World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:10, No:3, 2016

Statistical Feature Extraction Method for Wood Species Recognition System

Mohd Iz'aan Paiz Bin Zamri, Anis Salwa Mohd Khairuddin, Norrima Mokhtar, Rubiyah Yusof

***Abstract*—**Effective statistical feature extraction and classification are important in image-based automatic inspection and analysis. An automatic wood species recognition system is designed to perform wood inspection at custom checkpoints to avoid mislabeling of timber which will results to loss of income to the timber industry. The system focuses on analyzing the statistical pores properties of the wood images. This paper proposed a fuzzy-based feature extractor which mimics the experts' knowledge on wood texture to extract the properties of pores distribution from the wood surface texture. The proposed feature extractor consists of two steps namely pores extraction and fuzzy pores management. The total number of statistical features extracted from each wood image is 38 features. Then, a backpropagation neural network is used to classify the wood species based on the statistical features. A comprehensive set of experiments on a database composed of 5200 macroscopic images from 52 tropical wood species was used to evaluate the performance of the proposed feature extractor. The advantage of the proposed feature extraction technique is that it mimics the experts' interpretation on wood texture which allows human involvement when analyzing the wood texture. Experimental results show the efficiency of the proposed method.

***Keywords*—**Classification, fuzzy, inspection system, image analysis.

Open Science Index, Computer and Information Engineering Vol:10, No:3, 2016 publications.waset.org/10003749/pdf

I. INTRODUCTION

WITH more demands in timber industries and more tightly controlled international requirements, timber industries are required to meet tighter security requirements such as more accurate identification of the correct timber species, prevention of fraud and illegal logging. The wood texture is examined from the timber surface through a magnifier that has 10 times magnification [1]. Several automatic tropical wood species recognition systems have been developed by [2]-[7] to analyze the macroscopic features on the wood surface texture.

Statistical features are inherent to the content of images. Generally, different images will have different statistical properties [8]. The human-decision-based recognition system is basically based on visual inspection of the wood anatomy textures which can be presented as image data processing using statistical parameters representing the texture. In addition, fuzzy rules are capable of handling approximate data for sharply defined problems. Hence, the linguistic interpretation of the human behavior provided by a fuzzy model could be useful to experts in determining the wood species. Due to its simplicity and similarity to human reasoning, fuzzy logic algorithms have been applied in various imaging applications [9].

Anis Salwa Mohd Khairuddin and Norrima Mokhtar are with the Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia (e-mail: anissalwa@um.edu.my).

Mohd Iz'aan Paiz Bin Zamri, postgraduate student, is with the University of Malaya, Malaysia.

Rubiyah Yusof is with Malaysia Japan Institute of Technology Malaysia, Universiti Teknologi Malaysia.

In this paper, we address the problem of wood species recognition based on statistical features that enables experts' interference when analyzing the wood texture. Hence, we proposed a new fuzzy-based statistical feature extraction technique which focused on the pores distribution on the wood texture in order to mimic the experts' interpretation on wood texture. The contributions of this work are firstly, to improve the man-machine interface in environments for computer-aided training of human operators and secondly, the knowledge acquisition can be achieved for users by carefully checking the rules discovered from the training patterns. Therefore, the human operators are able to monitor the system by analyzing the statistical features generated from the proposed system.

II. PROPOSED METHODOLOGIES

A. Image Acquisition

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. The wood samples are in cubic form (approximately 1 inch by 1 inch by 1 inch in size). A specially designed portable camera with 10 times magnification is used to snap the images of the wood texture. The size of each image is 768 x 576 pixels. The wood images are pre-processed using homomorphic filters in order to enhance the image presentation. Homomorphic filters are used to sharpen the features on the image and flatten the lighting variations of an image. Therefore, illumination and reflectance on the images are removed.

B. Feature Extraction

The statistical feature extraction process consists of two steps namely pores extraction and fuzzy pores management (Fig. 1). 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 pores 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

Open Science Index, Computer and Information Engineering Vol:10, No:3, 2016 publications.waset.org/10003749/pdf

shown in Fig. 2. The pores extraction process extracted the statistical properties of pores from both binary images for each wood sample.

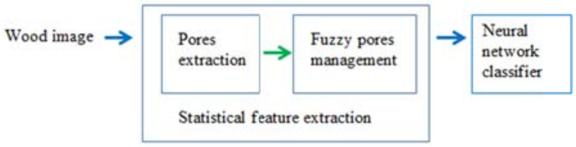
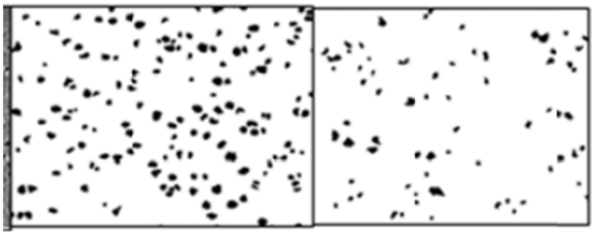


Fig. 1 The flowchart of the proposed wood recognition system



(a) (b)

Fig. 2 (a) Black pores image of *Palaquium stellatum* and (b) white pores image of *Palaquium stellatum*

In this research, the area of a region was detected and measured by pixels which mean we used the number of pixels inside a region to represent the area of region. The components in an image are classified to convert a binary image into a label matrix to compute region descriptors. 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 represents the area of the ith void. The ith void is defined as a pore if:

$$\frac{S_i-\hat{\mu}}{\hat{\sigma}} > \theta,\ \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} S_i \text{ and } \hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n} (S_i - \hat{\mu})^2 \quad (1)$$

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

The second step in the proposed statistical feature extraction is the fuzzy management process. Fuzzy if-else rules are employed to categorize the pores from both binary images into several features such as sizes of pores, types of pores and pores arrangements. Supposed we have 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: $x_p = [x_{p1}, x_{p2}, \ldots, x_{pn}]$, $p = 1,2,\ldots,m$. These training patterns are classified into M ($M \leq m$) classes. In this paper, the wood data has 38 feature variables with 52 classes, that is, n = 38, m = number of training samples, and M = 52. The 38 feature variables are the statistical features extracted from both binary images (black pores only and white pores only) such as:

1) Sizes of pores (6 features extracted from both binary images).

The measurement of pore size is defined by [1]. The pores are divided into 3 categories by using fuzzy rules: small pores, medium pores and large pores as:

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

2) Types of pores (6 features extracted from both binary images)

Pores may present in singles, pairs and multiple on the wood surface. Single pores are called solitary pores. Pair pores present as two pores combined and become a single pore. Chain pores present when multiple pores combine to form a long string of pores. The fuzzy rules are:

- IF pore is individual pore THEN it is a solitary pore
- IF pore is combination of 2 individual pores THEN pore is a pair pore
- IF pore is combination of 3 or more individual pores THEN pore is a chain pore.

3) Arrangement of pores (12 features extracted from both binary images).

The arrangements of pores are determined by calculating the mean distance between pores by using Euclidean distance approach. The pores arrangements are determined based on fuzzy rules as:

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

4) Number of vessels per square millimeter (mm) (2 features extracted from both binary images) is computed as:

$$\frac{\mathit{Num\ of\ pores\ in\ the\ image}}{\mathit{area\ of\ image\ in\ mm^2}} \quad (2)$$

5) Mean (4 features extracted from both binary images)

The average distance between pores and average size of pores are computed for each wood sample.

6) Standard deviation (4 features extracted from both binary images).

The standard deviation for pores distances and size of pores are computed for each wood sample.

7) Quantity of oval pores (2 features extracted from both binary images).
8) The pore is identified as an oval pore if $y \geq \frac{4}{3}x$ and $y < 2x$ where y and x represent the axis length perpendicular to each other.
9) Quantity of deformed pores (2 features extracted from both binary images).

The pore is identified as a deformed pore if $y \geq 2x$ where y and x represent the axis length perpendicular to each other.

Open Science Index, Computer and Information Engineering Vol:10, No:3, 2016 publications.waset.org/10003749/pdf

C. Classification

A supervised artificial neural network algorithm has been implemented in various identification and prediction problems such as iris image classification [10], thermographic crack detection [11], and classification of breast masses in mammograms [12]. In this paper, a back propagation neural network is used to classify the wood images. The training dataset consists of input signals assigned with corresponding target. The target is the desired output z (represent the 52 class of species). The trained network is then used to simulate the test data. The output of the simulation on the test data resulted to 52 class data. The classification accuracy of the proposed system is calculated based on percentage of correctly classified test samples over number of test samples as:

$$\text{Accuracy} = \frac{Num.of\ \text{correctly}\ classified\ samples}{Total\ num.of\ test\ samples} \times 100\ \% \qquad (3)$$

III. Results and Discussion

The wood database consists of 52 wood species where 100 images were taken from each wood species: 70 images for training and 30 images for testing. The wood surface has several distinctive features that may be used to discriminate the wood species. Fig. 3 shows the examples of types of pores and different sizes of pores extracted from a wood image. These statistical properties of pores on wood texture are used to classify the wood species by using the proposed automated wood species recognition system.

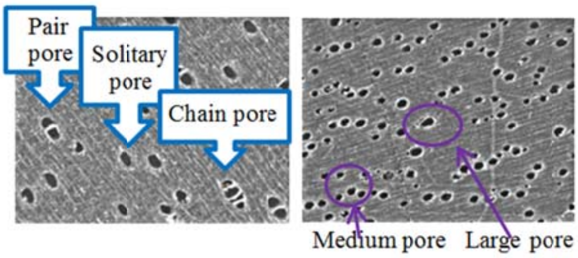


Fig. 3 Different types and sizes of pores on wood surface texture

This paper focused on analyzing statistical features of pores on 5200 wood images and used back-propagation neural network classifier to classify the wood images. The classification accuracy of the proposed system is tabulated in Table I. The neural network classifier is executed for several iterations. The number of iterations affects the generalization accuracy as shown in Table I. The highest classification accuracy is achieved at the 5^{th} iteration which is approximately 91.73 %. This shows that classifying wood species based on statistical properties of pores is reliable to be implemented in the timber industries.

TABLE I
The Classification Accuracy for Training and Testing Data

Experiment	1	2	3	4	5
Train (%)	95.88	95.00	94.86	96.13	95.82
Test (%)	90.58	89.10	89.94	90.83	**91.73**

Finally, the proposed system is benchmarked with the previous works. Table II summarizes the results published on the literature using image analysis approach to classify wood species. The previous works focused on extracting features based on the grey level of the images by using Grey Level Co-occurrence Matrix (GLCM) and Principle Component Analysis (PCA) techniques. As presented in Table II, majority of the previous works achieved classification accuracy of less than 80 % and all previous works used less than 2000 wood images in their experiments which are considered inadequate compared to the proposed system. Albeit the accuracy of the system proposed by [4] is 95 %, the accuracy dropped to 78% when the GLCM technique is implemented on the current wood database. Hence proving that GLCM feature extractor is inadequate due to the macroscopic anatomy image of wood is not uniform and contains variation of intensities which form certain repeated pattern called visual texture [13]. 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.

TABLE II
The 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 [2]	360 / 5	72 %
GLCM technique by [3]	360 / 5	72.8 %
GLCM technique by [4]	1949 / 20	95 %
GLCM technique by [6]	Not reported / 10	Not reported
PCA technique by [5]	180 / 60	70 % - 80 %
GLCM technique by [7]	1270 / 22	78 %
Proposed fuzzy-based statistical feature extraction	5200 / 52 species	91.73 %

IV. Conclusion

In this work, an automated wood species recognition system based on image analysis is presented. The proposed fuzzy-based statistical feature extraction which imitates experts' understanding of wood surface texture managed to extracts discriminant statistical features of wood pores which will be used for classification purposes. The supervised artificial neural network trained by back propagation algorithm able to classify 52 wood species with accuracy of approximately 91%. The simple experimental setup permits the adoption of the proposed method in real-time application. In spite of the improvement produced by the proposed fuzzy-based statistical feature extractor when compared to previous works that implemented GLCM and PCA feature extractors, it is clear that there is a lot of room for improvement. The challenge is how to extract most discriminant statistical features to mimic the experts' understanding on wood texture. For future works, we plan to investigate automatic techniques to extract the most discriminant wood features which will result to less misclassification rate and also a more robust classification module is needed to accommodate larger wood database.

ACKNOWLEDGMENT

The authors would like to thank Malaysian Ministry of Higher Education (MOHE) and University of Malaya for funding this research through 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

[1] P. K. B. Menon, A. Sulaiman, L. S. Choon, “Structure and identification of Malayan woods,” Malayan Forest Records No 25, Forest Research Institute Malaysia, Malaysia. (1993).
[2] J. Y. Tou, P. Y. Lau, Y. H. Tay. “Computer vision based wood recognition system.” Proceedings of International workshop on advanced image technology. 16. (2007).
[3] J. Y. Tou, Y. H. Tay, P. Y. Lau, “One-dimensional grey-level cooccurrence matrices for texture classification,” International symposium on information technology. pp 1–6. (2008).
[4] M. Khalid, Y. L. Lew, R. Yusof, and M. Nadaraj, “Design of an intelligent wood species recognition system,” Simulation Systems, Science & Technology. 9(3), pp. 9-19. (2008).
[5] Y. B. Tang, C. Cai, F. F. Zhao, “Wood Identification Based on PCA, 2DPCA and (2D) 2PCA”. Fifth International Conference on Image and Graphics. 784-789. (2009).
[6] R. Bremanath, B. Nithiya, and R. Saipriya, “Wood species recognition using GLCM and correlation.” International Conference on Advances in Recent Technologies in Communication and Computing. pp.615-619. (2009).
[7] P. L., Filho, L. S. Oliveira, Jr A. S. Britto, R. Sabourin, “Forest species recognition using color-based features.” 20th International Conference on Pattern Recognition. 4178–4181 (2010).
[8] F. Peng, J. T. Li, M. Long, “Identification of natural images and computer-generated graphics based on statistical and textural features.” Forensic Sciences. 60(2): 435-443. (2015).
[9] H. Yan and M. M. Gupta. “Guest editorial: Special section on neural networks and fuzzy logic for imaging applications,” Electronic Imaging. 6(3). (1997).
[10] Randy P. Broussard, Rober W. Ives. “Improving identification accuracy on low resolution and poor quality iris images using an artificial neural network-based matching metric.” Electronic Imaging. 20(1), (2011).
[11] B. O. Tranta, M. Sorger, P. O’Leary. “Thermographic crack detection and failure classification,” Electronic Imaging. 19(3). (2010).
[12] Tulio C. S. S. Andre’, R.M. Rangayyan. “Classification of breast masses in mammograms using neural networks with shape, edge sharpness and texture features,” Electronic Imaging. 15(1). (2005).
[13] M. Petrou, G.S. Sevilla, “Dealing with texture.” John Wiley & Sons Ltd. West Sussex, England (2006).

Open Science Index, Computer and Information Engineering Vol:10, No:3, 2016 publications.waset.org/10003749/pdf