Colour Statistics for Matching in Image Databases[1)]

Allan Hanbury and Walter G. Kropatsch
Pattern Recognition and Image Processing Group
Vienna University of Technology
Favoritenstraße 9/1832
A-1040 Vienna, Austria
{hanbury, krw}@prip.tuwien.ac.at, http://www.prip.tuwien.ac.at

Abstract:
The matching of colour images has important applications in both industrial inspection and multimedia. We present a framework for the matching of colour images represented by statistical descriptors of each colour component in a 3D-polar coordinate (hue, saturation and brightness) system, which includes the necessary adaptations to take the angular nature of the hue into account. The latter includes the calculation of hue statistics using the circular statistics formalism, and the adaptation of the colour distance function to take these angular values into account. It is shown that the correct use of circular statistics improves the performance of such a colour matching algorithm in a content-based image retrieval application. Lastly, an industrial wood matching application demonstrating possible simplifications of this colour matching algorithm is described.

1 Introduction

The processing and analysis of colour images is complicated by the fact that these images are multi-spectral, the large number of ways for representing such images, and the differences in colours that can be caused by differences in the acquisition setup [8]. The matching of colour images has important applications in both industrial inspection and multimedia.

When calculating colour differences between pairs of colour images, the difference between colour feature vectors extracted from these images is usually used [11]. The most commonly used feature is the colour histogram [2]. As these histograms are usually high-dimensional, the use of statistical moments of each colour component of an image has been suggested when colour feature vectors of lower dimension are required. For example, Fuh et al. [5] use the means and standard deviations of the R, G and B colour components. Stricker and Orengo

[1)] Part of this work was done while Allan Hanbury was with the Centre de Morphologie Mathématique, Ecole des Mines de Paris, France. It is supported by the Austrian Science Foundation (FWF) under grants P14445-MAT and P14662-INF.

[14] suggested calculating these statistical moments in a 3D-polar coordinate colour space, the HSV space, in order to reduce the correlation between colour components and separate the luminance and chrominance information to some extent. This approach was later reused in a series of tests by Ma et al. [9]. In both these papers, the fact that the hue in the HSV colour space is an angular value is ignored, and standard linear statistics are used. This could lead to some incorrect results when the hue values of an image lie on either side of the origin. For example, in an image which is predominantly red, i.e. which contains hue values clustered around the origin, the standard linear mean is 180°, in the cyan region of the hue circle. This problem has led Daul et al. [3] to claim that this type of matching is only applicable to hue-homogeneous surfaces, such as wood.

In this paper we present a framework for the matching of colour images represented in a 3D-polar coordinate colour space. This includes the specification of the necessary statistical descriptors (section 2) and the distance function (section 3) which takes the angular nature of the hue into account. The applications limited to hue-homogeneous surfaces can be seen to be a special case in this framework. As there is often much confusion as to which version of the HSV space should be used (Ma et al. [9] describe three versions), we employ the 3D-polar coordinate space which satisfies the prerequisites for use in image analysis developed in [7]. While algorithms for the filtering of images containing circular data have been developed [6, 10], no colour matching algorithm which explicitly takes the angular nature of the hue component of a 3D-polar coordinate colour space into account has been presented. The angle used by Androutsos et al. [1] is not the difference in hue values, but the angle between two RGB vectors at the origin. A further innovation is that the circular variance of the hue is taken into account in the automatic determination of the variances of each feature from a training set, which removes the necessity of having to set weights for the various feature differences manually, as done in [14]. We show in section 4 that in a content-based image retrieval [12] application, description of the hue channel using circular statistics improves the performance. We finally demonstrate, in section 5, an industrial application which matches hue-homogeneous surfaces. When the coloured surface being treated has an inherent constraint (here hue-homogeneity), the suggested algorithm can be simplified.

2 Colour image statistics

Colour images are most often represented in the RGB colour space. However, there is often a strong correlation between the values in these three channels [3]. The use of a colour representation which decorrelates the values in these channels, such as a 3D-polar coordinate representation in terms of hue, saturation and brightness, can therefore be recommended when colour statistics are to be calculated. A large number of commonly available 3D-polar coordinate colour spaces, such as HLS, HSV, etc., are nevertheless not suitable for use in calculating colour statistics [7]. In this paper, we use the IHLS (improved HLS) colour representation suggested in [7]. The main difficulty with using such a 3D-polar coordinate colour representation is that the hue is an angular value. This means that the hue values are cyclic, and that hues of 0° and 360°, even though they correspond to the same point on circle, have widely separated numerical values. The use of circular statistics [4] compensates for this problem. In this section we summarise the standard circular statistics descriptors.

The pth trigonometric moment of a set of n hue values H_i has direction

$$\hat{\mu}_p = \arctan\left(\frac{\bar{S}_p}{\bar{C}_p}\right) \tag{1}$$

where

$$\bar{C}_p = \frac{1}{n}\sum_i \cos pH_i, \quad \bar{S}_p = \frac{1}{n}\sum_i \sin pH_i \tag{2}$$

and magnitude

$$\hat{\rho}_p = \sqrt{\bar{S}_p^2 + \bar{C}_p^2} \tag{3}$$

The *mean direction* of a set of hue values is simply the direction of the first trigonometric moment $\hat{\mu}_1$, which corresponds to the direction of the resultant vector obtained by adding a unit vector in each of the directions in the set. The necessary care should be taken to expand the output of the arctan function into the range $[0, 360°]$. The *mean length* of this set of hue values is $\hat{\rho}_1$. It can be interpreted as the length of the actual resultant vector compared to the length of the resultant vector which would be obtained if all the H_i were identical. The value of the mean length is in the range $[0, 1]$ and can be used as an indicator of the dispersion of the data (similar to the variance). If $\hat{\rho}_1 = 1$, all the H_i are coincident. Conversely, a value of zero does not necessarily indicate a homogeneous data distribution, as certain non-homogeneous distributions can also result in this value. The *circular variance* is defined as

$$V = 1 - \hat{\rho}_1 \tag{4}$$

This variance differs from the standard linear statistical variance in being limited to the range $[0, 1]$, however it is similar in that lower values of V indicate that the data is less dispersed. The circular variance is converted into a circular standard deviation [4] by

$$\hat{\sigma} = \sqrt{-2\log(1 - V)} \tag{5}$$

Finally, the circular *skewness* is defined as

$$\hat{s} = \frac{\hat{\rho}_2 \sin(\hat{\mu}_2 - 2\hat{\mu}_1)}{(1 - \hat{\rho}_1)^{3/2}} \tag{6}$$

By making the length of the vectors dependent on the saturation associated with each hue, one can calculate a saturation-weighted hue mean direction. The hues associated with small saturation values will therefore have less influence on the direction of the resultant vector.

Given n pairs of values, the hue H_i and its associated saturation S_i, we proceed as before, except that instead of finding the resultant of a set of unit vectors, the vector with direction H_i is assigned length S_i. This weighting is done in the calculation of the first trigonometric moment by replacing equation 2 by

$$\bar{C}_1^S = \left(\frac{1}{\sum_{i=1}^n S_i}\right)\sum_{i=1}^n S_i \cos H_i, \quad \bar{S}_1^S = \left(\frac{1}{\sum_{i=1}^n S_i}\right)\sum_{i=1}^n S_i \sin H_i \tag{7}$$

and replacing $\bar{C}_1$ and $\bar{S}_1$ in equations 1 and 3 by $\bar{C}_1^S$ and $\bar{S}_1^S$ to obtain the saturation-weighted quantities denoted by $\hat{\mu}_1^S$ and $\hat{\rho}_1^S$.

3 Colour image matching

For each colour image, we calculate nine features in the IHLS representation. These are the standard linear statistical mean, variance and skewness for the luminance and saturation. For the hue, we calculate the circular versions of these quantities. In the experiments that follow, the performance when using the standard circular values $\hat{\mu}_1$, $\hat{\rho}_1$ and $\hat{s}$ is compared to that when using the saturation-weighted values, namely $\hat{\mu}_1^S$, $\hat{\rho}_1^S$ and $\hat{s}$ (the same measure of circular skewness is used in both these cases). A weighted Euclidean distance between the nine image features (the first three statistical moments of each channel) is used to quantify the colour difference between two images. The weights (i.e the variances) are determined using a training set containing manually grouped images. One should ensure that the angular-valued features are properly taken into account when calculating this distance.

For all features except for the hue mean direction ($\hat{\mu}_1$ or $\hat{\mu}_1^S$), the standard method for calculating a global diagonal covariance matrix given a training set containing pre-classified images is used. To calculate a global circular variance for the hue mean feature, we first calculate the circular variance V (equation 4) for each class in the training set, take the mean circular variance over all the classes, and finally use the square of the circular standard deviation (equation 5) as the intra-class variance. The distance between two colour images is then calculated as the sum of the squared differences of the corresponding features divided by the variance (i.e. the Mahalanobis distance with a diagonal covariance matrix). The only modification is that the difference between the hue mean features is calculated as the acute angle between them. In summary, each image feature vector contains 8 linear features which we notate by $x_1, \ldots, x_8$ and an angular feature notated as x_θ. After determining the intra-class variances $v_1, \ldots, v_8$ and v_θ of each feature as described above, the distance Δ_{AB} between the feature vectors of two colour images $\mathbf{x}^A$ and $\mathbf{x}^B$ is

$$\Delta_{AB}^2 = \sum_{i=1}^{8} \frac{\left(x_i^A - x_i^B\right)^2}{v_i} + \frac{\left[\angle\left(x_\theta^A, x_\theta^B\right)\right]^2}{v_\theta} \tag{8}$$

where the operator $\angle$ returns the acute angle between two angular values. The units of this angle and of v_θ should be the same.

Number of images in group	2	3	4	5	6	7	8	9	10	11
Number of groups	28	4	3	1	2	1	1	0	0	1

Table 1: The sizes of the groups of similar images in the VisTex database. The top row gives the number of images in a group, and the second gives the number of groups of this size.

4 Application 1: Content-based image retrieval

The use of the first three statistical moments of each channel of the HSV colour space as a nine-component image colour feature vector in a content-based image retrieval application has been suggested in [14], although linear statistics were used for the hue. We tested the suggested matching algorithm on the "Reference Textures" subsection of the VisTex database[1)]. This database contains a total of 167 images, of which 123 were manually grouped into 41 differently sized groups of similar images. The distribution of the group sizes (i.e. number of images in the group) is given in Table 1. We compared the image matching results using the nine colour image features described previously with different types of hue statistics calculations, namely: standard linear hue statistics, saturation-weighted hue circular statistics and standard hue circular statistics. The colour matching algorithm described in section 3 is used, with the intra-class variances calculated using half the manually defined groups in the test database.

For visualising the results, we showed the five images which had the smallest distance to the query image. A set of matching images retrieved using non-saturation-weighted circular statistics are shown in Figure 1 (the query image is on the left). The design of evaluation methods for content-based image retrieval algorithms is not standardised, and many possible strategies exist [13]. We designed a particularly challenging set of evaluation criteria for testing the proposed algorithm. The VisTex database has been manually grouped by taking into account all pertinent colour and texture features. However, even though we are matching only by colour here, we use this grouping to measure how well the automatic grouping without any texture features performs (and hence the future improvement when texture features are incorporated). Therefore, even though the white brick wall in third retrieval position in Figure 1 matches the white flowers in terms of colour, it is still counted as an incorrect retrieval as it does not belong to the same manually defined group of similar objects. The evaluation criterion used is to examine the N images calculated as being closest to the query image, with $N \in \{1, 2, 3, 4\}$, and determine the percentage of cases for which *all* N images belong to the same class as the query image. For the evaluation at each value of N, only manually-defined classes which contain at least $(N + 1)$ images are used. This means that the total number of query images used in each of the tests is 123 for $N = 1$, 67 for $N = 2$, 55 for $N = 3$ and 43 for $N = 4$. In Figure 1, for example, there is a correct retrieval for $N = 1$

[1)] `http://www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html`

Figure 1: The set of images retrieved for the query image on the left using distances calculated with standard circular statistics for the hue.

Method	$N = 1$	$N = 2$	$N = 3$	$N = 4$
Hue linear stats	66.7	44.8	27.3	30.2
Sat-weighted hue circ. stats	72.4	38.8	25.5	23.3
Hue circ. stats	72.4	55.2	34.5	32.6

Table 2: The percentage of correctly retrieved image groups for the VisTex database. N refers to the number of closest images to the query image examined.

and $N = 2$ and an incorrect retrieval for $N = 3$. As this manually-defined "white flowers" class contains only four images, it is not taken into account for the $N = 4$ evaluation. The results of this evaluation method using feature vectors which contain linear hue statistics, hue statistics calculated using the standard circular statistics formulae, and saturation-weighted hue statistics are presented in Table 2.

One can see that the use of circular hue statistics gives consistently better results than the use of linear statistics, with an improvement of up to 10%. The performance of the saturation-weighted hue statistics, while good for the retrieval of the closest image in the database, is disappointing when the number of closest images examined is greater than one.

5 Application 2: Wood colour matching

The wood colour matching application presented here was developed in the framework of the European project SCANMATCH. Due to the real-time requirements of the application, three simplifications were made to the colour matching algorithm described above. Firstly, only the first two statistical moments of each channel of the IHLS image were used. This was found to be enough to describe the features present: the mean quantifies the overall wood colour and the standard deviation quantifies the contrast between the wood and the veins. Secondly, as the hue of an image of wood covers a small range of values which does not vary much between different images, even of different species, we could use standard linear statistics on the hue channel (this is applicable to any problem for which the hue values are not distributed across the hue origin). Finally, as the colour characteristics of each piece of wood are so similar, no manual classification was done before determining the global diagonal covariance matrix used

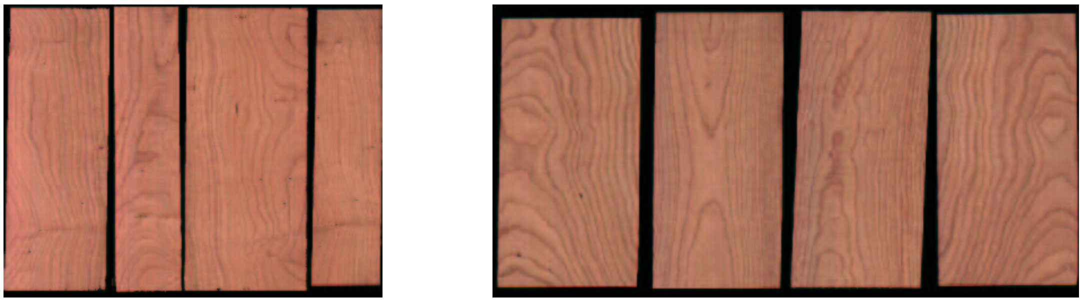

Figure 2: Two examples of the output of the wood colour matching algorithm on a database of 152 wild cherry plank images.

in calculating the distance between feature vectors. The whole database was simply taken to be one class. Two examples of this matching algorithm applied to a database of 152 wild cherry plank images are shown in Figure 2. The image on the left was selected by the user. The algorithm then chose the image from the database closest to this image, followed by the one closest to the second image, etc.

6 Conclusion

We have presented a framework for the matching of colour images represented in a 3D-polar coordinate colour space, unifying some independent work on the subject. Improvements on already presented methods include the correct use of circular statistics for the hue, and the formulation of a weighted Euclidean colour distance including angular hue differences. Matching of images with constraints, such as hue-homogeneity, is a special case of this general framework, as demonstrated by the industrial wood-matching application.

An improvement in image database retrieval performance is demonstrated on the VisTex database of colour textures, for which the classifier based on standard circular statistics is shown to have better agreement with the manual classification, especially when finding two or more images closest to the query image. The level of agreement between the manually classified images and automatically retrieved images is nevertheless rather low, this being because only colour information is considered in the automatic retrieval. Incorporation of texture information should improve this agreement. While the demonstrated improvement in retrieval performance is not large, it is of the order of what could be expected when one repairs some aberrant behaviour around the origin of a coordinate system.

It would be interesting to compare the performance of a similar colour matching algorithm in a perceptually uniform colour space such as L*a*b* or L*u*v*. The advantage of the current formulation in a 3D-polar coordinate representation of the RGB space is that one obtains a reasonable separation between the chromaticity and luminance components without the need for calibration of the white point and specification of the chromaticities of the primaries. The conversion to the IHLS coordinate system is also less calculation-intensive. For an industrial application in which each image is taken under identical lighting conditions with identical cameras, the extra complication introduced by a perceptually uniform colour space should not be necessary. In a perceptually uniform colour space, the use of image colour feature vectors containing the means of each of the channels will allow the perceptual uniformity to be taken into account in the calculation of the Euclidean distance between these vectors. If one includes contrast information in the form of standard deviations or higher order descriptors in the feature vectors, the direct advantage of the perceptual uniformity will be lost.

References

[1] D. Androutsos, K. N. Plataniotis, and A. N. Venetsanopoulos. A novel vector-based approach to color image retrieval using a vector angular-based distance measure. *Computer Vision and Image Understanding*, 75(1/2):46–58, July/August 1999.

[2] S. Boughorbel, N. Boujemaa, and C. Vertan. Histogram-based color signatures for image indexing. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU'2002*, 2002.

[3] C. Daul, R. Rösch, and B. Claus. Building a color classification system for textured and hue homogeneous surfaces: system calibration and algorithm. *Machine Vision and Applications*, 12:137–148, 2000.

[4] N. I. Fisher. *Statistical Analysis of Circular Data*. Cambridge University Press, 1993.

[5] C-S. Fuh, S-W. Cho, and K. Essig. Hierarchical color image region segmentation for content-based image retrieval system. *IEEE Transactions on Image Processing*, 9(1):156–162, 2000.

[6] A. Hanbury and J. Serra. Morphological operators on the unit circle. *IEEE Transactions on Image Processing*, 10(12):1842–1850, December 2001.

[7] A. Hanbury and J. Serra. A 3D-polar coordinate colour representation suitable for image analysis. Technical Report PRIP-TR-077, TU Wien, 2002.

[8] D. Lafon and T. Ramananantoandro. Color images. *Image Analysis and Stereology*, 21(Suppl 1):S61–S74, 2002.

[9] W-Y. Ma and H. Zhang. Benchmarking of image features for content-based retrieval. In *Proceedings of the 32nd Asilomar Conference on Signals, Systems and Computers*, volume 1, 1998.

[10] N. Nikolaidis and I. Pitas. Nonlinear processing and analysis of angular signals. *IEEE Transactions on Signal Processing*, 46(12):3181–3194, 1998.

[11] R. Schettini, G. Ciocca, and S. Zuffi. A survey of methods for colour image indexing and retrieval in image databases. In R. Luo and L. MacDonald, editors, *Color Imaging Science: Exploiting Digital Media*. John Wiley, 2001.

[12] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, December 2000.

[13] J. R. Smith. Image retrieval evaluation. In *Proceedings of the IEEE Workshop on Content-based Access of Image and Video Libraries*, 1998.

[14] M. Stricker and M. Orengo. Similarity of color images. In *Storage and Retrieval for Image and Video Databases (SPIE)*, pages 381–392, 1995.