated category for that colour patch.)

The results showed that CIECAM02 based model (16 WD %) gave a more accurate prediction than CIELAB model (21 WD %). Most importantly, CIECAM02 also gave a similar degree of accuracy as the visual results based upon a panel of ten observers (16 WD %).

Table 3 Performances of model in CIELAB and CIECAM02

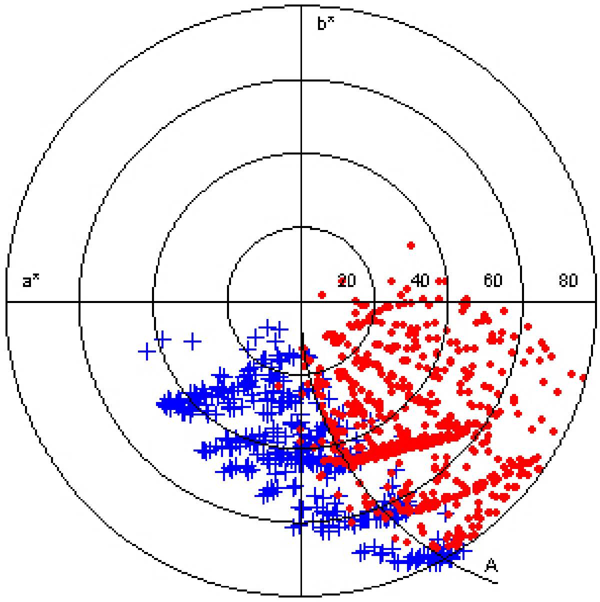
Basic colours	No. of basic colours	CIELAB WD	CIECAM02 WD
White	4	2	1
Grey	20	21	16
Black	22	8	13
Brown	194	59	39
Pink	156	60	53
Yellow	39	23	8
Red	196	56	55
Orange	128	46	49
Green	1136	104	78
Blue	375	107	72
Purple	646	123	84
Total	2916	609	468
%		21%	16%

Table 4 Substrate performance in WD% for each colour term in CIECAM02

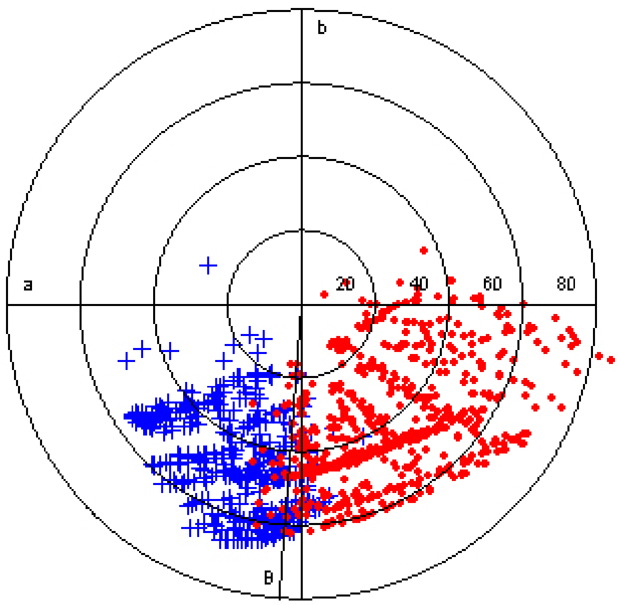
Basic colours	Plain	Bright	Design	Gloss
White	100	0	0	0
Grey	41	41	9	9
Black	0	13	40	47
Brown	27	30	23	20
Pink	29	28	22	21
Yellow	12	12	38	38
Red	26	28	26	19
Orange	27	25	22	26
Green	27	24	30	19
Blue	27	28	23	22
Purple	14	27	29	30
Mean	*30*	*23*	*24*	*23*
Mean (Except Neutral)	*24*	*25*	*27*	*24*

Although four substrates may lead to difference in appearance, such as gloss and colour gamut, it was also found that the average performance of each substrate is very similar as shown in Table 4 except the three neutral colours. (Note that, the percentages presented in Table 4 are the contributions of each substrate for total wrong decision numbers for the corresponding colour region.) In other words, the model is substrate independent, which is desired for colour management applications. There is an evidence to show that neutral colours are somewhat substrate dependent. There are also insufficient samples in these regions.

Fig. 5 a and b plot samples in the blue (pluses) and purple (points) regions in CIELAB and CIECAM02 Cartesian coordinates respectively. Their hue boundaries are also drawn in each space, line A in Fig.5-a, and line B in Fig.5-b. It can be seen that CIECAM02 performed better than CIELAB for those colours in the boundary. A curved boundary has to be applied to CIELAB. This implies that colours in a constant hue angle in the blue region of CIELAB space do not perceive as constant hue. This problem for CIELAB has addressed by other researchers.[13]



(a) CIELAB



(b) CIECAM02

***Figure5**. Plotting the comparison of blue (pluses), and purple (points) colours in CIELAB and CIECAM02 colour space.*

Conclusions

Colour category models based on CIECAM02 and CIELAB colour space were developed according to the 11 basic colour terms. Their performances were compared. CIECAM02 outperformed CIELAB, because the latter is lack of blue hue constancy. The CIECAM02 gives greater advantage than the other space like CIELAB. The current category model based on CIECAM02 is reliable. Further works need to be done by inclusion of more samples in the colour term boundaries to develop a more reliable and robust model.

The present experimental results showed that the colour category model based on CIECAM02 give very similar results for the four very different substrates studied except for neutral region.

There are some colours which have been memorised as preferred colour in human mind, such as skin, sky, vegetation and others.[11] These are memory colours. For image reproduction, the rendering of memory colours is also important.[17] The same experimental technique can be conducted to the study of important memory colour on different substrate, different media (such as display, different kinds of printer). The colour categorical model can be extended to involve these colours and capable of predicting colour appearance under different viewing conditions.

References

[1] A. Mojsilović, A Computational Model for Color Naming and Describing Color Composition of Images, IEEE Trans. Image Process., 14, pg. 690-699 (2005).

[2] ASTM E308-95, Standard Practice for Computing the Colors of Objects by Using the CIE System, American Society for Testing and Materials, Philadelphia, (1995).

[3] B. Atkinson, TC9.18 RGB Test Chart, released with GretagMacbeth ProfileMaker V5.0.1., (2003).

[4] B. Berlin and P. Kay, Basic Color Terms: Their Universality and Evolution, Berkeley: University of Califonia Press, (1969).

[5] CIE, A Colour Appearance Model for Colour Management Systems: CIECAM02, CIE Central Bureau, Vuenna, Austria, CIE Pub.159 (2004).

[6] CIE, Colorimetry, 3rd edition, CIE Central Bureau, Vienna, Austria, CIE Pub. 15:2004 (2004).

[7] D. Roberson, J. Davidoff and I. Davies, Color Categories are not universal: New evidence from Traditional and Western cultures, Proc. of SPIE, 9th Congress of International Colour Association, pg. 37-40 (2002).

[8] H. Lin, M.R. Luo, L.W. MacDonald and A W S. Tarrant, A Cross-Cultural Colour-Naming Study: Part III – A Colour – Naming Model, Color Res. Appl., 26, pg. 270-277 (2001).

[9] H. Motomura, O. Yamada and T. Fumoto, Categorical Color Mapping for Gamut Mapping, Proc. 5th CIC, pg. 50-55 (2004).

[10] H. Yaguchi, Color Categories in Various Color Spaces, Proc. 9th CIC, pg. 6-8 (2001).

[11] J. Pérez-Carpinell, M.D. de Fez, R. Baldoví and J.C. Soriano, Familiar Objects and Memory Color, Color Res. Appl. 23, 6, pg. 416-427 (1998).

[12] M. R. Luo, G. Cui and C. Li, Uniform Colour Spaces Based on CIECAM02 Colour Appearance Model, Dept. Colour and Polymer Chemistry, Univ. Leeds, UK (2006).

[13] P. Hung and R.S. Berns, Determination of Constant Hue Loci for a CRT Gamut and Their Predictions Using Color Appearance Spaces, Color Res. Appl. 20, pg. 285-295 (1995).

[14] R. W. G. Hunt, Measuring Colour, 3rd Edition (Fountain Press, England, 1998) pg. 315.

[15] S. N. Yendrikhovskij, Computing Color Categories from Statistics of Natural Images, J. Imaging Sci. Techno. 45, pg. 409-417 (2001).

[16] S.N.Yendrikhovskij, A Computational Model of Colour Categorization, Color Res. Appl, 26, pg. 235-238 (2001)

[17] S.N. Yendrikhovskij, F.J.J. Blommaert and H. de Ridder, Representation of Memory Prototype for an Object Color, Color Res. Appl., 24, 6, pg.393-410 (1999).

[18] Y. Chang, A Perceptual-based Image Description Method for Color-Image Retrieval Systems, Dept. Computer Science, TIT, Tokyo, Japan, (2005).

[19] Werner G. K. Backhaus, R. Kliegl and John S. Werner, Color Vision:Perspectives from Different Disciplines, Berlin; New York: Walter de Gruyter, (1998).

Author Biography

Zhaohui Wang received his MSc. degree in Imaging Science from the University of Derby (2004). He is currently a research student at the Dept. Colour and Polymer Chemistry in the University of Leeds. He is working on the field of image categorisation, scene referred profile/PCS, and memory colours on different medium.