

Fuzzy Colour Category Map for Content Based Image Retrieval

Matthew Seaborn Lee Hepplewhite

John Stonham

Department of Electrical and Electronic Engineering

Brunel University, Uxbridge, Middlesex, England

Matthew.Seaborn@brunel.ac.uk

Abstract

In this paper a new colour space for content based image retrieval is presented, which is based upon psychophysical research into human perception. It provides both the ability to measure similarity and determine dissimilarity, using fuzzy logic and psychologically based set theoretic similarity measurement. These properties are shown to be equal or superior to conventional colour spaces. Example applications are also demonstrated.

1 Introduction

Content based image retrieval (CBIR) is the retrieval of images that fit a user defined query utilising the visual content of the images in the database. This is done by isolating and extracting salient features from all the images within the database. When a query is made these features are searched to find the closest matches to the features defined by the user.

In smaller databases it is feasible to annotate the images with relevant keywords entered by hand, permitting retrieval on the same basis as a library catalogue. However, in large databases annotation becomes highly impractical and even if it were it could not possibly detail *all* of the features within the image, or as daVinci eloquently said ‘The poet ranks far below the painter in the representation of visible things’. Therefore it is clear features must be extracted from the images automatically.

Colour is by far the most used feature for image retrieval. All of the main CBIR systems, such as QBIC [3], Virage and VisualSEEk [10] have colour as their main feature from which other features are subsequently derived. Given the fundamental level at which colour is involved in the retrieval process, it is obviously important to quantify similarity between colours.

1.1 Colour Space

The existence of the colour space enables the description of a colour by a set of co-ordinates within that space, the most commonly known of which is the Red-Green-Blue (RGB) space used in monitors and the majority of image file formats.

Thus having defined a colour as a point in n -dimensional space ($n=3$ for RGB), common practice is to assume that colour similarity is inversely proportional to a distance metric in that space. However, this assumption proves inappropriate, particularly for the more commonly cited colour spaces (RGB etc.), since equal distances in colour space rarely match perceived equivalence in similarity. Several colour spaces have thus been proposed, such as CIE L^*a^*b , which attempt to achieve the desirable property of 'perceptual linearity'. Although the colour spaces proposed by the CIE almost meet the universally accepted Munsell standard, it is acknowledged that 'perceptual linearity' only occurs over small distances. Yet, because there is no alternative, large distances are commonly used to represent colour dissimilarity.

However, as Ewald Hering observed in the 1800's, human perception is opponent in nature; white vs. black, red vs. green and yellow vs. blue. As such there is no similarity between colours and their opponents, the colour 'reddish green' does not, and cannot, exist. As such there is no finite measurement of the 'perceptual distance' from red to green, or yellow to blue as they have no similarity what-so-ever.

The presence of these seemingly infinite distances would suggest that distances in colour space cannot be treated in a simple, Euclidean manner. The results of Wuerger *et al.* [14] supports this.

1.2 Colour Labels

Wertheimer [13] suggested that among perceptual stimuli there are 'ideal types' that are anchoring points for perception. Rosch [7] proposed that in certain perceptual domains, such as colour, salient prototypes develop nonarbitrarily. Indeed this was found to be so by Berlin and Kay [1] who, in their pioneering anthropological study, found that colour space was partitioned into a maximum of eleven basic colour categories of which three were achromatic (black, white, grey) and eight chromatic (red, green, yellow, blue, purple, orange, pink and brown). This partitioning was a universal tendency to group colour around specific focal regions.

However, despite the considerable amount of research generated by this finding, little was done in locating the basic colour categories in the colour spaces represented by specific colour order systems. There are two notable exceptions to this. The original work by Boynton *et al.* [2] identified the focal regions, centroids and consensus colours in the OSA colour space. This work was extended by Sturges *et al.* [11] into the more widely accepted Munsell colour space using more subjects and colour samples.

The aforementioned research has shown that, perceptually, colour is a combination of eleven basic categories and can be described either as one of these basic colours or as a combination thereof. This ability to describe what a colour *is* means that it can also say what it *isn't*.

Berlin and Kay's eleven basic colours provide a basis for a perceptually justified colour space, which could be utilised in image retrieval

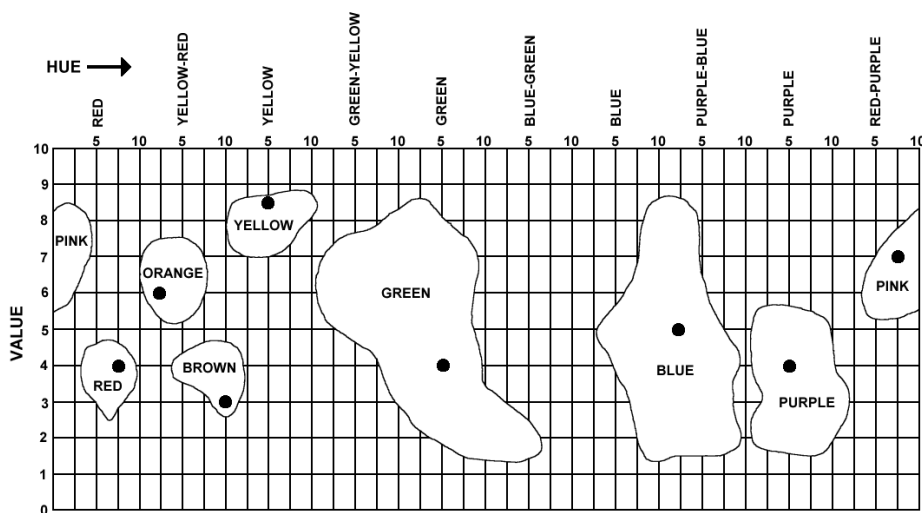


Figure 1: Location of consensus samples and focal colours on a two dimensional representation of the Munsell space as identified by Sturges *et al.*.

1.3 The Requirements of CBIR

The main purpose of content based image retrieval is to match the human perception: in the case of query-by-example, to find images that *look* like the example image the human user has chosen. However, when query-by-example is not possible the human needs to be able to describe their requirements to the retrieval system. In the case of colour this could be by an example colour, but it would be more convenient to give a description of a colour in a way easy for a human, such as an 'Orange - Red' or 'Blue'. Describing colour using the labelling found in human psychology by Berlin and Kay within a CBIR system has numerous advantages. It provides a perceptually salient, and low dimensional, method of segmenting a colour space for the purposes of generating such features as histograms, and provides an automatic annotation of colours for text based retrievals. An invaluable feature in database retrieval is the ability to discard potential retrievals that do not match the user description, which clearly can be done by labelling, but not by using a distance metric in a colour space, since that forces the entire database to be ranked in order of distance.

In this paper the Fuzzy Colour Category Map is presented and described and is shown to possess all of the above properties.

2 Colour Consensus Areas

Sturges *et al.* [11] identified consensus areas, groups of colours that were consistently labelled the same every time they were viewed by the test subjects, and details them in the Munsell colour space. These areas, shown in Figure 1, can be considered absolutely one colour without any hint of another colour present. It

should be noted that these areas are far from uniform in size or shape and that the boundaries are not equidistant from the focal points.

2.1 Munsell

The Munsell colour space is almost universally accepted as the standard, perceptually linear and uniform colour space to which others aspire. It is polar in nature, much like the more well known HSV, and has three dimensions; Hue (angle in the colour wheel), Value (luminance) and Chroma (strength of colour). However, unlike HSV, there is no simple transform for obtaining Munsell from RGB data. In order to achieve this the authors gamma corrected the RGB values to account for the non-linearity of the monitor used and transformed the result to the CIE standard xyY colour space. The Munsell Renotations were then obtained by interpolating from the data of Newhall, Nickerson and Judd [6].

2.2 Segmentation

It can be seen in Figure 1 that there is space between the consensus areas where the colour is not clearly labelled. Previously, the authors segmented the colour space based on which consensus label was closest [9] and this was shown to be successful in the generation of a histogram as a feature for image retrieval. However, while this allows for crisp colour descriptions such as 'Blue' and 'White' it does not allow for ones like 'Greenish Blue'. In essence it is too crisp, what is needed is some generalization. However, this generalization must also match human perception.

3 Fuzzy Colour Categories

The space between the consensus areas are neither one basic colour or another: they are a combination of those near it. It can be said that each point has a given membership in a given basic colour based on its distance from the basic colours area. To specify this membership the Fuzzy C-Means algorithm was chosen.

3.1 Fuzzy C-Means

It is proposed that eleven fuzzy patches, one for each basic colour, is applied to a colour space to represent the eleven basic colours identified by Berlin and Kay. Each and every point within the colour space then has a given membership in each of the patches representing how similar it is to the basic colour the patch represents. So a colour which is equally green and yellow would have 50% membership in green and 50% in yellow.

In this paper the authors use the Fuzzy C-Means algorithm to create these fuzzy patches. There are two functions in Fuzzy C-Means which define the fuzzy patches, one that defines the 'prototypes', the centre of the fuzzy patches, and another that defines any given point's membership to any given patch based on its distance from the patch's prototype. It is clear that the consensus areas defined by Sturges can be said to be the prototypes of the fuzzy patches of the human

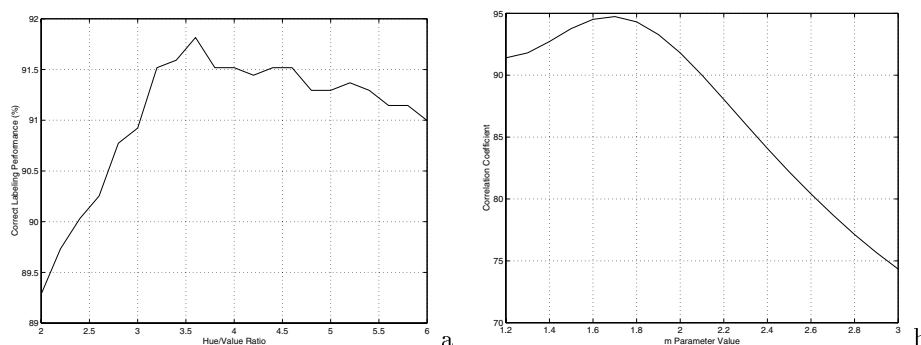


Figure 2: Psychophysical results for the generation of the FCCM. (a) Percentage of colour sample labels calculated by distance from consensus areas when compared with human defined ground truth. (b) Correlation between generated confusion matrix and Sturges' confusion matrix by varying the m parameter.

colour visual system. Thus the membership of any given point in the colour space to any given patch is defined as in Equation 1

$$\mu_{C_i}(x) = \frac{1}{\sum_{j=1}^k \left(\frac{\|x-v_i\|^2}{\|x-v_j\|^2} \right)^{\frac{1}{m-1}}} \quad 1 \leq i \leq k, x \in X \quad (1)$$

where x is a point in the colour space and v_i is the location of the i^{th} prototype and k is the number of patches. The parameter m is a weight that determines the membership of partial members of a patch.

3.2 Hue vs Value

To calculate the membership function it must be possible to calculate the distance from a the prototypes (consensus areas). The Munsell Renotation System implicitly defines the Hue dimension to be ten times more significant than the Value as Hue ranges from 1 to 100, Value 0 to 10. While this may be true over short distances, for the purposes of the fuzzy patches this is not necessarily so.

To determine the correct ratio a set of 1400 tiles representing all the colours in the Value range of 2 to 8.5 in steps of 0.5 and Hue of 1 to 100 in unit steps were taken. The chroma of each was the maximum attainable for each Hue - Value combination for the monitor. Values below 2 and above 8.5 were deemed too pale for the colour to be clearly seen on a monitor. This set was then presented to 9 individuals who were asked which of the eight chromatic basic colours they would label each sample. They were also given the *don't know* option. The *mode*, or most common response, for each colour was taken to be the ground truth of the label of that colour. If this resulted in a *don't know* then that sample was ignored for the purposes of calculating the performance of each of the distance ratios.

This was then compared to the label predicted by the relative distance from the edges of the consensus areas using a range of different distance ratios. The area/prototype to which the colour is the closest is defined as its label. The

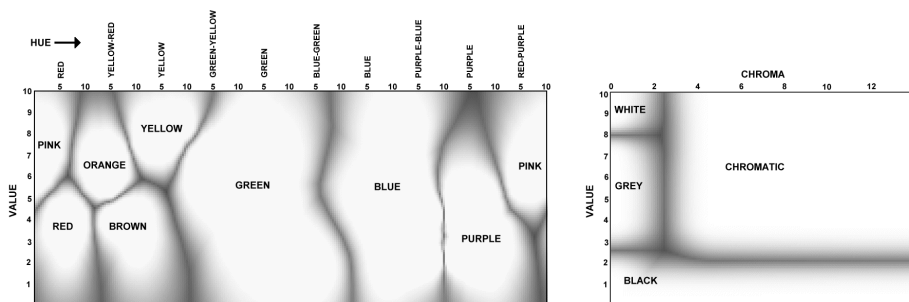


Figure 3: Maximum membership at all points in Munsell colour space.

percentage of correct calculated labels when compared with the ground truth is shown in Figure 2a. As can be seen it peaks at a Hue/Value ratio of 3.6.

3.3 Confusion and Entropy

Sturges [11] defines a *confusion matrix* identifying the number of times individuals were confused as to which of two basic colours a given sample most resembled. This matrix is used here to calculate the parameter m in Equation 1 which defines the distance 'squashing' property of the membership function. A value of 1 results in a step response which results in a point having 100% membership in the cluster with the nearest prototype. An m of 3 effectively results in a linear interpolation between prototypes.

The confusion between two fuzzy patches is defined in this paper as the *fuzzy entropy* of the two patches. Mathematically this is defined [5] as the ratio between the sizes of the underlap and overlap of the fuzzy set. Thus if A is the intersection of two fuzzy sets then the entropy is given by Equation 2.

$$FuzzyEntropy\{A\} = \frac{FuzzyCard\{A \cap A^c\}}{FuzzyCard\{A \cup A^c\}} = \frac{L^0(A, A_{near})}{L^\infty(A, A_{far})} \quad (2)$$

The Minkowski distancing metrics used were found to represent the colour confusability well.

For a range of values of m the fuzzy entropy matrix, as defined in Equation 2, was compared with the confusion matrix of chromatic colours found by Sturges by calculating the correlation between the generated confusion matrix and Sturges' confusion matrix. From the results shown in 2 it can be seen the maximum correlation occurs at $m = 1.7$.

The resulting fuzzy patches in the Munsell colour space can be seen in Figure 3.

4 Similarity Measurement

Numerous perceptual similarity metrics have been proposed, including distance, set theoretic etc. and a comparison of these was made by Jain *et al* [4]. Tversky [12] proposed a new method for the measurement of similarity which did not

use any distances, but relied on the presence and/or absence of features in the determination of similarity. This similarity metric allows the triangle inequality to exist and for objects to be completely dissimilar, as is found in perception. However, the features in Tversky's model were binary. Santini [8] extended this to the Fuzzy Feature Contrast (FCC) model and defined similarity as shown in Equation 3

$$\begin{aligned}
 S(\phi, \psi) = & \sum_{i=1}^p \min\{\mu_i(\phi), \mu_i(\psi)\} \\
 & - \alpha \sum_{i=1}^p \max\{\mu_i(\phi) - \mu_i(\psi), 0\} \\
 & - \beta \sum_{i=1}^p \max\{\mu_i(\psi) - \mu_i(\phi), 0\} \quad (3)
 \end{aligned}$$

Due to its obvious perceptual foundation this is the method utilised for similarity using the Fuzzy Colour Category Map.

5 Evaluation

To examine the performance to the Fuzzy Colour Category Map for colour similarity and discrimination, a test was performed much like the one performed by Wuerger *et al.* [14]. The test involved 600 sets of three randomly chosen, but different, colours shown to 9 individuals. The colours were presented in a row and the subjects were asked to identify which of the left and right colours were the most similar to the middle one. They were also given the option of both being equally similar and neither being similar. The results from these three tests were then correlated, and consistent responses were used in the compilation of the results.

5.1 Colour Similarity

The percentage of the colours that the both humans and the colour spaces found to be similar are shown in Figure 4a. The FCC parameters α and β were set to 0, making the similarity metric effectively the intersection of the fuzzy sets. In the cases where the FCCM gave a similarity of 1 for both colours in question (i.e. they were all in the same consensus area) the Munsell (HVC) colour space was used to measure similarity.

5.2 Colour Discrimination

Figure 4b shows the performance for each colour space separating similar from dissimilar colours, as defined by the human tests. For each of the colour spaces the discrimination point was set at the *optimal* distance to give the maximum performance.

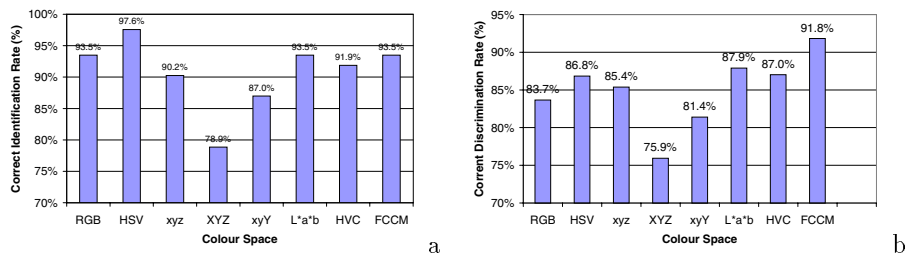


Figure 4: Similarity (a) and discrimination (b) performance of colour spaces when compared with human choices.

5.3 Discussion

FCCM is comparable to the other colour spaces in determining the most similar colour in the test. Upon examining the sets of colours that were incorrectly evaluated it was found that, with rare exceptions, the colours involved were blue and purple. The human results were indicating that what Sturges defined as consensus blues were similar to consensus purples. These colours, however, were always near the edge of the consensus area.

As predicted the FCCM was superior at discrimination to the other colour spaces using Euclidean distance as a dissimilarity measure. Examining the distribution of the similar/dissimilar colour distances shown in Figure 5 where it is clear that the performance of FCCM is much less dependant on the position of the thresholding boundary than even the best of the distance metrics.

The slight rise in FCCM defined dissimilar but humanly defined similar colours was again due to the purple-blue similarity mentioned above.

Thus it is clear that the FCCM performs comparably as a colour similarity measure whilst providing the ability to identify dissimilar images, which is of considerable use in database retrieval.

The main flaw has been seen to be the apparent perceptual similarity of purple and blues. Future work may be to examine the consensus areas of blues and purple with the view to revising them.

5.4 Example Applications

The FCCM is designed for use in the field of CBIR. Figure 6a demonstrates a query in the *PiffARO* 2 image retrieval system using FCCM colour labelling. The query is for 4 objects, one purple, one yellow, one pink and one red with the loose structural framework of the purple being to the right of the pink and the red below the yellow. The successful retrieval is shown in Figure 6b with the matching areas outlined.

Figure 7 demonstrates an alternative application of image segmentation/feature extraction. Using Tversky's measure and FCCM, the mean similarity of each pixel and its immediate neighbours is shown. Uniform areas have a high similarity where as edges have a low local neighbourhood similarity.

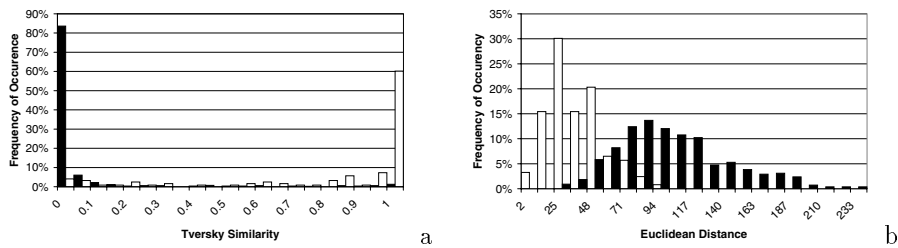


Figure 5: Histograms of the similarity/distance of human defined similar (shown in white) and dissimilar (shown in black) colours of (a) FCCM and (b) L^*a^*b .

6 Conclusion

In this paper the Fuzzy Colour Category Map has been presented. Based on psychophysical research it demonstrated the shortcomings of conventional colour spaces in the application of content based image retrieval, and shows that the FCCM fulfills the requirements that CBIR has of a colour space.

The FCCM is shown to be a versatile and powerful colour space with applications in colour similarity, colour feature representation and image feature extraction.

References

- [1] B. Berlin and P. Kay. *Basic color terms: their universality and evolution*. University of California Press, 1969.
- [2] R. M. Boynton and C. X. Olson. Locating basic colors in the o.s.a. space. *Color Research Applications*, 12:94–105, 1987.
- [3] C. Faloutsos, M. Flickner, D. Petkovic, W. Niblack, W. Equitz, and R. Barber. Efficient and effective querying by image content. Technical Report RJ 4953 (83074), IBM, August 1993.
- [4] R. Jain, S.N. Jayaram Murthy, and P.L.-J. Chen. Similarity measures for image databases. In *Proc. of SPIE on Storage and Retrieval for Image and Video Databases III*, volume 2420, pages 58–65, 1995.
- [5] B. Kosko. *Neural Networks and Fuzzy Systems: a dynamical systems approach to machine intelligence*. Prentice-Hall International, 1992.
- [6] S.M. Newhall, D. Nickerson, and D.B. Judd. Final report of the o.s.a. subcommittee on the spacing of munsell colors. *Journal of the Optical Society of America*, 33(7):385–481, July 1943.
- [7] E. Rosch. Cognitive reference points. *Cognitive Psychology*, pages 532–547, 1975.

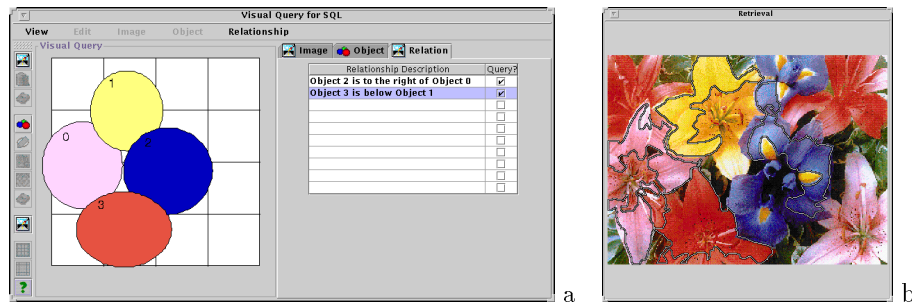


Figure 6: Example query (a) and retrieval (b) using FCCM colour labeling in the PiffARO 2 retrieval system.



Figure 7: Original image (a) and its local neighbourhood FCCM similarity (b).

- [8] S Santini and R Jain. Similarity queries in image databases. Technical report, Univeristy of California.
- [9] M Seaborn, L Hepplewhite, and J Stonham. Pisaro: perceptual colour and texture queries using stackable mosaic. In *Proceedings of the ICMCS*, volume 1, pages 171–176. IEEE Computer Society, June 1999.
- [10] J.R. Smith and S-F. Chang. Visualseek: a fully automated content-based image query system. Technical report, Columbia University, September 1996.
- [11] J Sturges and T W A Whitfield. Locating basic colours in the munsell space. *Color Research and Application*, 20:364–376, 1995.
- [12] A. Tversky. Features of similarity. *Psychological Review*, 84:327–352, July 1977.
- [13] M Wertheimer. *Numbers and numerical concepts in primitive peoples*. Harcourt, Brace and Co., 1938.
- [14] S.M. Wuerger, L.T. Maloney, and J. Krauskopf. Proximity judgments in color space: Tests of a euclidean color geometry. *Vision Research*, 35(6):827–835, 1995.