



Feature decision-making ant colony optimization system for an automated recognition of plant species



Mohammad Ali Jan Ghasab^{a,*}, Shamsul Khamis^b, Faruq Mohammad^c, Hessam Jahani Fariman^a

^a Department of Electrical Engineering, Universiti Putra Malaysia, 43400 (UPM), Serdang, Selangor, Malaysia

^b Unit of Biodiversity, Institute of Bioscience, Universiti Putra Malaysia, 43400 (UPM), Serdang, Selangor, Malaysia

^c Institute of Advanced Technology, Universiti Putra Malaysia, 43400 (UPM), Serdang, Selangor, Malaysia

ARTICLE INFO

Article history:

Available online 15 November 2014

Keywords:

Plant recognition
Feature subset selection
Ant colony optimization
Leaf analysis
Automatic leaf classification

ABSTRACT

In the present paper, an expert system for automatic recognition of different plant species through their leaf images is investigated by employing the ant colony optimization (ACO) as a feature decision-making algorithm. The ACO algorithm is employed to investigate inside the feature search space in order to obtain the best discriminant features for the recognition of individual species. In order to establish a feature search space, a set of feasible characteristics such as shape, morphology, texture and color are extracted from the leaf images. The selected features are used by support vector machine (SVM) to classify the species. The efficiency of the system was tested on around 2050 leaf images collected from two different plant databases, FCA and Flavia. The results of the study achieved an average accuracy of 95.53% from the ACO-based approach, confirming the potentials of using the proposed system for an automatic classification of various plant species.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The biodiversity of plants is an indispensable foundation for most of our terrestrial ecosystems and hence all the creatures living on earth are either directly or indirectly depended on plant species as they provide different forms of energy in a natural means. Since the plants maintains the physiology which converts the carbon dioxide into oxygen (essential for living organism) by taking light energy from the sun in a natural photosynthetic way and so, they considered to be the main suppliers of oxygen on earth. In addition, a remarkable majority of different plant species are utilized for a wide range of industrial applications which includes the nourishment, herbs and plant originated therapeutically active ingredients for medical/pharmaceutical sector, biofuels for sustainable energy and woody biomass for renewable energy production to mention some (Adam, Khamis, Ismail, & Hamid, 2012; Prochnow et al., 2009; Sulaiman et al., 2008).

Remarkably, due to the continuous destruction of our natural environment by means of increasing deforestation in recent years for the purposes of rapid industrialization and new constructions, the plant species especially are facing the endanger. Therefore, in

order to preserve the biodiversity of plant species, the field of plant taxonomy works for the identification of newfound species and further their categorization into existing plant families to which they belong to (Stuessy, 2009). In addition to the identification and naming of new plant species, the science of plant taxonomy in recent years also finding the ways for the usage of newly documented plants for industrial product development. The current process of new species identification in general requires experts, as they only can provide the taxonomic tags following the examination of related plant specimens. However for bringing up such experts, a long-term training and fully established financial resources are needed. On the other hand, by keeping in view of the diverse plants species, the use of traditional methods for classification purposes are time consuming. Therefore, it is very important to develop accurate, fast, and efficient system for an automatic identification of wide range of plant species, as the currently available methods are time consuming and expert based analysis.

In modern-day, the ubiquity of advanced technologies such as the digital cameras and hand-held computers has brought the idea of automatic plant identification closer to reality (Husin et al., 2012; White, Marino, & Feiner, 2007). In addition, the creation of multidisciplinary studies such as the image processing and machine learning technology in computer science sector also has motivated many scholars to perform research in this field to achieve systems of non-manual plant classification. The challenging task in this are the problems associated with the extraction

* Corresponding author. Tel.: +60 176621880/390191050; fax: +60 3 89466327.

E-mail addresses: mohammad.ali.jan.my@ieee.org (M. Ali Jan Ghasab), Khamis.Shamsul@gmail.com (S. Khamis), farooqm1983@gmail.com (F. Mohammad), h.jahaniFariman.1988@ieee.org (H. Jahani Fariman).

of discriminative features that can be applicable for distinguishing various plant species. Feature extraction from digital images, an inclusive term, functions to extract the innate properties of a group of objects out of their images that are common for those particular groups and uncommon in other groups (Du, Huang, Wang, & Gu, 2006). By making use of such set of features, all groups of objects can be classified into different classes to which they belong to. In this process, the system selects a set of more discriminant features for each group and hence better accuracy will be achieved from the classification part. Some of the feature extraction studies have shown that, of many different approaches available in plant taxonomy, the method based on leaf shape analysis seems to extract the best effective features. Also from the contexts of plant classification, leaves are the most popular part as they carry out plant's inherent properties and are readily accessible for examination rather than other parts and thus suits well for the implementation of "feature extraction" for plant identification (Gwo, Wei, & Li, 2013; Lee & Chen, 2006). The various different features from leaf images such as shape, color, morphology, texture, venation structure, etc have been extracted by scientists in order to evaluate the impacts of these attributes for the recognition of plant varieties (Arribas, Sánchez-Ferrero, Ruiz-Ruiz, & Gómez-Gil, 2011; Bruno, de Oliveira Plotze, Falvo, & de Castro, 2008; Kebapci, Yanikoglu, & Unal, 2011; Shabanzade, Zahedi, & Aghvami, 2011). Among the available methods in plant taxonomy, the centroid-contour distance method for representing leaf shape uses a contour descriptor, as it describes the contour of a shape by measuring chain of values from leaf's outline, starting from a point to trace the outline in either clockwise or counter-clockwise direction (Chaki & Parekh, 2011; Meade & Parnell, 2003). While doing this, the problems due to self-intersection arise when one part of the leaf covers other part of the same leaf, and the researchers attempted to solve this difficulty by assuming the darker part of the leaf as the overlapping part and tried to represent the true outline by compensating this covered area (Mokhtarian & Abbasi, 2004). However, the obligation of adequate brightness and the usage of backlit leaves in order to produce the darker region of the overlap are main limitations associated with this centroid-contour distance method while accessing the leaves for plant identification.

Quantitative geometrical features (QGFs) of a leaf are common features based on a two-dimensional contour model which considers the form and structure of the leaf shape to extract a set of geometrical features such as length, width and the boundary (Xiaofeng, Deshuang, & Du Ji-xiang, 2006). In this approach, a specific threshold is applied to the digital image of a leaf in order to determine its main region that can be separated easily from the background. As a result, the generated binary image of the leaf can be used for the extraction of shape features. Du et al. introduced digital morphological features (DMFs) that comprise the extracted geometric features of the leaf shape contour (Du, Wang, & Zhang, 2007). By using this leaf geometric characteristic approach, 12 different DMFs were extracted by Wu et al. to discern 32 different kinds of plant species (Wu et al., 2007). In addition to DMFs, Pauwels et al. introduced a more specific geometric feature, lobedness to further improve the classification process (Pauwels, de Zeeuw, & Ranguelova, 2009). Although the QGFs as intuitive features applicable for various leaf shapes and enjoy the considerable benefits of fast and easy analysis, however, the usage of these features to represent the shape of leaf is a matter of oversimplifying the task. Dealing with such features in general increases the lack of operating meaningful analysis due to their high correlation with each other (Cope, Corney, Clark, Remagnino, & Wilkin, 2012).

Regarding the texture and vein structure analysis of the leaves, independent component analysis (ICA) was employed as one of the technique (Li, Chi, & Feng, 2006). In this strategy, ICA is applied to patched plant leaves to learn the linear basis functions as a pattern

map by which the veins are detected. Continuously, the gray-level images are converted into these pattern maps in which the leaf, background, edge and the other pixels are categorized into various classes by pattern matching.

Further, Nam et al. proposed an adjacency matrix made from the end points of a venation and the intersection in order to model the correspondence between two plant leaf images (Nam, Hwang, & Kim, 2008). A graph-matching algorithm, "venation matching" is applied to obtain the venation similarities between the leaves. In order to measure this similarity, the distance and pattern between two graphs of the venation is considered, and then based on the weighted sum of values of venation similarity, the final correspondence is calculated. One of the effective methods proposed by Zheng et al. to obtain the visible structure of the leaf veins follows five successive steps to extract the vein network in an order that the gray-scale of the leaf image is gained from the colored image at first in order to remove the overlapped color over the entire set of veins of the leaf and the entire background (Zheng & Wang, 2010). In the second step, the process of gray-scale morphology is applied and during the third step, linear intensity tuning is performed to increase the gray quantity difference between the leaf background and its veins. Following this, a threshold is applied to segment the image using Otsus method and after processing the details, the vein features are extracted accordingly (Otsu, 1975). Other methods available for the extraction of texture features include the Gabor filters and wavelet transforming techniques (Casanova, de Mesquita Sá Junior, & Bruno, 2009; Ishak, Hussain, & Mustafa, 2009; Liu, Zhang, & Deng, 2009).

For the identification of unknown plant species, each type of species considerably needs its own inherent features in order to be recognized correctly. However, the selection of unsuitable features may result in a combination of irrelevant, redundant or misleading data which further limits the systems performance. Hence, presenting an efficient feature decision making system for automatic recognition of plant species classification is still a challenge (Cope et al., 2012). Although some methods such as Probabilistic Neural Network (PNN) with principal component analysis, support vector machine (SVM) utilizing Binary Decision Tree (BDT) and Fourier Moment, are already available for the identification, obtaining the faster and accurate results with those approaches is still a challenge as each method has its own limitations (Singh, Gupta, & Gupta, 2010).

Therefore, the objective of this project is to develop a layman technique for an automatic recognition of plant species. For that, we presented an expert algorithm method for a programmed recognition of diverse plant species based on the selection of discriminant features from the leaves. Based on the proposed method, we first obtained the raw data related to plant leaves through a scanner. Following the operation of a series of preprocessing on the collected images, a number of features such as the shape, morphology, texture, and color of the leaves were extracted which constitutes the feature search space. To make a decision from the selection of optimal subset of features, ant colony optimization (ACO) algorithm was employed to investigate the inside feature search space in order to determine the best discriminant features with respect to each database of species. Finally, the selected features are used by multi-class support vector machine (SVM) to classify the species. The general steps for recognition of plant species are shown in Fig. 1 as further sections give more detailed information about each part.

2. Materials and methods

The scheme for an automatic recognition of plants based on feature decision making method has been shown in Fig. 2. Initially,

Download English Version:

<https://daneshyari.com/en/article/382836>

Download Persian Version:

<https://daneshyari.com/article/382836>

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