

Statistical feature extraction method for wood species recognition system

Imanurfatihah Ibrahim¹ · Anis Salwa Mohd Khairuddin¹ · Hamzah Arof¹ · Rubiyah Yusof² · Effariza Hanafi¹

Received: 4 August 2015 / Published online: 22 March 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract A cascaded wood species recognition system using simple statistical properties of the wood texture is presented where a total of 24 statistical features are extracted from each wood sample. They are mainly vessel features that allow a broad initial grouping of wood texture using fuzzy logic. Then, a neural network classifier is used to refine the broad grouping into the final wood species classification. The proposed system emulates the classification approach normally taken by human experts when analyzing wood species based on texture. A comprehensive set of experiments was performed on a database composed of 3000 macroscopic images of 30 different wood species to evaluate the effectiveness of the system. Finally, its performance is compared with previous works in terms of classification accuracy.

1 Introduction

Illegal logging has been identified as a major cause of deforestation and loss of valuable biodiversity, ecosystem, resources and other assets associated with the tropical rainforest. Despite tighter conservation regulations, demand in wood products has continued to increase due to growing population. Therefore, intelligent systems for the recognition of wood species have been proposed and developed as

one of the measures to curb illegal logging. The systems should be able to identify the timber species accurately and approximate the age of the tree for a particular log.

Normally, experts identify the wood species based on the pattern of the wood surface texture (Menon et al. 1993). Experts who examine wood surface texture scrutinize several important characteristics of the wood such as the arrangement of vessel or pores, wood parenchyma or soft tissue, rays parenchyma, fibers, phloem, latex traces and intercellular canals. For instance, different wood species will have vessels of various size, arrangement, quantity and density on their surfaces.

The IAWA lists define the internationally standardized features for wood identification (Ruffinatto et al. 2015). This process of manual inspection is tedious and time consuming. Hence, several intelligent systems for the recognition of wood species have been developed based on two approaches: spectrum based and image based systems.

Spectral analysis is one of the most useful techniques in modern science. The method is based on the measurement of internal properties of products obtained by using nondestructive techniques in transmission mode such as γ -rays, X-rays, microwaves or ultrasound. This approach allows thorough inspection of the wood characteristics since internal features of wood are embedded within. The macroscopic wood features that are analyzed using spectrum analysis method provide a lot of information that can be used to identify the wood species efficiently (Fuentalba et al. 2004; Choffel 1999; Piuri and Scotti 2010; Bass and; Wheeler 2000; Rojas et al. 2011). Despite the efficiency of spectrum-based approaches, they are normally time consuming and require special equipment, settings and well trained experts who can interpret the data. Additionally, these techniques are more suitable to be carried out in laboratories rather than on-site in real applications.

✉ Anis Salwa Mohd Khairuddin
anissalwa@um.edu.my

¹ Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia

² Malaysia Japan Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

Image based wood species classification is simpler because wood surface cross section can be examined by human naked eye with the aid of a magnifying glass. The wood texture is examined from the timber surface using a magnifier with 10× magnification. Several automatic tropical wood recognition systems have been developed based on image analysis to furnish quantifiable, repeatable and reliable results. Examples are in the works of Hermanson and Wiedenhoef [2011](#); Tou et al. [2007](#), [2008](#); Khalid et al. [2008](#); Bremanath et al. [2009](#) and Filho et al. [2010](#). Approaches based on image analysis normally extract distinctive textural features from the grey level images of the wood surfaces to identify and discriminate the tree species (Wheeler [2011](#)). Myriad of textural features can be used to classify wood textures and they include statistical, structural, transform based, model based, multi resolution and numerous other ad-hoc features.

Statistical features are inherent to the distribution of grey levels in images and they are widely used to identify textural patterns (Peng et al. [2015](#)). Generally, different wood species have different statistical features that determine the wood types. Image based wood recognition systems are usually designed to emulate visual inspection of wood textures by human experts. Fuzzy logic can be incorporated into the system to provide linguistic interpretation normally used by experts when handling approximate data for vaguely defined problems (Anninou and Groumpos [2014](#); Batuwita et al. [2010](#)). For wood texture recognition involving many classes, fuzzy model is especially useful in initial data clustering (grouping) before the final classification of the wood species is made.

In this paper, a cascaded wood species classification method is proposed based on fuzzy logic and statistical features that imitate the way human experts perform inference when analyzing wood texture. First, vessel features are used

to form a broad initial grouping of the wood types. Then, a neural network classifier is used with the statistical features to further classify the wood types to arrive at the final classification of the wood species.

2 Materials and methods

2.1 Image acquisition of the wood surface texture

The first step in the proposed wood species recognition system is the image acquisition of the wood surface texture. One of the characteristics that remain unique to each wood species even after undergoing the chemical procedures is the surface texture. Different wood species will have variations in size, arrangement, quantity and density of vessels on the wood surface. These features are very important and can be used to identify the wood species. One of the important features used to identify a wood species is the vessel density. Certain wood species might have numerous vessels while some wood species can have very few vessels on the wood surface. Vessel density is determined based on the quantity of vessels per square millimeters on the wood surface. Figure 1 presents two different samples from two different wood species.

The samples are wood cubes whose side is approximately 1 inch in dimension. The treatment of the wood samples was done by sanding the wood surfaces. A specially designed portable camera is used to capture wood texture images at 10 times magnification. The camera is equipped with a systematic focusing function whereby the distance between the camera and the wood sample is fixed to 10 cm. The housing of the camera is made of a tube based on theoretical optimal object distance and hence, the object is just required to be laid against the camera housing.

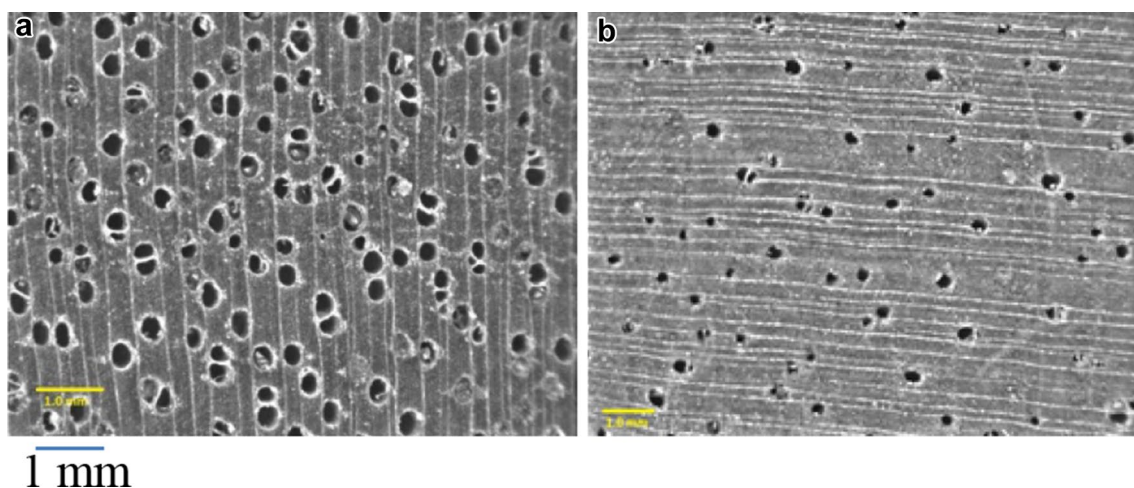


Fig. 1 Example of wood surface with **a** high density and **b** low density

The “tube function” is made in opaque material in order to cut all ambient light and its fluctuations. The size of each image is 768×576 pixels.

The original wood image is enhanced using homomorphic filtering technique to enhance the image presentation. Homomorphic filtering technique reduces all unwanted illumination and reflectance on the image (Woods and Gonzalez 2008). Homomorphic filtering uses a linear filter to do non-linear mapping to a different domain and later it was mapped back to the original domain. The algorithm reduces all unwanted illumination and reflectance on the image. Image brightness was also normalized, thus enhanced the contrast of the image. The algorithm on homomorphic filtering is explained in more detail in Woods and Gonzalez (2008). The summary of homomorphic filtering process is shown in Fig. 2.

Figure 3 shows the original image and filtered image after undergoing the homomorphic filtering process.

2.2 Proposed feature extraction

Vessels appearances are considered as one of the significant features that can be used by experts to identify the wood species. This is because vessels are exclusive for every species. Therefore, this research will focus on size of vessels, quantity of vessels, types of vessels and arrangement of

$$f(x,y) \rightarrow \ln \rightarrow DFT \rightarrow H(u,v) \rightarrow IDFT \rightarrow \exp \rightarrow g(x,y)$$

Fig. 2 The flowchart of homomorphic filtering process where *ln* natural log, *DFT* discrete fourier transform, *IDFT* inverse discrete Fourier Transform and *exp* exponential

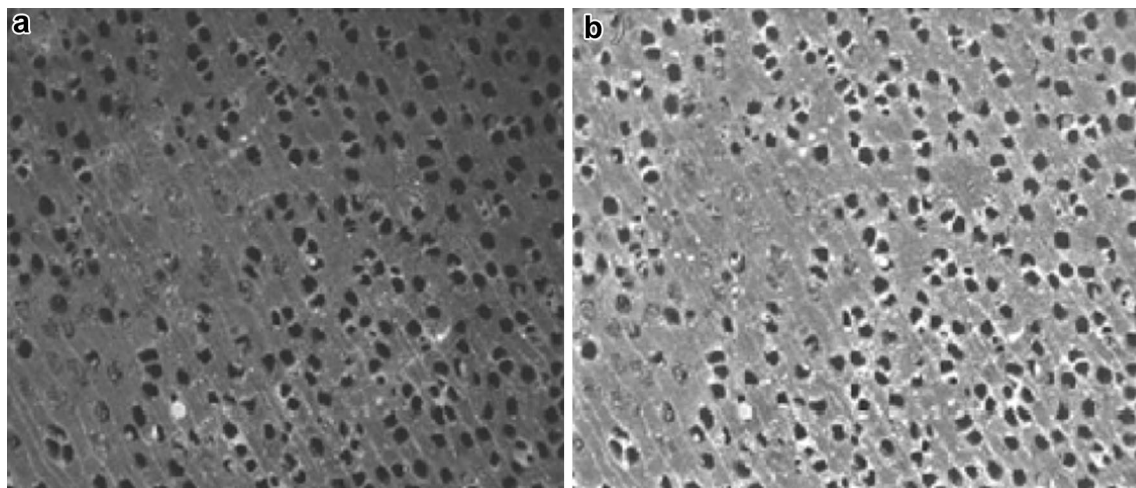


Fig. 3 a Original image of wood sample, b homomorphic image of wood sample

vessels on the wood texture. There are 24 features extracted from each wood image by using the proposed statistical feature extractor.

The statistical feature extraction process consists of two steps namely vessels extraction and fuzzy vessels management (Fig. 4). The statistical features will only allow distinct pores to be acknowledged as characteristics of a wood species. Then, a neural network classifier is used to classify the wood species based on the statistical wood features.

The first step in the proposed statistical feature extraction is the vessels extraction process. Basically, binary images are created from the homomorphic filtered image where only black pores are present and only white pores are present as shown in Fig. 5.

The binary images are created by using the default binarization based on Otsu’s method for optimal threshold selection. A threshold filter displays each pixel of an image in only one of two states, black or white. That state is set according to a particular threshold value. If the pixel’s brightness is greater than the threshold, the pixel is colored white, if less than, black (Ramli et al. 2015). This thresholding technique stores the intensities of the pixels in an array. The threshold is calculated by using total mean and variance. Based on this

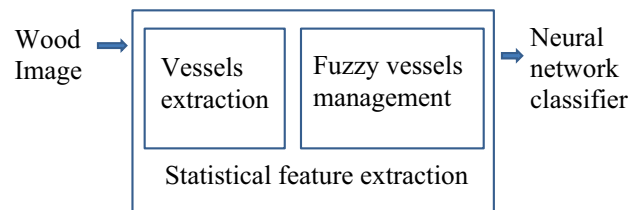


Fig. 4 Flowchart of the proposed wood recognition system

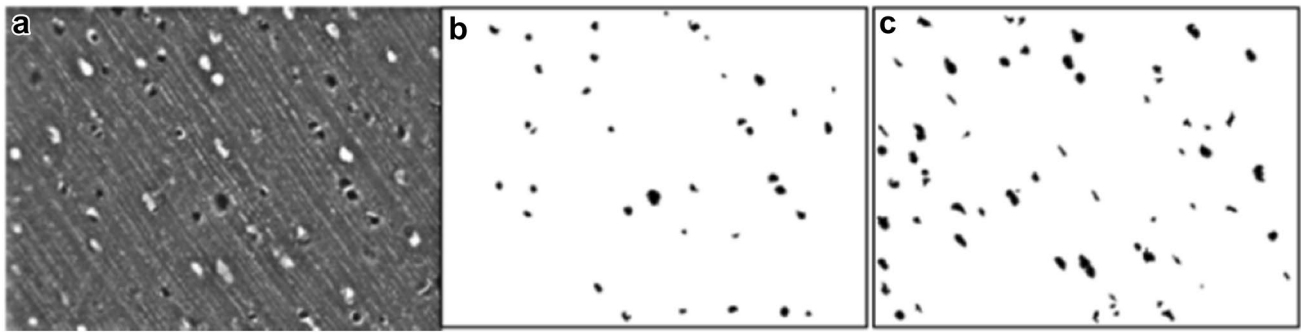


Fig. 5 a Homomorphic image, b binary images showing the black vessels only and c binary images showing the white vessels only

threshold value, each pixel is set to either 0 or 1. Therefore, the change of image takes place only once.

The following formulas are used to calculate the total mean and variance. The pixels are divided into 2 classes, C_1 with gray levels $[1, \dots, t]$ and C_2 with gray levels $[t+1, \dots, L]$. The probability distribution for the two classes is:

$$C_1 = \frac{p_1}{w_1}(t), \dots, \frac{p_t}{w_1}(t) \text{ and} \quad (1)$$

$$C_2 = \frac{p_{t+1}}{w_2}(t), \dots, \frac{p_L}{w_2}(t) \quad (2)$$

where $w_1(t) = \sum_{i=1}^t p_i$ and $w_2(t) = \sum_{i=t+1}^L p_i$. The means for the two classes are

$$\mu_1 = \frac{\sum_{i=1}^t ip_i}{w_1(t)} \text{ and } \mu_2 = \frac{\sum_{i=t+1}^L ip_i}{w_2(t)} \quad (3)$$

Using discriminant analysis, Otsu defined the between-class variance of the threshold image as

$$\sigma_B^2 = w_1(\mu_1 - \mu_t)^2 + w_2(\mu_2 - \mu_t)^2 \quad (4)$$

For bi-level thresholding, Otsu verified that the optimal threshold t^* is chosen so that the between-class variance is maximized; that is $t^* = \text{Arg Max}_{1 < t < L} \{\sigma_B^2(t)\}$.

After thresholding, the area of a detected region is approximated by the number of pixels it contains. Then the region is labelled automatically before its statistical parameters can be computed. Labelling is necessary to avoid a detected one from being analyzed twice.

Let $\hat{\sigma}$ represent the estimated standard deviation of all the void area, $\hat{\mu}$ represent the estimated mean of all the void area, and S_i represent the area of the i th void. The i th void is defined as a pore if:

$$\frac{s_i - \hat{\mu}}{\hat{\sigma}} > \theta, \quad (5)$$

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n S_i \quad (6)$$

$$\hat{\sigma} = \frac{1}{n} \sum_{i=1}^n (S_i - \hat{\mu})^2 \quad (7)$$

where θ is a weight used to adjust the threshold of the void area.

In order to establish the cardinal directions within each image, the special properties of Gabor filters were used to estimate the best directionality distribution of wood vessels. A 2-dimension (2D) Gabor filter is capable of detecting boundaries within a 2D image $I(x,y)$ which can be formed by convolving the Gabor function with the image on a pixel basis, so that the filter output at an orientation θ can be represented by (Gdyczynski and Manbachi 2014):

$$V(x,y;\theta) = G(x,y;\theta) * I(x,y) \quad (8)$$

where $*$ represents a 2D convolution. When such a function is convolved at various orientations to an image function, the highest Gabor coefficient output is obtained when the edges are perpendicular to the angle of the filter. Hence, by scanning the wood image over a range of filter angles (0° , 45° , 90° , 135° , 180°), the angles at which the edges occur are determined.

Then features related to the statistical properties of vessels are extracted from both binary images for each wood image. The second step in the proposed statistical feature extraction is the fuzzy management process. The fuzzy algorithm is explained in Kuncheva (2000). Let $\Omega = \{\omega_1, \dots, \omega_c\}$ be a set of class labels. Let $x = [x_1, \dots, x_n]^T \in \mathbb{R}^n$ be a vector describing an object such as an image. Each component x expresses the value of a feature such as size, quantity, length, etc. The fuzzy management system is a mapping of $D: \mathbb{R}^n \rightarrow \Omega$. The canonical model of the fuzzy system shall be considered as a black box at the input where x is submitted and the output produces values of c discriminant functions $g_1(x), \dots, g_c(x)$, expressing the support for the respective classes. The maximum membership rule assigns x to the class with the highest support. Fuzzy if-else rules are employed to categorize the vessels from both binary

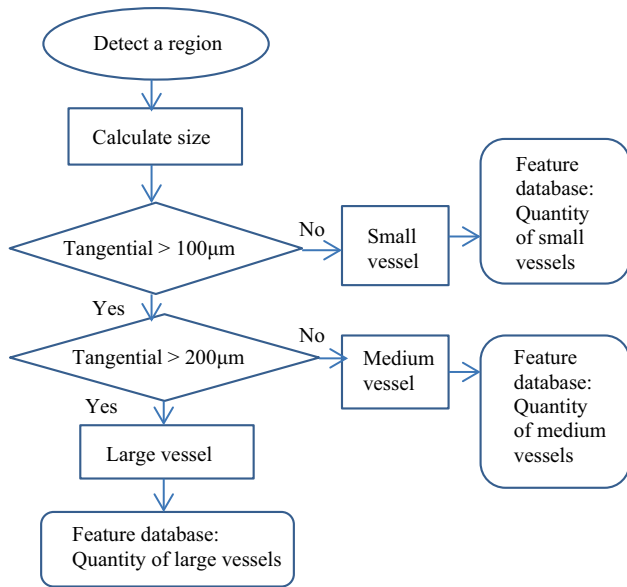


Fig. 6 Flowchart of the statistical feature extraction process for vessel sizes

images into several features such as sizes of vessels, types of vessels and vessels arrangements.

Suppose a set of input images from a wood database with n -feature variables and m image samples. The input data (wood sample features) x_p on the pattern space is represented by the following pattern matrix: $x_p = [x_{p1}, x_{p2}, \dots, x_{pn}]$, $p = 1, 2, \dots, m$. These training patterns are classified into M ($M \leq m$) classes. In this paper, the wood data has 24 feature variables with 2100 training wood data and 30 classes, that is, $n = 24$, $m = 2100$, and $M = 30$. The 24 feature variables consist of statistical features such as sizes of vessels, types of vessels and vessel arrangements which are extracted from both binary images (black vessels only and white vessels only).

2.2.1 Sizes of vessels

The vessel size is also used to distinctively define the wood species. The flowchart of the statistical feature extraction for vessels sizes is presented in Fig. 6. The measurement

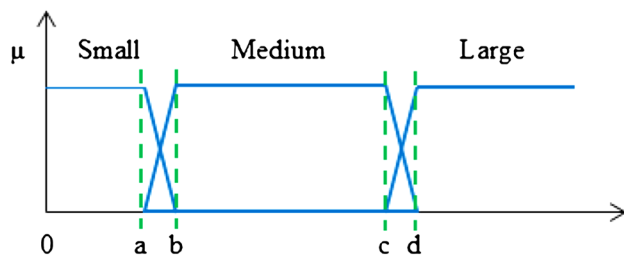


Fig. 7 Membership function to determine the size of pores

of vessel size is defined by Menon et al. (1993). The size of vessels is measured based on the tangential of vessels.

The grouping model is obtained using fuzzy IF-THEN rules can be written as below:

IF x_{p1} is $A_{j1}(x_{p1})$ THEN x_{p1} belongs to group M .

Where all variables are defined as follows:

x_{p1} input tangential of vessel of a wood sample.

$A_{j1}(x_{p1})$ antecedent linguistic membership value.

Group M consequent group.

The vessels are divided into 3 categories by using fuzzy rules: small vessels, medium vessels and large vessels as below.

- IF the tangential of vessel is below 100 μm , THEN the vessel is considered as small vessel.
- IF the tangential of vessel is between 100 and 200 μm , THEN the vessel is considered as medium vessel.
- IF the tangential of vessel is above 200 μm , THEN the vessel is considered as large vessel.

The membership function (MF) deals with mapping an input space to a membership value between 0 and 1. The membership function $\mu_A(x)$ describes the membership of the elements x of the base set X in the fuzzy set A , whereby

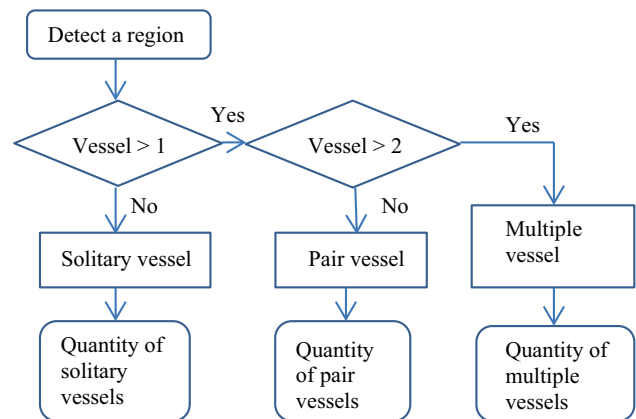


Fig. 8 Flowchart of the statistical feature extraction process for types of vessels

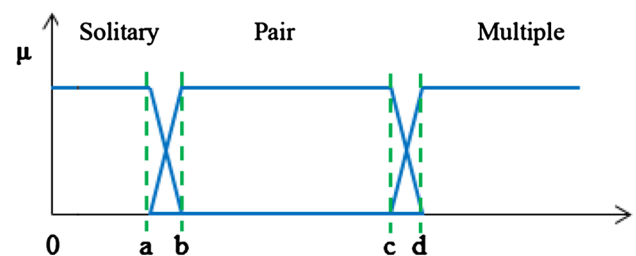
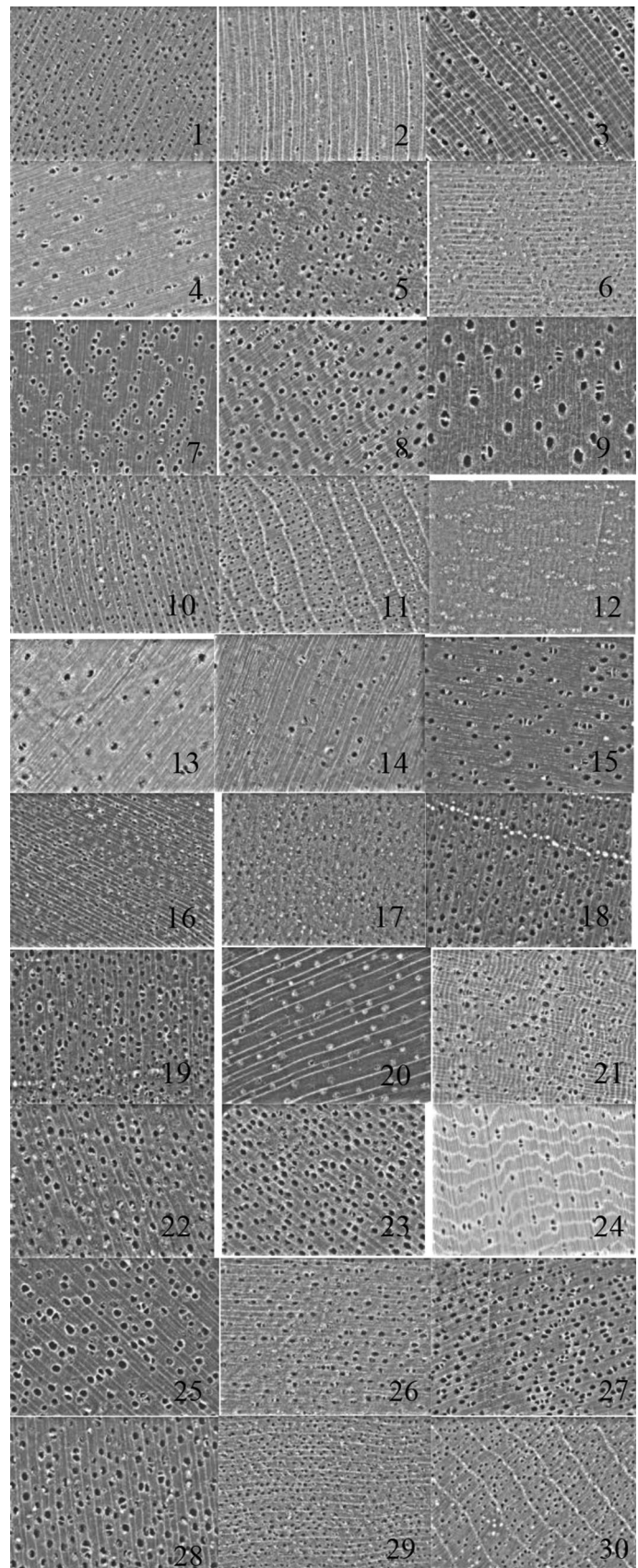


Fig. 9 Membership function to determine the types of vessels

Fig. 10 Wood texture for 30 different wood species. Species numbering refers to wood species presented in Table 1



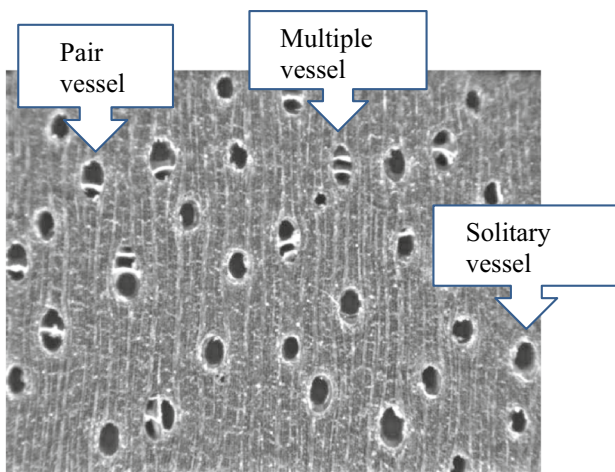


Fig. 11 Types of vessels on wood surface texture

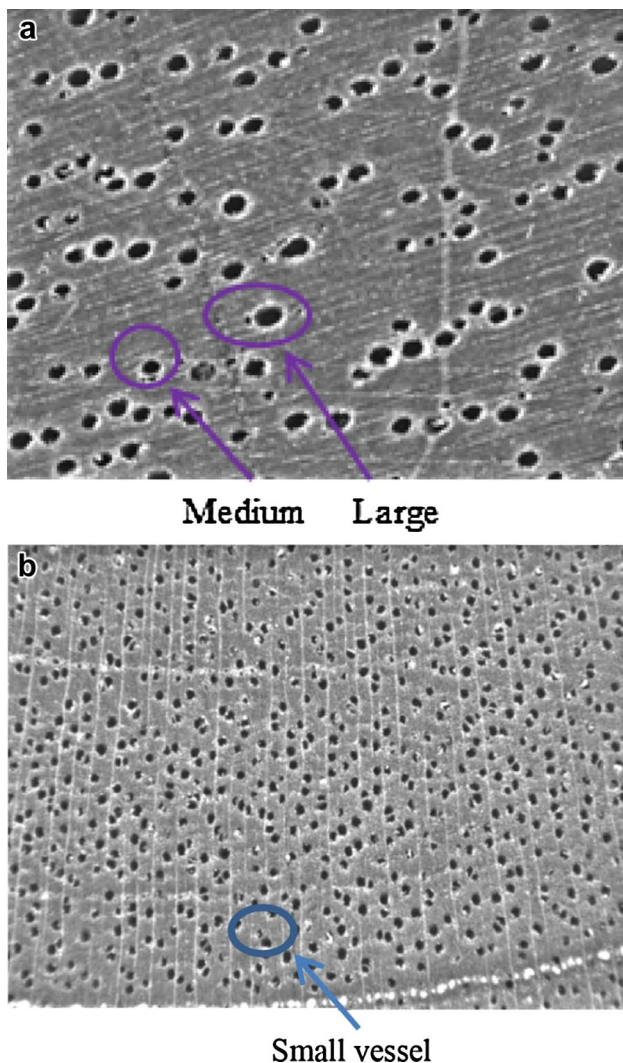


Fig. 12 Different sizes of vessels on wood surface texture

for $\mu_A(x)$ a large class of functions can be taken. Trapezoidal membership function is applied as shown in Fig. 7. For example: the linguistic label “medium” is associated with a possibility distribution defined by four parameters (a, b, c, d).

The parameters of a membership function as shown in Fig. 7 are defined as $a=95, b=105, c=195$ and $d=205$, respectively. The parameters are determined based on the measurement defined by Menon et al. (1993). Each wood image may consist of 3 types of vessels sizes: small vessels, medium vessels and large vessels. Hence, the quantity of vessels for each vessel size is computed for each wood sample. For example, $x_1 = v[p_1 \ p_2 \ p_3]$ where x_1 represents a wood sample, p_1 represents the quantity of small vessels extracted from the wood image, p_2 represents the quantity of medium vessels extracted from the wood image, and p_3 represents the quantity of large vessels extracted from the wood image.

2.2.2 Types of vessels

Vessels may be present in singles, pairs and multiple on the wood surface. Single vessels are called solitary vessels. Pair vessels are present as two vessels combined and become a single vessel. Multiple vessels are present when more than 2 vessels combined to form a single vessel. There are three types of vessels extracted from the wood texture, namely solitary vessel, pair vessel and multiple vessel. The flow-chart of the statistical feature extraction for types of vessels is presented in Fig. 8.

The grouping model is obtained using fuzzy IF-THEN rules can be written as below:

IF x_{p1} is $A_{j1}(x_{p1})$ THEN x_{p1} belongs to group M .

Where all variables are defined as follows:

x_{p1} input number of vessels within a region of a wood sample.

$A_{j1}(x_{p1})$ antecedent linguistic membership value.

Group M consequent group.

The types of vessel are determined based on fuzzy rules as below while the membership function is shown in Fig. 9:

- IF vessel is individual vessel THEN it is a solitary vessel.
- IF vessel is combination of 2 individual vessels THEN vessel is a pair vessel.
- IF vessel is combination of 3 or more individual vessels THEN vessel is a multiple vessel.

Each wood image may consist of 3 types of vessels: solitary vessels, pair vessels and multiple vessels. Hence, the quantity of each type of vessels is computed for each wood sample. For example, $x_1 = [q_1 \ q_2 \ q_3]$ where x_1 represents a

Table 1 Example of 6 statistical features extracted from each wood image for 30 different wood species

Num	Wood species	f1	f2	f3	f4	f5	f6
1	<i>Dehaasia curtisii</i>	109	157	0	13	3	0
2	<i>Dyera costulata</i>	10	35	0	24	8	0
3	<i>Endospermum diadenum</i>	0	49	26	5	14	0
4	<i>Hevea brasiliensis</i>	4	26	21	3	2	2
5	<i>Palaquium stellatum</i>	16	81	15	31	39	4
6	<i>Neobalanocarpas Heimii</i>	50	187	0	56	74	2
7	<i>Calophyllum kunstleri</i>	10	75	32	4	7	0
8	<i>Cynometra malaccensis</i>	5	94	42	4	14	1
9	<i>Durio spp</i>	6	12	40	3	2	0
10	<i>Hopea pachycarpa</i>	194	115	0	48	64	3
11	<i>Kokoona sessilis</i>	278	105	0	21	84	10
12	<i>Madhuca utilis</i>	23	54	0	71	115	2
13	<i>M. caesia</i>	4	33	30	23	25	2
14	<i>M. quadrifida</i>	0	19	14	9	27	0
15	<i>Pometia ridleyi</i>	24	66	30	6	9	3
16	<i>Pentace floribunda</i>	1	64	3	20	11	1
17	<i>Scorodocarpus borneensis</i>	33	48	7	22	29	2
18	<i>Shorea guiso</i>	4	61	12	4	32	1
19	<i>Shorea collina</i>	12	84	36	8	19	1
20	<i>Artocarpus dadah</i>	2	11	36	37	38	0
21	<i>Dialum kingie</i>	8	40	70	3	16	1
22	<i>Dipterocarpus baudii</i>	13	153	10	67	113	2
23	<i>Dryobalanops oblongifolia</i>	34	198	44	11	22	0
24	<i>Lophopetalum sp</i>	41	59	1	38	39	1
25	<i>Parashorea stellate</i>	13	137	80	22	57	2
26	<i>Pentace curtisii</i>	3	62	69	10	33	1
27	<i>Shorea laevis</i>	16	286	12	8	22	0
28	<i>Shorea uliginosa</i>	24	151	11	9	31	1
29	<i>Sp of lauraceaceae</i>	216	106	4	31	37	0
30	<i>Kokoona reflexa</i>	256	73	3	61	85	0

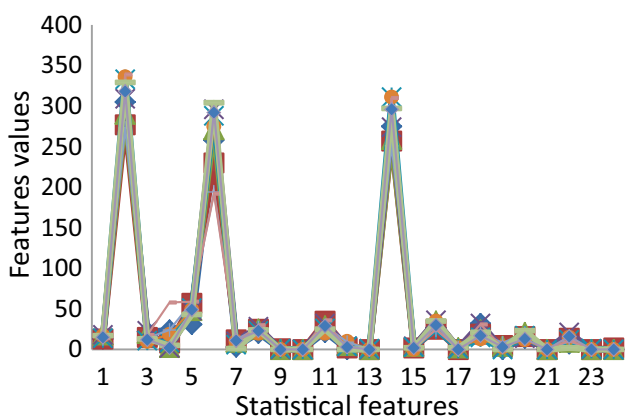


Fig. 13 Graph of 24 statistical features for 10 samples of wood species *Shorea laevis*

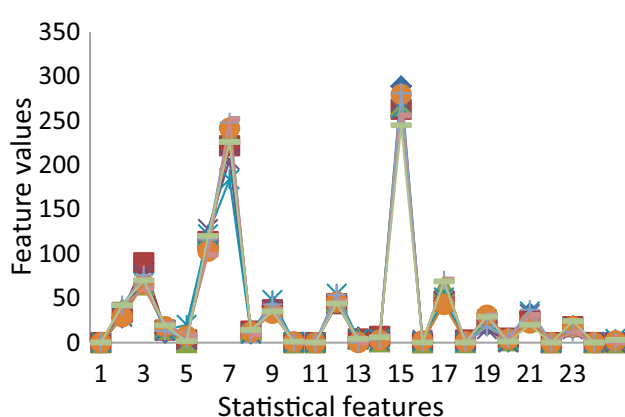


Fig. 14 Graph of 24 statistical features for 10 samples of wood species *Palaquium stellatum*

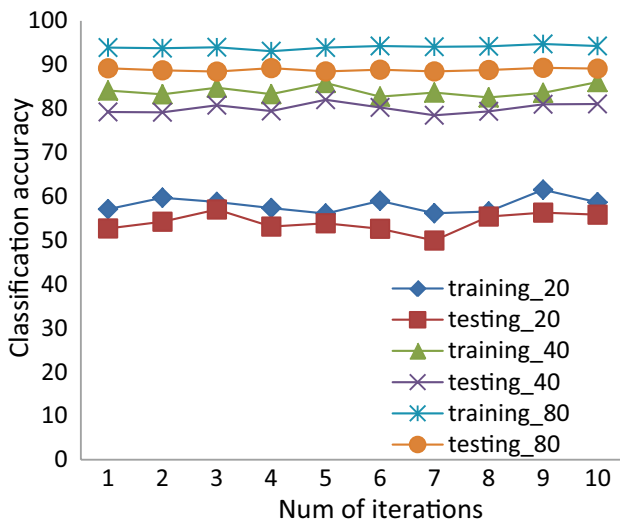


Fig. 15 Graph of classification accuracies for training samples and testing samples for different number of neurons in hidden layer

wood sample, q_1 represents the quantity of solitary vessels extracted from the wood image, q_2 represents the quantity of pair vessels extracted from the wood image, and q_3 represents the quantity of multiple vessels extracted from the wood image.

2.2.3 Vessels arrangement

There are 6 types of vessel arrangements for the wood texture, namely smallsmall, medmed, bigbig, smallmed, smallbig, medbig. The arrangements of vessels are determined by calculating the mean distance between vessels by using Euclidean distance approach. The vessels arrangements are determined based on fuzzy rules as below:

- IF distance between small vessel to small vessel THEN it is smallsmall.
- IF distance between medium vessel to medium vessel THEN it is medmed.
- IF distance between large vessel to large vessel THEN it is bigbig.
- IF distance between small vessel to medium vessel THEN it is smallmed.
- IF distance between small vessel to large vessel THEN it is smallbig.
- IF distance between medium vessel to large vessel THEN it is medbig.

Each wood image may consist of 6 types of arrangements: smallsmall, medmed, bigbig, smallmed, smallbig, medbig. Hence, the quantity of each type of arrangements is computed for each wood sample. For example, $x_1 = [r_1 r_2 r_3 r_4 r_5 r_6]$ where x_1 represents a wood sample, r_1 represents

Table 2 Confusion matrix for 30 wood species for 30 test samples

Known wood species (species label in wood data)	Predicted wood species (30 samples)	
	Correct species (samples)	Other species (samples)
<i>Dehaasia curtisii</i>	28	2
<i>Dyera costulata</i>	28	2
<i>Endospermum diadenum</i>	28	2
<i>Hevea brasiliensis</i>	28	2
<i>Palaquium stellatum</i>	28	2
<i>Pometia ridleyi</i>	28	2
<i>Calophyllum kunstleri</i>	27	3
<i>Cynometra malaccensis</i>	27	3
<i>Durio spp</i>	27	3
<i>Hopea pachycarpa</i>	27	3
<i>Kokoona sessilis</i>	27	3
<i>Madhuca utilis</i>	27	3
<i>M. caesia</i>	27	3
<i>M. quadrifida</i>	27	3
<i>Neobalanocarpus heimii</i>	27	3
<i>Pentace floribunda</i>	27	3
<i>Scorodocarpus borneensis</i>	27	3
<i>Shorea guiso</i>	27	3
<i>Shorea collina</i>	27	3
<i>Artocarpus dadah</i>	26	4
<i>Dialum kingie</i>	26	4
<i>Dipterocarpus baudii</i>	26	4
<i>Dryobalanops oblongifolia</i>	26	4
<i>Lophopetalum sp</i>	26	4
<i>Parashorea stellate</i>	26	4
<i>Pentace curtisii</i>	26	4
<i>Shorea laevis</i>	26	4
<i>Shorea uliginosa</i>	26	4
<i>Sp of lauraceaceae</i>	26	4
<i>Kokoona reflexa</i>	25	5

the quantity of smallsmall arrangement from the wood image, r_2 represents the quantity of smallmed arrangement extracted from the wood image, r_3 represents the quantity of smallbig arrangement extracted from the wood image, r_4 represents the quantity of medmed arrangement extracted from the wood image, r_5 represents the quantity of medbig arrangement extracted from the wood image, and r_6 represents the quantity of bigbig arrangement extracted from the wood image.

The total number of statistical features extracted from each wood image is 24 features (12 features from black vessels image and 12 features from white vessels image). The next step after feature extraction is the classification process where the wood species are classified into its own species based on the statistical features.

Table 3 Comparison of classification accuracy and database between the proposed feature extractor with previous works

Previous works	Wood database (Images/species)	Classification accuracy (%)
GLCM technique by Tou et al. (2007)	360/5	72%
GLCM technique by Tou et al. (2008)	360/5	72.8%
GLCM technique by Khalid et al. (2008)	1949/20	95%
GLCM technique by Bremanath et al. (2009)	Not reported/10	Not reported
PCA technique by Tang et al. (2009)	180/60	70–80%
GLCM technique by Filho et al. (2010)	1270/22	78%
Proposed fuzzy-based statistical feature extraction	3000/30 species	89.3%

2.3 Classification

A supervised artificial neural network algorithm has been implemented in various identification and prediction problems such as prediction of tension properties of cork (Iglesias et al. 2015) and trend prediction of foreign exchange rates (Zafeiriou and Kalles 2013). In this paper, a back propagation neural network is used for classification purposes. The training dataset consists of input signals assigned to corresponding target. The target is the desired output z (representing the 30 classes of species). The first layer has weights coming from the input. Each subsequent layer has a weight coming from the previous layer. The last layer is the network output. Adaption is done which updates weights with the specified learning function. The trained network is then used to simulate for the test data. The output of the simulation resulted in 30 class data for the test data. Comparison of accuracy is computed to determine the performance of the proposed system with previous systems (Antikainen et al. 2015). The classification accuracy is calculated based on percentage of correctly classified test samples over number of test samples as:

$$\text{Accuracy} = \frac{\text{Num. of correctly classified samples}}{\text{Total num. of test samples}} \times 100 \quad (9)$$

Besides that, a confusion matrix is tabulated to summarize the misclassification of each wood species. The confusion matrix is calculated by applying the classification model to test data in which the target values are already known. These target values are compared with the predicted (classified) target values.

3 Results and discussion

The wood database consists of 30 wood species where 100 images were taken from each wood species: 70 images for training and 30 images for testing. The training images are

classified into 30 species based on the statistical features extracted using the proposed statistical feature extractor. Then, unknown test images are classified based on the training images stored in the database. Figure 10 shows the wood texture images of 30 wood species used in this research.

The wood texture which is the wood grain form on the timber surface is very important in identifying wood species. The wood texture will not change even after going through several chemical procedures. The wood surface has several distinctive features that may be used to discriminate the wood species. Figure 11 shows the examples of types of vessels extracted from a wood image while Fig. 12 shows the examples of different sizes of vessels extracted from a wood image. These statistical properties of vessels on wood texture are used to classify the wood species by using the proposed automated wood species recognition system.

Table 1 tabulates examples of statistical features extracted from the 30 wood images presented in Fig. 10. As shown in Table 1, the statistical features extracted are the quantity of small, medium and large vessels for black vessels images which are represented as f_1 , f_2 and f_3 , respectively followed by the quantity of small, medium and large vessels for white vessels images which are represented as f_4 , f_5 and f_6 , respectively.

Variations in statistical features within inter wood species are shown in Figs. 13 and 14. The graphs represent 24 statistical features extracted from 2 different wood species. There are 10 samples taken from the same wood species and it can be seen that the statistical features of the same wood species are almost similar to each other. On the other hand, the statistical features vary significantly for different wood species which will be used for classification purposes.

However, there are small variations within the wood samples of the same species for the 4th feature and the 6th feature in Fig. 13 which correspond to quantity of pair vessels and quantity of multiple vessels, respectively. This is due to the nonlinearity of wood texture. The nonlinearity in wood features is affected by the variation

within wood species that are associated to its place of growth and age. Younger trees will have smaller vessels on its cross section and have lighter color. Trees that grow on lower ground will acquire more nourishment, as a result, they grow faster and the trunks are not as hard as trees that grow on the hills. Trees on the hill fight for nutrients hence they grow at slower pace. Their vessels are smaller than the trees that grow on lower ground. Nonlinearity in wood features could also exist when the wood samples are taken from different parts of the same tree (Fichtler and Worbes 2012; Wheeler and Baas 1998).

Number of neurons used in the neural network classifier may affect the classification performance. There are 3 different quantities of neurons (20 neurons, 40 neurons, 80 neurons) tested in the proposed back propagation neural network. As the number of neurons in the hidden layer was increased, the classification accuracy improved significantly. This is because more neurons are required to process large size database. Besides that, 10 iterations were performed for training and testing the wood data. The number of iterations affects the generalization accuracy as shown in Fig. 15. The highest classification accuracy of the test dataset is approximately 89% when using 80 neurons in hidden layer.

In order to examine the classification performance for each wood species, a confusion matrix is tabulated as shown in Table 2. The confusion matrix shows how the predictions are made by the proposed wood species recognition system. There are 30 testing samples used for each wood species. The highest misclassification rate of 16.7% belongs to wood species *Kokoona reflexa*. This is due to the nonlinearities in wood texture whereby the statistical properties of vessels for wood species *Kokoona reflexa* are very close to several wood species such as *Kokoona sessilis* which contributed to the misclassification of wood species.

As shown in Table 2, this paper focused on analyzing statistical features of vessels on 3000 wood images which results in 89.3% classification accuracy for 30 wood species. This shows that classifying wood species based on statistical properties of vessels is adequate and reliable to be implemented in the timber industries. The advantage of the proposed system compared to previous works is that the proposed system enables human intervention when analyzing the wood texture and to aid human–machine interface.

Finally, the proposed system is benchmarked with the previous works. Table 3 summarizes the results published in literature using image analysis approach to classify wood species. The results show that the proposed system managed to classify wood species more effectively compared to previous works that implemented feature extractor approach based on GLCM and PCA techniques.

4 Conclusion

In this work, a cascaded wood species recognition system based on image analysis is presented. The proposed classification method imitates the approach taken by human experts in analyzing wood surface texture. First, statistical features are extracted from each wood species and used in a broad initial grouping of wood texture using fuzzy logic. Then a neural network classifier is used to refine the broad grouping into the final wood species classification. The supervised artificial neural network trained by back propagation algorithm is able to classify 30 wood species with an accuracy of approximately 89%. In spite of the improvement introduced by the proposed method compared to the performance of previous works using GLCM and PCA features, it is clear that there is a lot of room for improvement. It can be observed that the statistical features used for classification do not identify a species definitively in all cases. For instance, in an image of a wood species, there exist small, medium and large vessels in different quantities. Perhaps the pattern that these vessels form, rather than their statistics is more important. The challenge now is to find better features and gain better understanding of how human experts classify wood textures. Gabor filters or wavelet transform can be used to extract multiresolution textural features of vessels and parenchyma that form the wood pattern. Color information can also help in the discrimination of wood species that look alike.

Acknowledgements The authors would like to thank Malaysian Ministry of Higher Education (MOHE) and University of Malaya for funding this research through BKP Grant (BK047-204) and UMRG Grant (RP023-2012B). The authors also would like to thank Forest Research Institute of Malaysia (FRIM) for providing us with the wood samples.

References

- Anninou AP, Groumpos PP (2014) Modeling of Parkinson's disease using fuzzy cognitive maps and non-linear hebbian learning. *Int J Artr Intell Tools* 23(5): 1450010
- Antikainen T, Rohumaa A, Hughes M, Kairi M (2015) Comparison of the accuracy of two online industrial veneer moisture content and density measurement systems. *Eur J Wood Prod* 73:61–68
- Baas P, Wheeler E (2000) Dicotyledonous wood anatomy and the APG system of angiosperm classification. *Bot J Linn Soc* 134(1–2):3–17
- Batuwita R, Palade V, Bandara DC (2010) A customizable fuzzy system for offline handwritten character recognition. *Int J Artif Intell Tools* 20(3):425–455
- Bremanath R, Nithiya B, Saipriya R (2009) Wood species recognition using GLCM and correlation. *Int Confer Adv Recent Technol Commun Comput* 615–619
- Choffel D (1999) Automation of wood mechanical grading—coupling of vision and microwave devices. *Proceedings of SPIE* 1999. *Int Soc Optic Eng* 3836:114–121

- Fichtler E, Worbes M (2012) Wood anatomical variables in tropical trees and their relation to site conditions and individual tree morphology. *IAWA J* 33(2):119–140
- Filho PL, Oliveira LS, Britto Jr AS, Sabourin R (2010) Forest species recognition using color-based features. In: 20th International Conference on Pattern Recognition 4178–4181
- Fuentealba C, Simon C, Choffe, D, Charpentier P, Masson D (2004) Wood products identification by internal characteristics readings. *Proceed IEEE Int Confer Indust Technol* 2:763–768
- Gdyczynski CM, Manbachi A (2014) On estimating the directionality distribution in pedicle trabecular bone from micro-CT images. *Physiol Measur* 35(12):2415–2428
- Hermanson JC, Wiedenhoeft AC (2011) A brief review of machine vision in the context of automated wood identification systems. *IAWA J* 32:233–250
- Iglesias C, Anjos O, Martinez J, Pereira H, Taboada J (2015) Prediction of tension properties of cork from its physical properties using neural networks. *Eur J Wood Prod* 73:347–356
- Khalid M, Eileen LYL, Yusof R, Nadaraj M (2008) Design of an Intelligent Wood Species Recognition System. *Int J Simul Syst Sci Technol* 9(3):9–19
- Kuncheva LI (2000) How good are fuzzy if-then classifiers? *IEEE Trans Syst Man Cybern Part B-Cybern* 30(4): 501–509
- Menon PKB, Sulaiman A, Choon LS (1993) Structure and identification of Malayan woods. *Malayan Forest Records* No 25. Forest Research Institute Malaysia, Malaysia
- Peng F, Li JT, Long M (2015) Identification of natural images and computer-generated graphics based on statistical and textural features. *Forens Sci* 60(2): 435–443
- Piuri V, Scotti F (2010) Design of an automatic wood types classification system by using fluorescence spectra. *IEEE Trans Syst Man Cybern Part C Appl Rev* 40(3): 358–366
- Ramli R, Arof H, Ibrahim F, Idris MYI, Khairuddin ASM (2015) Classification of eyelid position and eyeball movement using EEG signals. *Malaysian. J Comp Sci* 28(1):28–45
- Rojas JAM, Alpuente J, Postigo D, Rojas IM, Vignote S (2011) Wood species identification using stress-wave analysis in the audible range. *Appl Acoust* 72:934–942
- Ruffinatto F, Crivellaro A, Wiedenhoeft AC (2015) Review of macroscopic features for hardwood and soft-wood identification and a proposal for a new character list. *IAWA J* 36:208–241
- Tang YB, Cai C, Zhao FF (2009) Wood Identification Based on PCA, 2DPCA and (2D)²PCA. *Fifth International Conference on Image and Graphics* 784–789
- Tou JY, Lau PY, Tay YH (2007) Computer vision based wood recognition system. In: *International workshop on advanced image technology* 16
- Tou JY, Tay YH, Lau PY (2008) One-dimensional grey-level cooccurrence matrices for texture classification. In: *International symposium on information technology* 1–6
- Wheeler EA (2011) In-sideWood - a web resource for hardwood anatomy. *IAWA J* 32 (2): 199–211.
- Wheeler EA, Baas P (1998) Wood identification- a review. *IAWA J* 19(3):241–264
- Zafeiriou T, Kalles D (2013) Short-term trend prediction of foreign exchange rates with a neural-network based ensemble of financial technical indicators. *Int J Artif Intell Tools* 22(3):1350016
- Woods RE, Gonzalez RC (2008) *Digital Image Processing* (3rd ed). Prentice Hall