WALAILAK JOURNAL
http://wjst.wu.ac.th Article

Colour Extraction of Agarwood Images for Fuzzy C-Means Classification

Mohamad Razi MAD AMIN[1], Siti Khairunniza BEJO[1,*], Wan Ishak WAN ISMAIL[1] and Shamsiah MASHOHOR[2]

[1]*Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia*
[2]*Department of Computer and Communication Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia*

([*]Corresponding author; e-mail: skbejo@eng.upm.edu.my)

Received: 26 December 2011, Revised: 29 April 2012, Accepted: 9 December 2012

Abstract

Colour is one of the quality features used in determining agarwood quality and grade. This research investigates the relationship of agarwood physical colour properties with its price. Colour features of agarwood images taken from Red, Green, Blue (RGB), Hue, Saturation, Intensity (HSI) and Commission Internationale de l'Eclairage standard L,*a,*b colorspace (CIELAB) has been extracted by Fuzzy C-Means (FCM) classification. The performance of these colorspaces has been determined using five cluster validity indices. One hundred and forty agarwood images consisting of seven different prices have been analyzed. From the experiment, it has been shown that CIELAB colorspace with four numbers of clusters gave more consistent and accurate results compared to the others. It also gave a significant relationship when tested using analysis of variance (ANOVA) and Duncan Multiple Range Test (DMRT). The method performs best when classifying lower price agarwood with component *L* for RM250 and RM800, *b* for RM350 and RM2500 while *a* for RM900. Overall, the proposed method proved that there is a significant relationship between agarwood price and its physical colour properties, which thus shows that the image processing has an enormous potential to be used in the agarwood chips grading task for the future development.

Keywords: Agarwood, colorspace, fuzzy c-means, classification, cluster validity

Introduction

The agarwood has been treasured and traded for its complex and pleasing incense which is commonly used in the perfume industry. Most of the quality of agarwood based products depends on the agarwood grade. High grade agarwood usually produces a pleasant aroma and holds the aroma for several days. The agarwood quality is graded based on several set of properties which are the amount of resin content, wood colour, fragrance, size, species, country of origin and etc. **Table 1** shows the properties being used to grade agarwood in Malaysia [1], Taiwan [2], Indonesia [3,4], Papua New Guinea (PNG) [5] and Japan [6,7].

Table 1 Properties used to grade agarwood.

Properties	Malaysia[1]		Taiwan[1,2]	Indonesia[1]	Papua New Guinea[1,3]	Japan[1]
	Literature	Jasben company				
Resin Content	√	√	√	√		√
Colour	√	√		√	√	√
Fragrance	√	√	√			
Size	√		√			
Country of origin	√					
Density	√				√	
Product purity	√					
Fibre		√				
Weight		√		√		√
Age			√			
Location			√			
Scarceness			√			
Shape					√	√
Flammability					√	

Note: [1]Aquilaria Malaccensis, [2]Aquilaria Crassana, [3]Aquilaria Falaria

Generally, the process of agarwood grading starts with the determination of the wood sinking ability [1,2,6-7]. The process is done by placing the agarwood chips in water. The high grade agarwood will sink while low grade agarwood floats. After that, the agarwood is graded through colour characteristics. The darker physical colour of agarwood with the same sinking ability gives higher grade and prices. The logic behind the manual method is that, high grade agarwood commonly constitutes a large amount of resin content hence affecting its density and physical colour properties [1,2]. In some cases, the fragrance will be evaluated by burning a small portion of agarwood chip. The higher grade will give a pleasant and long duration scent. Size of agarwood chips also influence the agarwood price. Usually, the larger size agarwood of the same grade will have higher price. The scarceness and country of origin will sometimes be selected as one of the grading criteria as it is believed that agarwood from Vietnam gives a higher price than other countries. However, there is still no systematic standard that relates agarwood colour and its grade reported. It is also difficult to create a standard because the colour pattern of agarwood is influenced by a lot of parameters i.e. species, defects, illumination, temperature, radiation and chemical substances [8]. In the grading process, human visualization and experience are relied on highly in determining the agarwood grade. This has led to misclassification and inconsistent grade due to fatigue and wrong judgment [9,10]. For that reason, only colour properties of *aquilaria malaccensis* species and seven different prices will be analyzed in this study.

The application of image processing and pattern recognition has been used in automatic grading technology. Preliminary work on automatic agarwood grade determination was done by Abdullah *et al.* [9] using percentage of overall image intensity in a grayscale colorspace. Five different agarwood grades were grouped based on five ranges of intensity which are 0 - 25 for grade A, 26 - 31 for grade B, 32 - 44 for grade C, 45 - 63 for grade D and 64 - 150 for grade E. The experimental result gave 80 % overall accuracy. However, the drawback of the system is it only considers several grades of agarwood chips in the market.

The colour of the natural product is influenced by wood species. Furthermore, chemical substances and environment consequently produce non-homogenous patterns in colour. Therefore, the definitions of colour boundaries in the natural product become a difficult task. The concept of natural colour obviously fits well with the description of the fuzzy approach. In this study, a fuzzy c-means algorithm will be used as the image segmentation process in order to classify each of the image pixels into their homogenous regions [11]. The main problem in segmentation of agarwood images was the difficulty in sequencing the images into satisfactory colour properties. The final outcome of the segmentation should produce a region that is homogenous (colour similarity) and has a low variation or stable through the sequence. In addition, the number of regions or clusters should be as small as possible in order the decrease the processing time.

The unsupervised classification can be used to automatically classify the agarwood colour pixels into several colour sets [12]. Fuzzy C-Means (FCM) is one of the popular unsupervised methods used in classification. Each pixel is classified based on its membership degree in the range of 0 to 1. These membership degrees represent the probability of that pixel belonging to a specific cluster. Kuo *et al.* [13] has applied FCM to cluster the printed fabrics images in RGB colorspace [13]. A similar approach has been done by Ronghua *et al.* [14] for yarn-dyed fabrics in three different colour model i.e. Red, Green, Blue (RGB), Hue, Saturation, Intensity (HSI) and CIELAB. Based on the experiment, the CIELAB colour model gave a more effective colour extraction of yarn-dyed fabrics compared to the other models. Kang *et al.* [15] has applied FCM classification in HSI colour model for dental plaque quantification. Another study done by Saikumar *et al.* [16] used FCM algorithm and CIELAB colorspace to find meaningful region of satellite imaginary. On the other hand, several researchers have proposed a combination of FCM with another image processing algorithm or pattern recognition process to find the best region of interest based on colour similarity of a subject [17-19]. As an example, Patmavathi and Muthukumar [17] used a combination of FCM and a thresholding algorithm to identify underwater images. Grayscale images were used as the input. Chuai-Aree *et al.* [18] has introduced a combination of FCM and neighbourhood smoothing technique in segmenting a document. It gave satisfactory results in defining text, image and background. Ghaleh and Behrad [19] combined FCM with an active contour model in finding a lip contour. RGB lip colour images were used. The algorithm showed good lip segmentation even in different speakers with different conditions of illumination. A comparable study of FCM, combination of FCM with morphology and nearest neighbor method has been done by Sopharak *et al.* [20] to detect exudates in the eyes using HSI colour images. The result showed that FCM gave the highest classification sensitivity (97.2 %) while a combination of a FCM and morphology algorithm gave the highest classification specificity (99.2 %). Therefore, it can be concluded that the use of FCM in image classification is promising. It can be used to group pixels with similar properties into the same cluster.

Cluster validation is a very important issue in clustering analysis. Pal and Bezdek [21] has introduced two types of cluster validity indices which are Partition Coefficient (PC) and Classification Entropy (CE). Other work done by Gunderson has proposed separation coefficient indices which consider the data geometrical properties. The other indices that were introduced such as Fukuyama and Sugeno (FS), Xie and Beni's (XB), Fuzzy Hypervolume (FHV), Partition Density (PD), Separation (S), Separation Compaction (SC), Dunn Index (DI), etc. Most of these indices measure the compactness and separation in the cluster properties [22-24]. The compactness of the cluster measures the variance within the cluster which indicates how different each object in the cluster. The separation measures the distance between each cluster (isolation). Good clustering normally will produce a high compactness within a cluster and high isolation between clusters [23,25].

The literature has shown that most of the countries that supply agarwood use colour as one of the criteria in deciding the agarwood grade. Therefore, the objective of this study is to extract suitable colour features of agarwood images for agarwood grade classification.

Material

One hundred and fourty agarwood images consist of seven different prices i.e. RM250, RM350, RM800, RM900, RM1000, RM2500 and RM3500 have been used in this study. All of the samples have been graded by an expert at the Malaysian Institute for Nuclear Technology Research (MINT). As shown in **Figure 1**, it is hard for the non-expert to classify its price since no clear trend can be observed. All algorithms were coded using a Matlab programming language and the program was run on the PC Intel Core 2 Duo 2.2 GHz processor with 2 GB of RAM.

Methodology

In general, the methodology consists of six main steps as follows:

Step 1: Image acquisition
Step 2: Image segmentation
Step 3: Colour transformation
Step 4: FCM clustering
Step 5: Performance test
Step 6: Statistical analysis

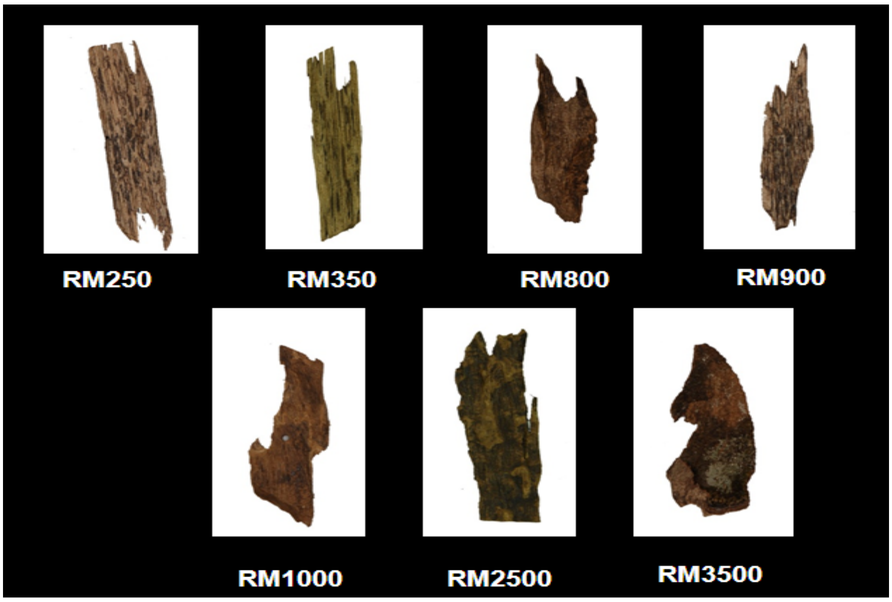


Figure 1 Agarwood images in different prices.

The agarwood images were acquired using a Nikon D200 CCD camera with an ISO setting of 200, 105 mm focal length and without a flash. The images were stored in a JPEG format. The camera was setup at 1.19 m with a 90° angle between the camera and the samples. It was conducted in a controlled room size (3.35 × 2 m) with a 2 unit overhead fluorescent lighting (Model: Philips, TLD 36W/54) 3.65 m from the ground.

Image segmentation was performed to eliminate background and unwanted object in the image. In this study, histogram thresholding was used and applied to the RGB colour images. The process involved constructing a histogram distribution of each RGB band for each agarwood price. The image was thresholded as in Eq. (1):

$$\mathrm{Im}(R,G,B) = \begin{cases} \mathrm{Im}(R,G,B), & \mathrm{Im}(R,G,B) \le TH \\ 255, & Im(R,G,B) > TH \end{cases} \tag{1}$$

where *Im(R,G,B)* is the original pixel values and *TH* is the threshold value (180). Since the agarwood colour properties is in the range of black to brownish colour which tends to be represented by a lower pixel intensity value, therefore, unwanted pixels (background) were assigned to 255 (white). Only pixels with a value less than or equal to 180 were analyzed during the grading process. Since the focus of this research is to extract the best colour features without considering the final structure of the agarwood image, the holes that existed in the image after applying the threshold will not be considered since it does not affect the final classification result.

In this study, three colorspaces i.e. RGB, HSI and CIELAB were analyzed. The HSI colour component can be transformed from RGB colorspace coordinate as in Eq. (2) - (5) [26,27].

$$I = \frac{1}{3}(R + G + B) \tag{2}$$

$$S = 1 - \frac{3}{R + G + B} \tag{3}$$

$$H = \begin{cases} \theta & if \quad B \leq G \\ 360 - \theta & if \quad B > G \end{cases} \tag{4}$$

$$\emptyset = cos^{-1}\left\{\frac{\frac{1}{2}[(R-G)]+(R-B)}{[(R-G)^2 + (R-B)(G-B)]}\right\} \tag{5}$$

A specific organization called CIE (Commission Internationale de l'Eclairage) has determined standard values that are used worldwide to measure colour.

The three coordinates used by CIE are called L*, a* and b* and the color measurement method is called CIELAB. This colour component is independent and has the ability to measure low colour contrast in an image. The L*, a* and b* component can be obtained by converting the RGB colorspace coordinates into the cube root XYZ matrix transformations (Eq. (6)) and then the L, $a*$ and $b*$ component is derived as in Eq. (7) - (9) [26,27]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6}$$

$$L = 116\left[\sqrt[3]{\frac{Y}{Y_0}} - \left[\frac{16}{116}\right]; \frac{Y}{Y_0} > 0.01\right] \tag{7}$$

$$a* = 500\left[\sqrt[3]{\frac{X}{X_0}} - \sqrt[3]{\frac{Y}{Y_0}}\right]; \frac{X}{X_0} > 0.01 \tag{8}$$

$$b* = 200\left[\sqrt[3]{\frac{Y}{Y_0}} - \sqrt[3]{\frac{Z}{Z_0}}\right]; \frac{Z}{Z_0} > 0.01 \tag{9}$$

where, X_o, Y_o and Z_o are the X, Y and Z values with standard white reference (0.9642, 1, 0.8249).

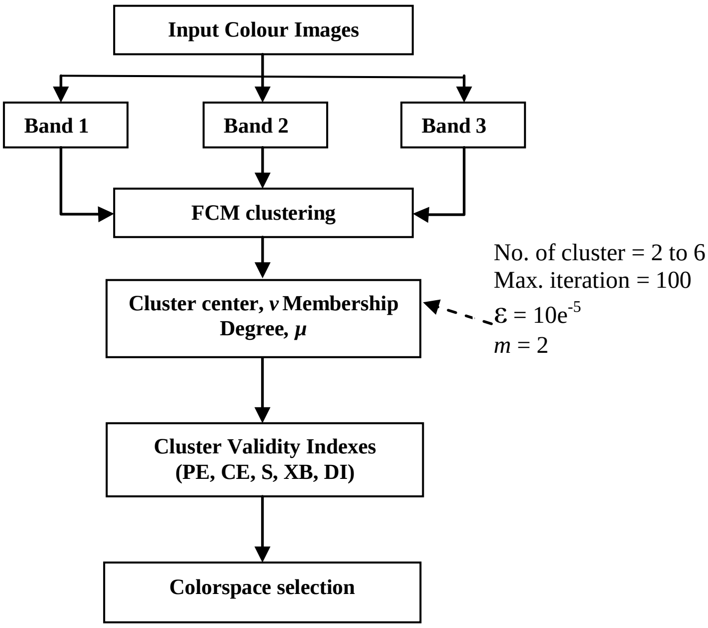


Figure 2 Methodology of colour clustering using FCM algorithm.

Figure 2 shows the flowchart of the colour clustering process used in this study. The purpose of the clustering algorithm in this study is to find the agarwood image colour cluster center without using human intervention in order to classify the colour boundary. For this research, it is an important process since until today no accurate information available on colour standard of agarwood has been published. In this study, colour images for all the selected colorspaces (RGB, HSI and CIElab) were used as the input in the FCM clustering algorithm. Each colour band is transformed to its membership degree in the range of 0 to 1. These membership degrees represent the probability of that the pixel belongs to a specific cluster. For instance, if data $X = (x_1, x_2, x_3, \ldots. x_n)$ denotes an image with n pixels to be classified into c cluster where x_i represents multispectral data, the algorithm for an iterative optimization that minimizes the cost function is defined as in Eq. (10):

$$J_m(U,V) = \sum_{j=1}^{n}\sum_{i=1}^{c} u_{ij}^{m} \left\|x_j - v_i\right\|^2 \quad (10)$$

where μ_{ij} represents the membership of pixel x_j in the ith number of cluster, v_i is the ith cluster center and m is a weighing exponent, $1 \leq m \leq \infty$. The cost function is to assign pixels with high membership degrees with the closer ith cluster center and assign pixels with low membership degrees to the farthest ith cluster center. In FCM methods, the membership degree depends on the distance between the pixel and each cluster center. The membership degree and cluster centers are defined as in Eq. (11) - (12) respectively:

$$\mu_{ij} = \left[\sum_{K=1}^{C}\left(\frac{\left\|X_j - V_i\right\|^2}{\left\|X_j - V_k\right\|^2}\right)^{\frac{2}{m}-1}\right]^{-1} \quad (11)$$

$$v_i = \frac{\sum_{j}^{n}(\mu_{ij})^m x_j}{\sum_{j=1}^{n}(\mu_{ij})^m} \quad (12)$$

The FCM algorithm starts with an initial guess of the cluster center. The determination of cluster centers is done by iteratively minimizing the sum squared distances between the guess cluster center and the calculation cluster center. This will move all the cluster centers to the correct location. The algorithm will stop when one of the numbers of iterations or the termination condition is met. In this study, two to six numbers of clusters have been tested and analyzed. The weighing exponent, number of iterations and the termination tolerance (ε) has been setup to be 2, 100 and $10e^{-5}$, respectively.

After the clustering process, all of the pixels will have their own membership degree, μ. Meanwhile, each image will have several cluster centers, v based on the number of clusters applied. These data will be used as the input in cluster validity index determination. Since the focus of the colour classification is to find the colour set of agarwood based on colour similarity, therefore, the main criteria of the selected indices were based on the variance determination within clusters and distance measurement between each cluster. Five cluster validity indices i.e. PC, CE, S, XB and DI which met this criterion have been selected.

PC index measures the amount of overlapping between clusters where $1/c \leq PC(c) \leq 1$. The $\max_{2\leq c\leq n-1} PC(c)$ produces the best clustering performance for the corresponding datasets. The index is defined in Eq. (13) [23,28]:

$$PC(c) = \frac{1}{N}\sum_{i=1}^{c}\sum_{j=1}^{N}\left(\mu_{ij}\right)^2 \tag{13}$$

The CE index measures the fuzziness of the cluster partition where $0 \leq CE\ (c) \leq \log_2 c$. The $\min_{2\leq c\leq n-1} CE(c)$ produces the best clustering performance for the corresponding datasets. The index is defined in Eq. (14) [23,28]:

$$CE(c) = -\frac{1}{N}\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}\,log\left(\mu_{ij}\right) \tag{14}$$

The S index is the index which uses a minimum-distance separation to produce the best clustering performance. The index is defined in Eq. (15) [28]:

$$S(c) = \frac{\sum_{i=1}^{c}\sum_{j=1}^{n}\left(\mu_{ij}\right)^2\left\|x_j-v_i\right\|^2}{N_{min_{i,k}}\|V_k-V_i\|^2} \tag{15}$$

The XB index aims to quantify the ratio of the total variation within clusters and the separation of clusters. The best clustering performance is the minimum value of the index [29]. The index is defined in Eq. (16) [24,29,30]:

$$XB(c) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\mu_{ij}^{m}\left\|x_j-v_i\right\|^2}{n_{min_{ij}}\left\|v_i-v_j\right\|^2} \tag{16}$$

Dunn's Index (DI) was originally proposed to be used to identify compact and well separated clusters [24,29,30]. Here $d(c_i, c_j)$ is the minimum intercluster distance from cluster c_i to c_j, and $diam(c_k)$ is the maximum intracluster distance from one point in cluster c_k to another. The maximum DI value is taken as the optimal number of the clusters and thus gives the best cluster performance. It is defined in Eq. (17). The main drawback of DI is that it is computationally expensive and time consuming as c and N increase [24].

$$D_{n_c} = \min_{i=1\ldots n_c}\left[\min_{j=i+1\ldots n_c}\frac{d(c_i,c_j)}{\max_{k=1\ldots n_c} diam(c_k)}\right] \tag{17}$$

In this study, all statistical analysis was performed using SPSS 16.0. An analysis of variance (ANOVA) test was conducted to assess whether the means of the dependent variable were significantly different among group or not [31]. Duncan Multiple Range Test (DMRT) has been chosen to find specific colour features for each agarwood price since it is a result guided test that compares the treatment means while controlling the comparison-wise error rate [28]. The data used for statistical analysis was the cluster center result after applying the FCM algorithm using CIELAB colorspace and 4 cluster numbers as an input.

Results and discussion

Figure 3 shows a sample histogram distribution for the red band of RM 250 agarwood sample. In general, the pixels are normally distributed in a bell shape. It shown that the agarwood physical colour properties have less tendency to produce unusually extreme values and thus not show any specific colour pattern. Since the background of the image is white (255), thus, based on the histogram, a threshold value of 180 provides better representation to differentiate between the agarwood image and its background. This threshold value is then applied to all the agarwood images used. The constant threshold value is important in this analysis in order to eliminate bias in further analysis and making it easy to develop an automated grading system in the future. Therefore, only pixels with a value less than or equal to 180 in all the RGB bands will be used in the analysis. The HSI and CIELAB images were then transformed by considering only this set of pixels.

Twenty samples of each agarwood grades were then classified into two to six numbers of clusters by using FCM classification. The best colorspace was determined based on the maximum frequency of optimal cluster validity of each index used. **Table 2** shows the frequency of occurrences for each cluster validity index in RGB, HSI and CIELAB for the different numbers of clusters.

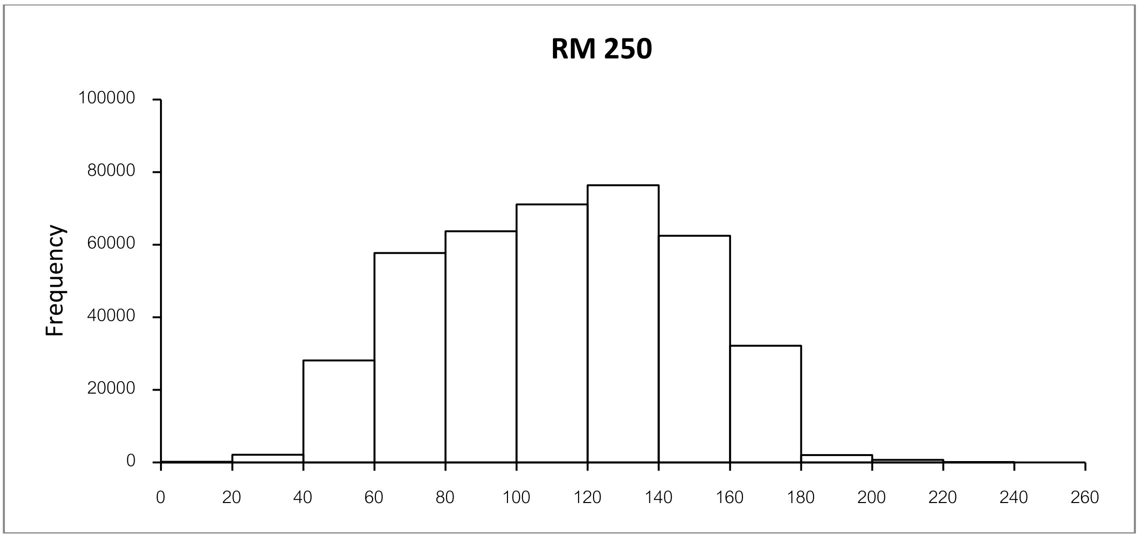
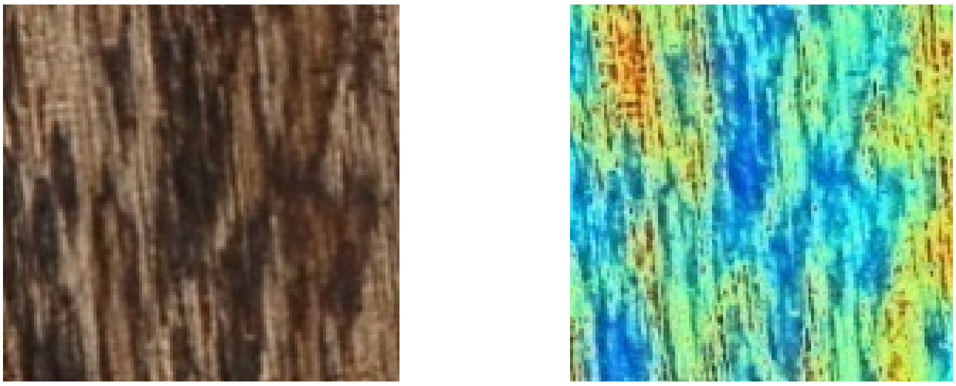


Figure 3 Example of red band histogram taken from RM 250 sample of agarwood.

Table 2 Results of cluster validity indices and its total number of occurrences.

Indices	Colorspace	No. of Clusters, *c*				
		2	3	4	5	6
PC	RGB	95	13	1	-	-
	HSI	45	127	139	140	140
	CIELAB	-	-	-	-	-
CE	RGB	90	10	-	-	-
	HSI	50	130	140	140	140
	CIELAB	-	-	-	-	-
S	RGB	-	-	-	-	-
	HSI	119	29	-	-	-
	CIELAB	21	111	140	140	140
XB	RGB	3	12	10	2	4
	HSI	22	24	22	19	20
	CIELAB	115	104	108	117	116
DI	RGB	1	2	3	8	15
	HSI	41	38	25	22	14
	CIELAB	98	101	111	103	113

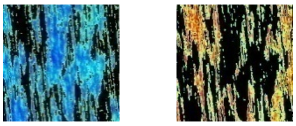
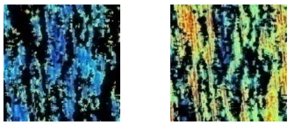
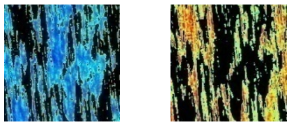
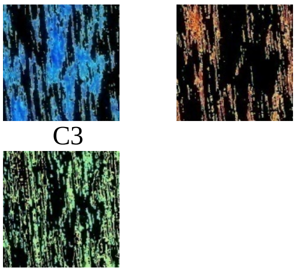
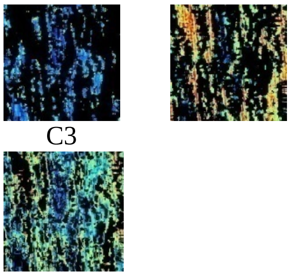
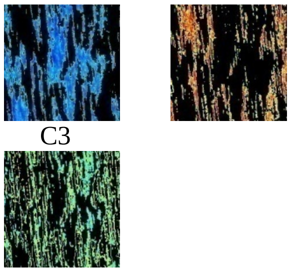
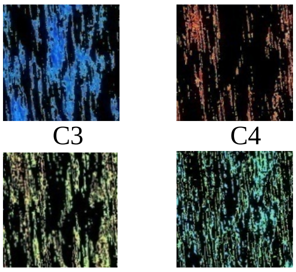
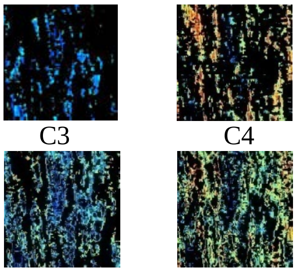
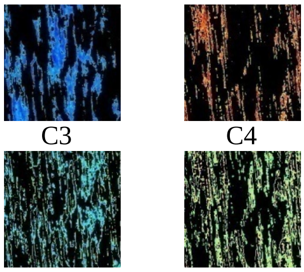
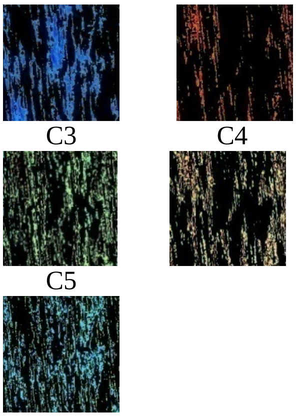
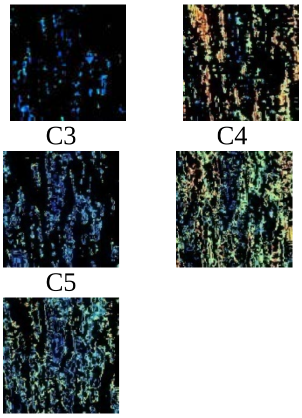
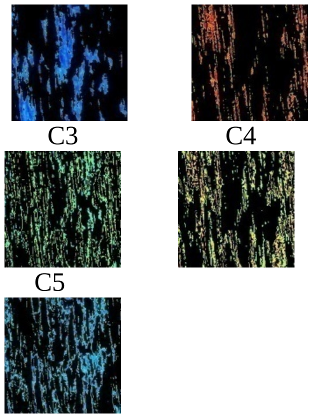
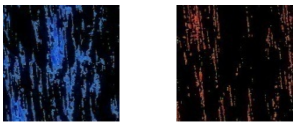
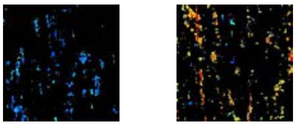
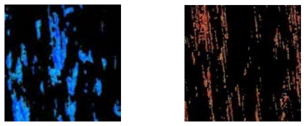
Results from **Table 2** show that, for the PC index, RGB colorspace give the best performance at 2 clusters (95 total numbers of occurrences) followed by HSI colourspace (45 total numbers of occurrences). Meanwhile, as the number of clusters was increased from 3 to 6, the PC index showed optimal occurrences at HSI colorspace. For the CE index, the same pattern as the result of PC index is found. Conversely for the S index, only the HSI colorspace is best at 2 clusters but when the number of clusters increases from 3 to 6, CIELAB produces the highest number of optimal occurrences. From the results of the PC and CE indexes, it is clear that neither of these indices give an optimal result for the CIELAB colorspace, this is due to the agarwood colour characteristics (low colour contrast) which results in colour transformation that does not give an advantage to the CIELAB colorspace since it produces a low correlation image (independent) compared to RGB colorspace (high correlation image). The weaknesses of PC and CE indexes which only consider the membership degree make it biased to the RGB colorspace. The disadvantages of PC and CE indices which tend to monotonically decrease with an increasing number of clusters which also affect the final decision. For the XB and DI indices, it shows that the CIELAB colorspace produces the highest number of occurrences for all numbers of clusters tested. The high number of occurrences of CIELAB colorspace in S, XB and DI indices indicates strongly that colour classification using CIELAB colorspace is more accurate.

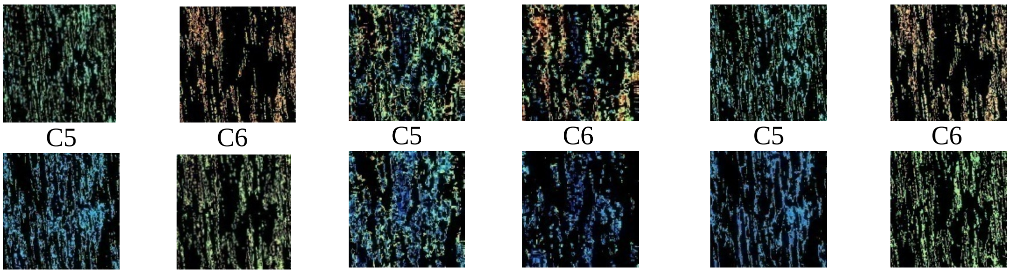


(a) (b)

Figure 4 (a) Agarwood image (cropped), (b) Agarwood image after applying colour indicing using colourmap function (jet, 255).

Table 3 Resulting agarwood images after the FCM segmentation process.

No. of Clusters	RGB Colorspace	HSI Colorspace	CIELAB Colorspace
c = 2	C1 C2 	C1 C2 	C1 C2 
c = 3	C1 C2 C3 	C1 C2 C3 	C1 C2 C3 
c = 4	C1 C2 C3 C4 	C1 C2 C3 C4 	C1 C2 C3 C4 
c = 5	C1 C2 C3 C4 C5 	C1 C2 C3 C4 C5 	C1 C2 C3 C4 C5 
c = 6	C1 C2  C3 C4	C1 C2  C3 C4	C1 C2  C3 C4



From the observation and comparison of **Table 3** and **Figure 4**, RGB and CIELAB colorspace provide high consistency and accuracy in classifying the agarwood image while the HSI colorspace performed poorest for each number of clusters used. The inconsistency of the HSI colorspace was due to the discontinuity of colour pixels transformation especially for pixels value near to the HSI cylinder axis. To the naked eye, RGB and CIELAB do not perform significantly different in terms of classification consistency and accuracy.

Table 4 Total number of occurrences for the indices used in the CIELAB colorspace.

No. of Cluster	2	3	4	5	6
Total occurrences	213	326	359	360	369

To determine the best number of clusters for the FCM algorithm, the result of every index in the CIELAB colorspace need to be considered. The rule in the selection of number of clusters for FCM processing is that the number of clusters should not be too high but must give adequate consistency and accuracy. Too high a number will be reflected in processing time and will not give significantly better result in the final output. From observation of **Tables 3** and **4**, four clusters are acceptable to be used based on its consistency and accuracy compared to the other clusters. Therefore, it will be used for further analysis in this study.

Table 5 shows the result of ANOVA of CIELAB colour component and the number of clusters. The observed significant values for all the colour components and number of clusters are less than 0.05 ($p < 0.05$). Therefore, there is a significant effect among the colour cluster for each band on the agarwood price. However, it does not isolate which number of clusters and colour component differences are.

Table 5 Result of ANOVA for agarwood data.

		Sum of squares	df	Mean square	F	Sig.
L1	Between Groups	322.481	6	53.747	4.116	0.001
	Within Groups	1736.702	133	13.058		
	Total	2059.183	139			
a1	Between Groups	1225.004	6	204.167	85.571	0
	Within Groups	317.33	133	2.386		
	Total	1542.334	139			
b1	Between Groups	280.35	6	46.725	4.028	0.001
	Within Groups	1542.649	133	11.599		
	Total	1822.999	139			
L2	Between Groups	975.492	6	162.582	10.399	0
	Within Groups	2079.461	133	15.635		
	Total	3054.953	139			
a2	Between Groups	1947.865	6	324.644	188.155	0
	Within Groups	229.479	133	1.725		
	Total	2177.344	139			
b2	Between Groups	615.322	6	102.554	12.376	0
	Within Groups	1102.063	133	8.286		
	Total	1717.385	139			
L3	Between Groups	1999.515	6	333.253	14.399	0
	Within Groups	3078.16	133	23.144		
	Total	5077.675	139			
a3	Between Groups	2294.94	6	382.49	202.218	0
	Within Groups	251.566	133	1.891		
	Total	2546.506	139			
b3	Between Groups	740.088	6	123.348	23.62	0
	Within Groups	694.557	133	5.222		
	Total	1434.645	139			
L4	Between Groups	2283.907	6	380.651	10.64	0
	Within Groups	4758.321	133	35.777		
	Total	7042.228	139			
a4	Between Groups	2146.344	6	357.724	132.731	0
	Within Groups	358.449	133	2.695		
	Total	2504.793	139			
b4	Between Groups	1021.801	6	170.3	13.879	0
	Within Groups	1631.933	133	12.27		
	Total	2653.734	139			

Table 6 Results of DMRT for agarwood data.

CIELAB Colour Components	No. of cluster, c	Agarwood Grade						
		RM 250	RM 350	RM 800	RM 900	RM 1000	RM 2500	RM 3500
L	1	25.564 *a*	21.788 *b*	20.633 *b*	22.739 *b*	21.482 *b*	21.432 *b*	21.286 *b*
	2	34.898 *a*	30.144 *b*	26.575 *d*	29.550 *bc*	27.456 *bcd*	27.373 *d*	27.751 *bcd*
	3	45.772 *a*	39.553 *b*	33.425 *d*	37.238 *bc*	35.472 *cd*	35.318 *cd*	36.063 *bcd*
	4	56.863 *a*	51.288 *b*	**42.991 *d***	47.042 *c*	48.905 *bc*	46.530 *cd*	48.047 *bc*
a	1	8.337 *a*	0.633 *c*	6.893 *b*	5.973 *b*	6.746 *b*	0.456 *c*	6.698 *b*
	2	9.861 *a*	1.163 *d*	9.291 *ab*	7.657 *c*	9.099 *ab*	0.485 *d*	8.866 *b*
	3	8.800 *a*	0.281 *c*	9.481 *a*	7.674 *b*	9.610 *a*	-0.006 *c*	9.430 *a*
	4	7.545 *a*	-1.091 *b*	7.797 *a*	6.800 *a*	7.561 *a*	-1.217 *b*	7.882 *a*
b	1	13.248 *a*	13.449 *a*	10.854 *bc*	11.222 *abc*	9.931 *bc*	11.934 *ab*	9.482 *c*
	2	18.360 *b*	20.686 *a*	15.878 *cde*	16.099 *cd*	14.707 *de*	17.286 *bc*	14.083 *e*
	3	19.822 *bc*	24.544 *a*	18.869 *cd*	18.466 *cd*	17.878 *d*	**20.988 *b***	17.248 *d*
	4	19.832 *b*	24.818 *a*	18.302 *bc*	19.418 *bc*	17.314 *c*	23.194 *a*	17.664 *bc*

Follow-up tests were conducted to evaluate pair-wise differences among the means (colour component, number of cluster and agarwood price) by applying Duncan's Multiple Range Test (DMRT). The result of DMRT is shown in **Table 6**. Different letters in the same row indicates which comparisons were significant at 95 % confidence intervals. Based on **Table 6**, several agarwood price have clearly produced specific colour features (shaded) which is RM250 [L1: 25.564±3.50, L2: 34.898±2.89, L3: 45.772±2.97, L4: 56.863±3.08, a1: 8.337±1.69], RM350 [b2: 20.686±3.46, b3: 24.544±2.89] and RM900 [a2: 7.657±1.40, a3: 7.674±1.22]. Meanwhile, RM800 [L4: 42.991±3.78] and RM2500 [b3: 20.988±2.71] shows a colour similarity with RM2500 and RM250 respectively but differ greatly in price (bold). This group has a potential to be classified if related price is removed through their specific colour features. For example, RM2500 can be classified after removing RM250 using L1 and followed by b3 colour features. However, RM1000 and RM3500 do not show any significant relationships between CIELAB colour component, number of clusters and price. As a summary, the low price of agarwood gave significant relationships with their physical colour properties. The method might be improved by considering other properties used during a grading process such as density, resin content and fragrance.

Conclusions

Physical colour properties of agarwood image on three different colorspaces have been studied. Their capabilities for FCM classification has been evaluated using five types of cluster validity indices. The results show that the CIELAB colorspace produces the highest frequency of optimal occurrences for most of the indices used. Meanwhile, four clusters produced a high classification in consistency and accuracy in clustering of the agarwood image. The statistical analysis has shown that there are relationships between agarwood price and its physical colour properties especially for RM250 for the *L* colour component. Meanwhile, the *b* colour component is acceptable for RM350 and the *a* colour component is acceptable for RM900 agarwood price. All of these prices can be classified using these specific colour components and its cluster number. Based on this research, it can be concluded that image processing has a potential to be used in the agarwood grading task. Thus, it will help in the future development of an automatic agarwood grading system especially for the lower priced agarwood types.

Acknowledgements

The authors wish to thank and are grateful for the samples and expertise given by Malaysian Institute for Nuclear Technology Research (MINT).

References

[1] LSL Chua. *Agarwood (Aquilaria Malaccensis) in Malaysia*, Forest Research Institute Malaysia, 2008.

[2] TE Asia-Taipei and TS Asia. *The Trade and Use of Agarwood in Taiwan, Province of China*. CITES; 2005.

[3] T Soeharto and AC Newton. Conservation and sustainable use of tropical trees in the genus Aquilaria II. The impact of gaharu harvesting in Indonesia. *Biol. Conservat.* 2001; **97**, 29-41.

[4] T Soeharto and AC Newton. The gaharu trade in Indonesia: Is it sustainable? *Econ. Bot.* 2002; **56**, 271-84.

[5] F Zich and J Compton. The Final Frontier: Towards Sustainable Management of Papua New Guinea's Agarwood Resource. *In*: TRAFFIC Oceania and the WWF South Pacific Programme, 2001.

[6] A Barden, NA Anak, T Mulliken and M Song. Heart of the Matter: Agarwood Use and Trade and CITES Implementation for Aquilaria Malaccensis. *TRAFFIC International Publication*, 2000.

[7] K Chakrabarty, A Kumar and V Menon. Trade in Agarwood. In: TRAFFIC India, New Delhi, 1992.

[8] W Kurdthongmee. Colour classification of rubberwood boards for fingerjoint manufacturing using a SOM neural network and image processing. *Comput. Electron. Agr.* 2008; **64**, 85-7.

[9] A Abdullah, NKN Ismail, TAA Kadir, JM Zain, NA Jusoh and NM Ali. Agar wood grade determination system using image processing technique. *In*: Proceedings of Electrical Engineering and Informatics. Institut Teknologi Bandung, Indonesia. 2007, p. 427-9.

[10] DG Donovan and RK Puri. Learning from traditional knowledge of non-timber forest products: Penan benalui and the autecology of *Aquilaria* in Indonesian Borneo. *Ecol. Soc.* 2004; **9**, Art 3.

[11] NR Pal and SK Pal. A review on image segmentation techniques. *Pattern Recognit.* 1993; **26**, 1277-94.

[12] Y Chahir and A Elmoataz. Skin-color detection using fuzzy clustering. *In*: Proceedings of International Symposium on Control, Communications, and Signal Processing. Marrakech, Morocco, 2006.

[13] CFJ Kuo, CY Shih, CY Kao and JY Lee. Color and pattern analysis of printed fabric by an unsupervised clustering method. *Textil. Res. J.* 2005; **75**, 9-12.

[14] Z Ronghua, C Hongwu, Z Xiaoting, P Ruru and L Jihong. Unsupervised color classification for yarn-dyed fabric based on FCM algorithm. *In*: Proceedings of International Conference on Artificial Intelligence and Computational Intelligence, Sanya, China. 2010, p. 497-501.

[15] JY Kang, LQ Min, QX Luan, X Li and JZ Liu. Dental plaque quantification using FCM-based classification in HSI color space. *In*: Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China. 2008, p. 78-81.

[16] T Saikumar, P Yugander, P Sreenivasa and B Smitha. Colour based image segmentation using fuzzy c-means clustering. *In*: Proceedings of International Conference on Computer and Software Modeling, Singapore. 2011, p. 180-5.

[17] G Padmavathi, Mr Muthukumar. Image segmentation using fuzzy c means clustering method with thresholding for underwater images. *Int. J. Adv. Network. Appl.* 2010; **2**, 514-8.

[18] S Chuai-Aree, C Lursinsap, P Sophatsathit and S Siripant. Fuzzy C-means: a statistical feature classification of text and image segmentation method. *Int. J. Uncertainty, Fuzziness and Knowledge* 2001; **9**, 661-71.

[19] VEC Ghaleh and A Behrad. Lip contour extraction using RGB color space and fuzzy c-means clustering. *In*: Proceedings of the 9th International Conference on Cybernetic Intelligent Systems, Reading, UK. 2010.

[20] A Sopharak, B Uyyanonvara and S Barman. Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy C-means clustering. *Sensors* 2009; **9**, 2148-61.

[21] NR Pal and JC Bezdek. On cluster validity for the fuzzy c-means model. *IEEE Trans. Fuzzy Syst.* 1995; **3**, 370-9.
[22] KL Wu, MS Yang and JN Hsieh. Robust cluster validity indexes. *Pattern Recognit.* 2009; **42**, 2541-9.
[23] Y Zhang, W Wang, X Zhang and Y Li. A cluster validity index for fuzzy clustering. *Inform. Sci.* 2008; **178**, 1205-13.
[24] CY Yen and KJ Cios. Image recognition system based on novel measures of image similarity and cluster validity. *Neurocomputing* 2008; **72**, 401-11.
[25] W Wang and Y Zhang. On fuzzy cluster validity indices. *Fuzzy Set. Syst.* 2007; **158**, 2095-117.
[26] HD Cheng, XH Jiang, Y Sun and J Wang. Color image segmentation: advances and prospects. *Pattern Recognit.* 2001; **34**, 2259-81.
[27] RC Gonzalez and RE Woods. *Digital Image Processing*. 2nd ed. Prentice Hall, New Jersey, 2002, p. 190.
[28] KL Wu and MS Yang. A cluster validity index for fuzzy clustering. *Pattern Recognit. Lett.* 2005; **26**, 1275-91.
[29] CH Chou, MC Su and E Lai. A new cluster validity measure and its application to image compression. *Pattern Anal. Appl.* 2004; **7**, 205-15.
[30] S Thilagamani and N Shanthi. Literature survey on enhancing cluster quality. *Int. J. Comput. Sci. Eng.* 2010; **2**, 1999-2002.
[31] KA Gomez and AA Gomez. *Statistical Procedures for Agricultural Research*. 2nd ed. John Wiley & Sons, New York, 1984, p. 680.