



# Application of kernel-genetic algorithm as nonlinear feature selection in tropical wood species recognition system



Rubiyah Yusof<sup>a,\*</sup>, Marzuki Khalid<sup>a</sup>, Anis Salwa M. Khairuddin<sup>b</sup>

<sup>a</sup> Center for Artificial Intelligence and Robotics, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

<sup>b</sup> Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia

## ARTICLE INFO

### Article history:

Received 9 April 2012

Received in revised form 7 January 2013

Accepted 10 January 2013

### Keywords:

Wood recognition system

Texture identification

Nonlinear feature selection

Genetic algorithm

## ABSTRACT

Classifying tropical wood species pose a considerable economic challenge and failure to classify the wood species accurately can have significant effects on timber industries. Previous works on tropical wood species recognition systems considered methods for classification of linear features of the wood species. However, tropical wood species are known to exhibit nonlinear features due to several factors such as age of the tree, samples taken from different parts of the tree, etc. to address the nonlinear features of the tropical wood species, a Kernel-Genetic Algorithm (K-GA) technique for feature selection is proposed. This method combines the Kernel Discriminant Analysis (KDA) technique with Genetic Algorithm (GA) to generate nonlinear wood features and at the same time reduce dimension of the wood database. The proposed system achieved a classification accuracy of 98.69%, showing marked improvement to the work done previously.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

With more demands and more tightly controlled international requirements, many of the tropical countries are required to meet tighter security requirements as well as higher technical demands such as more accurate recognition of the correct timber species, prevention of fraud and illegal logging, and Environmental Investigation Agency (EIA) requirements, to name a few. In many timber industries, one of the major problems is to find good wood graders. Currently, very few certified officers are involved in the traditional wood identification process. The process of training up experienced officers in performing the job is difficult since the job is no longer considered lucrative and rather laborious. Moreover, the possibility of biasness and mistakes in human recognition system has to be considered. Besides that, it is impractical and cost effective for a human to analyze and identify large amount of timber species. Hence automatic wood species recognition system is needed to overcome the errors caused by traditional wood identification system which is based solely on human expertise.

Image analysis techniques are frequently used in classifying agricultural data (Ruiz et al., 2011; Camargo and Smith, 2009). Within the last decade, numerous wood species recognition have been proposed in classifying wood species based on image analysis

(Pan and Kudo, 2011; Wang and Bai, 2007; Bremanath et al., 2009). These studies concentrates on classifying tropical wood species based on the macroscopic features of the wood sample. However, these works are applied on less than 10 samples of wood species and hence the extent of the problem of identifying the tropical wood species has not been explored.

Khalid et al. (2008) developed an automatic tropical wood species recognition system based on 20 wood species using Gray Level Co-occurrences Matrix (GLCM) as feature extractor to classify 20 tropical wood species. In a later development, Khairuddin et al. (2011) implemented the fusion of basic gray level aura matrix (BGLAM) and statistical properties of pores distribution (SPPD) to extract features from wood surface for 52 tropical wood species. Genetic Algorithm (GA) based feature selection is introduced to classify the 52 tropical wood species. Although the accuracy of the recognition of the tropical wood species improved with the use of the GA for feature selection, the system proposed by Khairuddin et al. (2011) requires large memory space due to the large dataset. Moreover, the system does not consider the nonlinear features of the tropical wood species directly.

There are three main challenges in the design of tropical wood recognition system based on image processing which are large training data size, finding suitable feature extraction for nonlinear wood texture representation and classification method of a new wood image based on the chosen feature representation. Feature selection techniques have been gaining a lot of interest in the area of image classification due to the ability to deal with large data size problem through the reduction of the number of features. A good

\* Corresponding author. Tel.: +60 3 26913710; fax: +60 3 26970815.

E-mail addresses: [rubiyah@ic.utm.my](mailto:rubiyah@ic.utm.my) (R. Yusof), [marzuki@utm.my](mailto:marzuki@utm.my) (M. Khalid), [anissalwa@um.edu.my](mailto:anissalwa@um.edu.my) (A.S. M. Khairuddin).

feature selection technique facilitates classification of the images by eliminating redundant features that can impede recognition.

Genetic Algorithm (GA) has been shown to be an effective method for feature selection through previous works by Sun et al., 2005; Lu and Zhao, 2008; Zheng et al., 2005; Grefenstette et al., 2005; Hussein et al., 2001. It is a robust technique and can work in a large database. GA is widely used as feature selection technique in various applications such as classification of wood species (Khairuddin et al., 2011), classification of medical and biochemistry data (Raymer et al., 2000), classification of protein-bound water molecules (Raymer et al., 1997), and face recognition system (Kanan and Faez, 2008). However, all of these methods are used for feature selection in linear feature extraction methods.

As the problems of pattern classification based on images are becoming more complex, the solution using linear models is no longer adequate. In these types of problems, the target values cannot be expressed with simple linear combination of features. GA could not achieve good classification accuracy due to the fact that basic GA feature selection does not involve any functions that provide solutions for nonlinearity. To address the problem, researchers have used GA neural network as a nonlinear feature selection due to the ability of the neural network to solve for nonlinearities. However, the optimization of the GA based neural network models is a complex procedure. A novel genetic algorithm–kernel partial least square (GA–KPLS) has been proposed as a novel nonlinear feature selection (Mehdi and Kyani, 2007). However, this technique is only suited for the analysis of genomic and proteomic data.

To this end, various nonlinear approaches have been developed to solve nonlinear problems. One of the nonlinear approaches is using kernel representations for these complex problems. Kernel representations increase the computational power of linear approaches by transforming data to the higher dimensional feature space, which enables the use of some linear approaches to solve nonlinear problems. A number of nonlinear feature extraction techniques based on kernel trick have shown to be effective in reducing the dimension of a dataset nonlinearly and results in high classification accuracy such as in (Mu et al., 2010; Jiang et al., 2008; Lu et al., 2003; Baudat and Anouar, 2000; Yu and Xu, 2008). Kernel Discriminant Analysis (KDA) based on generalized singular value decomposition (GSVD) technique which is a nonlinear extension of linear discriminant analysis (LDA) based on kernel function and GSVD has shown to be effective in solving a multiclass problem (Park and Park, 2005; Yang and Tian, 2011; Wang et al., 2007). When implementing KDA technique, the possibility of singularity problem of total scatter matrix needs to be handled. The singularity problem occurs when the number of data points is smaller than the dimension of the data space. GSVD technique has been proposed to overcome this problem in (Park and Park, 2005; Howland and Park, 2004; Paige and Saunders, 1981; Ye et al., 2004).

In order to use a GA based feature selection method for solving a nonlinear problem in image processing, nonlinearity should be provided in the system via appropriate transformations like kernel transformations. To overcome the earlier mentioned problems for automatic tropical wood recognition system, this paper presents a novel approach of nonlinear feature selection in the tropical wood species recognition system. The strong capability of the GA in feature selection is effectively combined with the capability of the KDA which can perform nonlinear dimension reduction without substantially downgrading the system's performance. The proposed kernel-GA method acts as an enhanced feature selection, to perform nonlinear feature selection.

This paper is structured as follows. Section 2 describes the data acquisition module while Section 3 focuses on the proposed Kernel-GA for nonlinear feature selection and the classification of

tropical wood species. Finally, Section 4 presented the experimental results and discussion of the performances of the overall system.

## 2. Nonlinear wood image acquisition

### 2.1. Nonlinear features in classifying tropical wood species

The distribution of wood images, under a perceivable variation in viewpoint, and illumination, is highly nonlinear and complex. The nonlinearity in tropical wood features is caused by the variation within wood species that are directly related to its age and place of growth. Younger trees will have slightly smaller pores on its cross section and has lighter color. Trees that grow on lower ground gets more nourishments therefore they grew faster and the trunk are not as hard as trees that grew on the hills. Trees on hill struggles for nutrients therefore they grew in slower pace. Their pores are also smaller than the trees that grew on lower ground. Nonlinearity in wood features could also be present when the wood samples are taken from different part of the same tree. The examples of nonlinear features are the presence of resin canal in some wood samples, different color tones, density and size of pores. Some of the wood samples are also filled with dust, dammar or deposit. Size of pores of certain trees also changes through the years. In this research, the wood samples are taken randomly from the same or different trees.

Even when two samples taken from the same tree, but at different position, one nearer to the pith and another is nearer to the bark, may have some differences. Fig. 1 is the texture of wood species *Shorea laevis* taken from these two different areas. The area near to the pith is called heartwood and the area near to the bark is called sapwood. For some tropical wood species, wood in heartwood region has darker color intensity, Fig. 1a, than in sapwood, Fig. 1b. The density of pores is also higher at the heartwood.

In this paper, the nonlinear features are composed of basic gray level aura matrix (BGLAM) features and statistical properties of pores distribution (SPPD) features which will be explained in Sections 2.3.2 and 2.3.3 respectively.

### 2.2. Image acquisition system

The wood samples for this research are obtained from the Forest Research Institute of Malaysia (FRIM). There wood samples are in cubic form (approximately 1 in. by 1 in. in size). One of the main problems in image acquisition is to maintain the quality of the images captured as well as the portability of the capturing device. The image captured must be consistent with the images in the database which are used for training purposes. The difference in the quality of images will affect the rate of recognition accuracy greatly. In order to maintain the same quality of image, factors like magnification or field of view, pattern and format and distance from the camera to wood sample should be kept as constant and consistent as possible.

Hence, a specially designed portable camera with 10 times magnification is used to capture the wood texture images. A USB camera is chosen since USB technology is much more cost effective. The USB camera power is supplied by the USB port; hence, no additional battery is required. The camera is equipped with a systematic focusing function whereby the distance between the camera and the wood sample is fixed and need not be adjusted. The housing of the camera is designed to cut all ambient light and its fluctuations. The camera is equipped with lighting system which provides sufficient lighting homogeneity.

The data acquisition process, Fig. 2, where wood images are captured from each wood cubes by using the portable camera.

Download English Version:

<https://daneshyari.com/en/article/6541046>

Download Persian Version:

<https://daneshyari.com/article/6541046>

[Daneshyari.com](https://daneshyari.com)