Volume 4, Issue 9, September 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Plant Identification and Classification Using Fuzzy Methods of Segmentation

Er. Simranjit Kaur Dhindsa*
Asra College of Engineering & Technology
India

Rajbir Singh
Asra College of Engineering & Technology
India

Abstract-Plants play an important role in our environment. Without plants there will be no existence of the earth's ecology. But in recent days, many types of plants are at the risk of extinction. To protect plants and catalogue various types of flora diversities, a plant database is an important step towards conservation of earth's biosphere. There are a huge number of plant spices world wide. To handle such volumes of information, development of a quick and efficient classification method has become an area of active research. In addition to the conservation aspect, recognition of plants is also necessary to utilize their medicinal properties and using them as sources of alter native energy sources like bio-fuel. There are several ways to recognize a plant, like flower, root , leaf , fruit etc. In recent times computer vision methodologies and pattern recognition techniques have been applied towards automated procedures of plant recognition. We have proposed a method for recognition and classification of leaves. This is a review paper regarding the work we have to done in our research.

Keywords:-Plant recognition, fuzzy clustering, computer vision.

I. INTRODUCTION

Leaf Analysis is a matter of interest for scientists. Classification and nomenclature of leaf is used only if leaves can be recognized. Botanists recognize leaves based on their knowledge and expertise, but for laymen leaf recognition is still a complicated task. Leaf recognition can be made simpler by using computer aided automation. A leaf recognition system should be based on a leaf classification system because there are more than one-half million of plants inhabiting the Earth and recognition without classification is a complex task.

Biodiversity of plants

Biodiversity [6] of plant shape is mainly attributable to biodiversity of leaf shape and the shape of floral organs, the modified leaves. However, the exact mechanisms of leaf-shape determination remain unclear due to the complexity of flat-structure organogenesis that includes the simultaneous cell cycling and cell enlargement in primordia. Recent studies in development and moleculer genetics have revealed several important aspects of leaf-shape control mechanisms. For example, understanding of polar control in leaf-blade expansion has advanced greatly.

Plant Kingdom [9] is mainly classified into two.

This type of plant classification is done according to how they reproduce.

1. Spore bearing plants (Algae, mosses, ferns and their relatives)
2. Seed bearing plants (Conifers and flowering plants)

Plant classification according to this criterion, can be pictorially represented as follows:

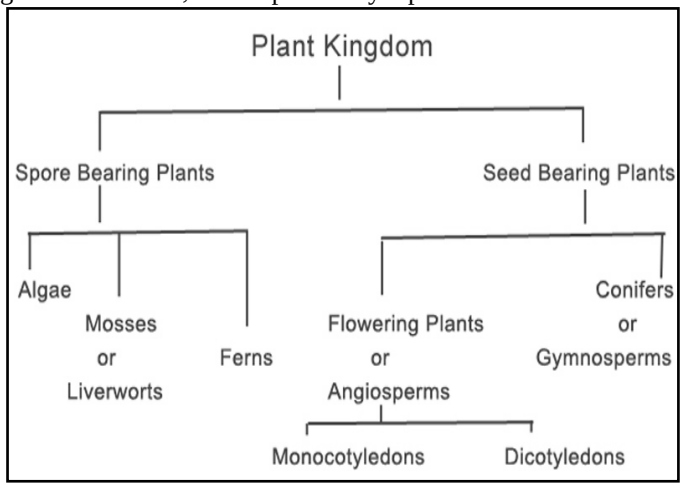

Figure 1.1 Plant Kingdom

Need of Biodiversity Biological diversity is important because of the way relationships between species and habitats combine to provide yet more variation in the living world. Any human activity that diminishes this 'bio-diversity' could therefore impoverish our own quality of life, reduce the resources available to us and ultimately jeopardise the survival of our descendants.

We should seek to conserve biodiversity [6] because:

- It confers direct benefits as natural processors protect our planet.
- It provides the raw material of food, clothing and medicines.
- It enhances our quality of life, by adding variety to our surroundings.
- It helps shape our culture and inspires our poets, painters and writers.
- It is heritage we should not deny to the next generation.
- It is affected enormously by what we do.

II. PLANTS IDENTIFICATION BY EXAMINING LEAVES

Plant species identification [8] requires recognizing the plant by various characteristics, such as size, form, leaf shape, flower colour, odor, etc., and linking it with a common or so-called scientific name. Correct identification provides basic information about size, shape and texture of a plant and can be helpful in protecting it from various types of pests and diseases. Plant species classification can be done through various ways like flower, root, leaf, fruit etc. Botanists adopted traditional classification method such as morphologic anatomy , cell biology and molecular biological approaches for doing plant species classification and recognition which is time consuming, troublesome and less efficient.

However advancement in computer technologies improves the process of plant species identification by designing automatic recognition system of plants. Plants classification can be done according to the structures of their leaf, bark, flower shapes, colours, textures and seedling morph. But if the plant classification is based on only two dimensional images then shapes of flowers, seedling and morph of plants are unsuitable because of their complex three dimensional structures. Since the plant leaves are two dimensional in nature they are well suited for classification of various plant species.

Plants species classification based on leaf image is the first choice as compred with other methods, such as cell and molecule biology methods. Sampling leaves and capturing leaves are low cost and convenient. Leaf image can be easily transferred to computer that can automatically extract features using various image processing techniques.

Leaves are obvious choice for identifying plants since they are so easily observed. They usually consist of two parts [6]. The blades wide are the most obvious part of leaf. The "stalk" or petiole by which blade is attached to the stem. There are three main types of venation in leaves:

1. In parallel-veinedleaves, the veins run parallel to each other. This condition is characteristics of the mono cotyledoneae. Parallel veins may run lengthwise on the leaf, as in Eucharis grandiflora(Amazon lily), or they may be parallel, but directed outward from the midrib to the margin (penniparallel).
2. Pinnately-veinedleaves have a single primary vein or midrib, from which smaller veins branch off, like the divisions of a feather. Examples are Eriobotrya japonica (loquat) and camellia japonica (camellia).
3. Palmately-veinedleaves have several principal veins radiating from the base of the leaf blade, as in Acer rubrum(red maple) and carica papaya (papaya).

Computer Vision

Computer vision [4] is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, or multidimensional data from a medical scanner. Computer vision, also known as machine vision, consists of three parts: measurement of features, classes of leaves based on the extracted features. Furthermore, the system uses the results of the classification scheme in identifying the class of new leaf images.

Measurement/Extraction of Features

Image processing technologies are used to extract a set of features/measurement that characterize or represent the image [1]. The values of these features provide a concise representation about the information in the image. For example, a set of features that characterize a triangle could be the length if each side of the triangle.

Pattern classification

Pattern classification [10] is the organization of patterns into groups of pattern sharing the same set of properties. Given a set of measurement of an unknown object and the knowledge unknown object belongs could be made.

For example, if a set of features/measurements is extracted from a leaf, a decision about the possible class of the leaf can be made. Pattern classification may be statistical or syntactic.

Pattern Recognition

Pattern recognition [10] is the process of classifying data or patterns based on the knowledge/information extracted from patterns. The patterns to be classified usually groups of measurements or observations defining points in an appropriate multidimensional space.

In this thesis, pattern recognition is implemented on a set of test images in order to validate and evaluate the performance

of the underlying classification scheme.

III. COLOR QUANTIZATION AND FUZZY CLUSTERING METHODS

Color quantization

A color image quantization is a process that reduces the number of distinct colors used in an image, usually with the intention that the new image should be as visually similar as possible to the original image.

The process of color image quantization is often broken into four phases, Heckbert [2].

Phase 1 is sampling the original image for color statistics.

Phase 2 is choosing a color map based on those statistics.

Phase 3 is mapping the colors to their representative in the color map.

Phase 4 is quantization and drawing the new image.

Phase 4 is a trivial matter regardless of the quantization method. The other three phase however are more strongly connected. In particular the method used for Phases 1 and 2 will determine the best method for accomplishing phase 3.

In general algorithms for color quantization can be broken into two categories: Uniform and Non-Uniform.

- Uniform: Here the color space is broken into equal sized regions where the number of regions, Nr is less than or equal to K.
- Non-Uniform: Here the manner in which the color space is divided is dependent on the distribution of colors in the image.

IV. FUZZY METHODS OF CLUSTERING

Fuzzy Algorithms:- In this no exist and one object is assigned to more than one cluster [3]. The most recent and effective fuzzy clustering techniques has been explained below.

- **Fuzzy clustering:-** Fuzzy clustering is also known as soft clustering. In this data elements can belong to more than one cluster, and associated with each element is a set of membership levels. These indicate the strength of the association between that data element and a particular cluster. Fuzzy clustering is a process of assigning these membership levels, and then using them to assign data elements to one or more clusters. Fuzzy clustering methods are Fuzzy C-Means and Kernelized Fuzzy C-Means.
- **Fuzzy C-Means(FCM) Algorithm:-** Fuzzy C-Mean clustering algorithm is one of the most widely used fuzzy clustering algorithm. This algorithm works by assigning membership to each data point corresponding to each cluster centre on the basis of distance between the cluster centre and the data point. More the data is near to the cluster centre more is its membership towards the particular cluster centre.

Kernelized Fuzzy C-Means (KFCM)

Algorithm:-

KFCM is an algorithm which is generated from FCM by modifying the objective function using Kernel induced distance matrix instead of Euclidean distance in FCM and thus the corresponding algorithm is derived and called as the kernelized fuzzy c-means (KFCM) algorithm, which is more robust than FCM.

V. LITERATURE REVIEW

Studies have been conducted in the past decade on automation of plant classification and recognition. A handful of these studies were about the extraction of a single feature from the image of a plant part such as the leaf, or the flower. Some studies were about the extraction of multiple features but from a single family of plants. Some studies focused on image-based plant classification, while others focused on image-based plant recognition.

Brendon J et al [11] detailed the image processing, and neural network classification methods applied to the task of identifying the pest that caused the damaged to apple fruits and leaves in orchards. They have used wavelet based analysis of fruit/leaf images consist of the fruit and the insect which is damaging it so therefore it is a classification problem which tries to identify fruit/leaf having particular pests damaging it. The objective of this thesis have been to take advantage of taking images of the fruit/leaf without doing manual labour in terms of inspection and climbing trees and manually checking the pests infected. This person is also using neural network based classifier for identifying different types of pests on particular set of fruits/leafs and the accuracy result is greater than 90%.

L. Tang et al [12] introduced low level features (texture features) to classify different types of grass weeds leaves. A pattern recognition system composed of a Gabor wavelet feature extractor and a feed forward back propagation ANN classifier was developed to classify weeds into broadleaf and grass classes. Particularly, a Gabor wavelet filter bank was designed to obtain joint space-frequency characteristics from a set of validation images demonstrated the potential of the method. The results of paper for each type of broad leaf and grass weed ranges from 88-92%. When compared with other statistical methods of using co-occurrence matrices, the developed feature extraction algorithm is computationally efficient and thus presents advantages in meeting real-time requirements.

Stephen Gang Wu et al [13] introduced PNN (Probabilistic Neural Network) for classification of leaf images based on 12 leaf features. A principal component analysis was used for reducing the 12 dimensions into five dimensions for faster processing. The 12 features used were physiological length, physiological width, leaf area, and leaf perimeter, smooth factor. Aspect ratio, form factor, rectangularity, narrow factor, perimeter ratio of diameter, perimeter ratio of physiological length and physiological width, and vein features. The use of scale variant features such as physiological length, physiological width, area, and perimeter might constraint this approach into using standard size for the leaf

image. If a standard size image is not used then the features of the same leaf vary with different sizes images. Since it is the same leaf, the value of all the features are expected to be same. The values for these features vary for different scaled versions of the same leaf image. In this research, features used are scaling, translation and rotation invariant.

D. Warren et al [5] introduced an automatic computerized system that used as its input 10 images of each chrysanthemum species for testing the variation in the images. In this study, features such as shape, size and color of the flower, petal, and leaf were described mathematically. Different rose features were extracted and used in the recognition scheme for pattern recognition. The study, however, was limited to chrysanthemum species only.

T. Satiohet al [7] studied an automatic method for recognition wild flowers. This recognition required two images: a frontal flower image and a leaf image taken by a digital camera. Seventeen features, eight from the flower and nine from the leaf, were fed to a neural network. This research yielded an accuracy of 95% on 20 pairs of pictures from 16 wild flowers. These studies dealt with a single or group of similar plant species only.

VI. CONCLUSION

Botanists recognize leaves based on their knowledge and expertise, but for laymen leaf recognition is still a complicated task. A leaf recognition system should be based on a leaf classification system because there are more than one-half million of plants inhabiting the earth and recognition without classification is a complex task. In my research work, those salient features of leaves will be studied that may be used as a basis for leaves classification and recognition will be studied to develop an approach that produces the best classification algorithm. Following objectives will be measured after developing the algorithm which will be based on fuzzy methods of clustering.

- To develop a representative dataset of leaf.
- Do scoring based leaf architecture.
- Develop data for machine.
- Run machine algorithm.

Our objective in research have been to develop a leaf classifier which works on the principal of extracting information based on its architecture using image processing to classify correctly in the plant kingdom. The idea is to develop a classifier which has less computation and complexity.

REFERENCES

[1] YotismitaChaki, Ranjan Parekh (2011)" Plant Leaf Recognition using shape based Features and Network classifiers" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol.2.

[2] Heckbert, P. "Color Image Quantization for Frame Buffer Display", Computer Graphics, Vol 16, #3, pp.297-303, 1982.

[3] Ivancsy, Renata and Kovacs, Ferenc (2006) "Clustering Techniques Utilized in web Usage Mining" proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrir, Spain.

[4] D.Warren (1997) "Automated leaf shape description for variety testing in chrysanthemums" in proc. 6th Int. Conf. Image Process and its Applicant. , Duplin, Ireland.

[5] D.Warren (1997) "Automated leaf shape description for variety testing in chrysanthemums" in proc. 6th Int. Conf. Image Process and its Applicant. , Duplin, Ireland.

[6] http://clamp.ibcas.ac.cn/

[7] T.Satiosh and T.kaneko (2000), "Automatic Recognition of Wild Flowers", ICPR, Vol.2, pp. 2507-2510.

[8] Gurpreet kaur, Himanshumonga (2012) "Classification of Biological Species Based on Leaf Architecture" IRACST- International Journal Of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555 Vol. 2, No.2.

[9] Miroslav BENCO, Robert HUDEC (2007) "Novel Method For Color Textures Features Extraction Based on GLCM" Radio engineering, Vol. 16, No. 4.

[10] www.mendeley.com/research/patternrecognition.

[11] Brendon J. Woodford, Nikola K. Kasabov and C. Howard Wearing (1997) "Fruit Image Analysis using Wavelets" in Proc. 6th Int. Conf. Image Process. And It's Applicant, Duplin, Ireland.

[12] L.Tang, L. Tian, B. L. Steward (2003) "Classification of broadleaf and grass weeds usingGabor Wavelets and an artificial neural network" Transactions of the ASAE Vol. 46(4): 1247-1254 _ 2003 American Society of Agricultural Engineers ISSN 0001-2351.

[13] Stephen Gang wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and Qiao-Liang Xiang (2007) "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network", arXiv:0707.4289vl [cs.AI].

