International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 04 | Apr 2019 www.irjet.net p-ISSN: 2395-0072

Feature Extraction Of Leaves Using Image Processing

Ritika Dasmohapatra[1], Pooja Dhere[2]

[1]*Ritika Dasmohapatra, B-Tech Student, Dept. of Electronics Engineering, VIT, Pune, Maharashtra, India*

[2]*Pooja Dhere, B-Tech Student, Dept. of Electronics Engineering, VIT, Pune, Maharashtra, India*

Abstract - *Imagine someone hiking or trekking somewhere, where he/she finds a weird leaf or any leaf he/she is not aware of. This person has always been bad in biology but would like to know more about that plant. Its main features? Is it rare? Is it protected? some parameters related to that leaf? Etc. By simply taking a picture of the leaf with a Digital Camera, he or she could feed it to the database in his computer and then get all the information regarding the leaf image through a leaf recognition application. In recent decades, digital image processing, image analysis and machine vision have been sharply developed, and they have become a very important part of Computer Vision and Artificial Intelligence and the interface between human and machine grounded theory and applied technology. These technologies have been applied widely in industries, medicine, agriculture and many more fields. The Project is elaboration of such an application in agriculture field has been attempted. Taking the photograph and recognition of leaves from photographs implies several steps, starting with image preprocessing, feature extraction, plant identification and testing and finally obtaining the results implemented in MATLAB. The main AIM of the project is providing a detailed identification to hikers, campers, etc. While a botanist could be content with a system for recording a plant species discovered in its natural habitat, to be identified.*

Key Words*: *Features, Computer Vision, Image processing, botanist, MATLAB, etc .

1.INTRODUCTION

Plants exist everywhere on Earth where we live, as well as places where we don't. Many of them carry significant information and play vital role for the development of human society. The current situation is that many plants are at the risk of extinction. So it is very necessary to set up a database regarding plant features so as to know the details regarding the plant. The human visual system can easily interpret the variations in contrast, brightness and shading in the photograph and segmenting the object from its background. Let's imagine a person taking a field trip, camping or trekking and finding a bush or a plant on the ground, he or she would like to know whether it's a weed or poisonous or any other plant but have no idea about what kind of plant it could be. With a good digital camera and a recognition program, one could get some useful information regarding the features of the leaves of the plant. There are a huge number of plant species on the Earth. Research is going on to handle such volumes of information, development of a quick and efficient classification method. In addition to the conservation aspect, recognition of plants is also necessary to utilize their medicinal properties and using them as sources of alternative energy sources like biofuel. There are several ways to recognize a plant, like flower, root, leaf, fruit etc. The technique we are using in the project is Leaf features Extraction.

1.1 BACKGROUND

Since recent decades, digital image processing, image analysis and machine vision have been widely developed, and they have become a very important part of Computer Vision and Artificial intelligence and plays a role of interface between human and machine applied technology. These technologies have been applied widely in industry and medicine, but rarely used in agriculture or natural habitats. Despite the importance of the subject of identifying plant using digital image processing, and although this has been studied for last many years, the progress seems to be quite timid. Methods are too specific. The ideal method would be able to identify any kind of plant but that has been found to be very basic methodology. Operation conditions are very strict. Many images used to develop new methods are collected under the strict conditions of lighting, angle of capture, distance between object and capture device, among others. This is a common practice and was used in early Research. However, in real world real Time applications, those conditions are almost impossible to be enforced. Thus, it is a problem that many studies never get to the point of testing and upgrading the method to deal with more realistic conditions, because this limits their scope greatly. Lack of technical knowledge about more sophisticated technical tools. The simplest solution for a problem is usually the preferable one. In the case of image processing in Computer Vision help to solve some problems by using only morphological operations that are quite mathematical, which are easy to understand as well as easily implemented. Nowadays, computer vision and pattern recognition techniques have been applied towards automated procedures of plant recognition. Digital image processing is the use of the algorithms and procedures for operations such as image enhancement, image compression, image analysis, etc. Digital images have wide

application areas including pattern recognition, computer vision, industrial automation and healthcare industries. One of the most common methods in leaf feature extraction is based on morphological features of leaf. One can easily capture the leaf image and transfer to a computer and a computer can extract features automatically by image processing techniques. The aim of the project is to develop a Leaf recognition program based on specific characteristics extracted from photography. Hence its an approach where the plant is identified based on its leaf features such as area and classification. The main purpose of this program is to use MATLAB resources. There are many advantages of using MATLAB and The result proves this thing very well.

1.2 RELATED WORK

Many research have been working and made an attempt for plant identification. Some techniques identify the plants based on plant image color histogram, edge features and its texture information. They also classify the plants as trees, shrubs and herbs using complicated classifier algorithms. The algorithm used in our project is quite simple to understand and implement too. Figure 1 shows the flow chart of the proposed algorithm[3].

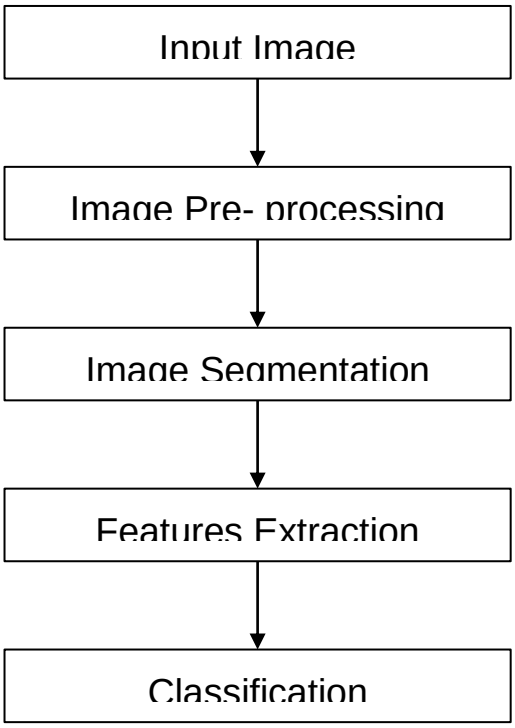

Fig. 1. Flow chart of the proposed Algorithm

2. STAGES OF THE SYSTEM:

Leaf recognition applications are widely used by botanist. However, the program developed by them in that case is specific for only one task and is of no use in more general applications. Actually, these simple techniques focus only on a few features, and are not efficient in a more general purpose. Consequently more general image classification methods are used, because it is a widespread topic, and there are a lot of well-known features such as colour Histogram, HOG (Histogram of oriented gradient), Shape descriptors, OCR (Optical Character Recognition). In order to focus on the main structure of the program, the MATLAB implementation, the database retrieving and specific feature creation, we will use the built functions available in MATLAB for Digital Image Processing. The final program provides a algorithm for segmentation as seen in Figure 2. In addition, like in most of the image processing and recognition programs, firstly a database of plant or leaf picture has to be done, then a learning method to extract the features for the database, and finally matching method to retrieve the best match from the database been provided. Several additional small programs have been applied to gather information for results.

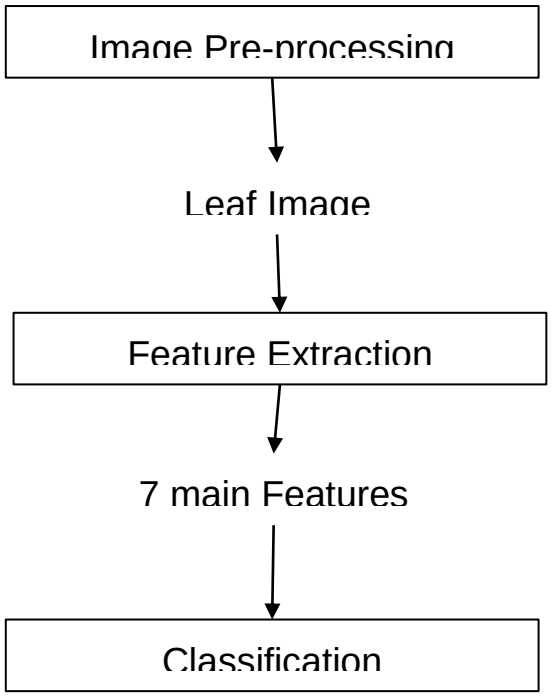

Fig. 2. Main stages of System.

3. IMPORTANCE

In leaves recognition research, a lot has been done about general features detection, extraction and recognition between different classes of objects. In case of specific domain recognition, taking into account the unique characteristics that belong to this category, improves the performance of the system. Biologists also emphasize the

importance of leaves; indeed their size, their shape, their disposition can vary very much and be a good mean for differentiating similar blooms[5]. In case of a species domain recognition, taking into account the unique characteristics of the tree/plant/leaf that belong to this category, improves the performance of the system. Despite the high technical aspect of this project, dealing with leaves, gives it a biological notation[5]. Some basic knowledge about leaves have to be learned and concepts about how the biologists themselves recognize leaves has to be studied. Biologists receive a large number of requests to identify plants for people, many species of plants look very similar on their leaves, and botanists will turn to identifying the species based on their structure or other morphologies. There are three main parts to a leaf:

1. The baseis the part which is the point at which the leaf is joined to the stem[6].

2. The stalk or petiole is the thin section joining the base to the lamina - it is generally cylindrical or semicircular in form[6].

3. The lamina or leaf blade is the wide part of the leaf[6].

4. SYSTEM OVERVIEW:

Identifying a landscape or garden plant requires recognizing the plant by one or more characteristics, such as size, form, leaf shape, leaf colour, flower colour, odor, etc., and linking that recognition with a name, either a common name or scientific name. Accurate identification of a cultivated plant can be very helpful in knowing how it grows e.g., size shape, texture, etc. as well as how to care and protect it from pests, damages and diseases. First let's look at some common characteristics of leaves of plants that are useful in identifying the features and characteristics of the plants. Now if the same was in a botany class dealing with plant systematics, the field of study concerned with identification, naming, classification, and evolution of plants, we would spend a good deal of time on the reproductive parts of plants, i.e., mostly the various parts of the flowers, i.e., ovary, stigma, etc. Structural similarity of reproductive parts is an important means by which plants are categorized, grouped, named, and hence identified[6]. Pattern recognition is a very vital field within computer vision, and the aim of pattern recognition/classification is to classify or recognize the patterns based on extracted features from them. The pattern recognition technique involves three steps:

(1) Image Pre-processing
(2) Feature Extraction
(3) Classification.

4.1. IMAGE PRE- PROCESSING

Image pre processing involves some method to convert the original image into some informative images. Before the operations, some of the leaf images are rotated manually for helping the program to arrange leaf apex direction to the right side. Afterwards, automatic preprocessing techniques are applied to the leaf images that is input to the system.

4.1.1. CONVERTING RGB TO BINARY IMAGE

The leaf image is captured by scanners or digital cameras. Since we have not found any digitizing device to save the image in a lossless compression format, the image format here is JPEG, png or JPG. There is no restriction on the direction of leaves when photoing as well as while capturing the image. An RGB image is firstly converted into a grayscale image. The formula to convert image from RGB to grayscale is:

$$gray = 0.2989 \cdot R + 0.5870 \cdot G + 0.1140 \cdot B$$

where R, G, B correspond to the color of the pixel Red, Green and Blue respectively. The level to convert grayscale into binary image is determined according to the RGB histogram[8] as shown in the figure below.

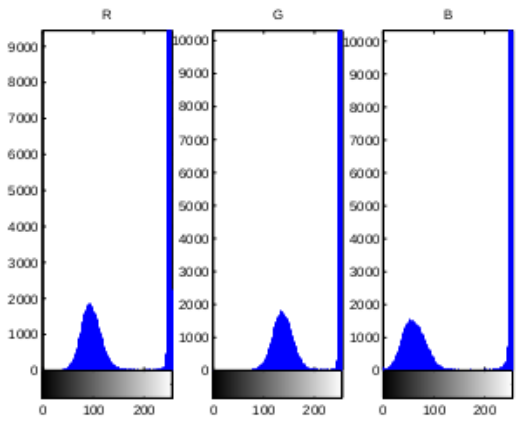

Fig. 3. RGB Histogram

4.1.2. IMAGE SEGMENTATION

When mentioning the leaf shape, the first thing appears in your mind might be the shape, size and margin of a leaf. Convolving the image with a Laplacian filter of following 3 × 3 spatial mask we can have the margin of the leaf image. The figure below shows the pre processing techniques.

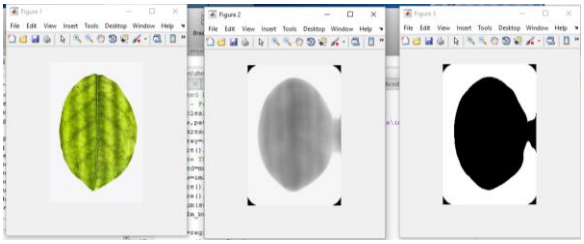

Fig. 4. Output of Image Pre-processing.

4.2. FEATURE EXTRACTION

After Image pre-processing techniques, in pattern recognition, the important and essential task is to measure the properties of an object because objects have to be detected by computed properties. In the feature extraction step, the task is to describe the regions based on chosen representation, e.g. a region may be represented by its boundary and its boundary is described by its properties (features) such as color, shape and texture, etc. There are two types of representations, an external representation and internal representation. An external representation is chosen on the basis of shape characteristics. An internal representation is selected on the basis of regional properties such as color and texture. Sometimes the data is used directly to obtain the descriptors such as in determining the texture of a region, the aim of description is to quantify a representation of an object[12]. This implies, one can compute results based on their properties such length, width, area and so on.

- Area: Area represents number of pixels in the leaf region. Binary form of our leaf image has black background and white leaf. In this image, number of white pixels represents the area of the leaf.
- Major Axis: Major axis is denoted as a line, which lies between apex and base of the leaf.
- Minor Axis: Minor axis of the ellipse that has the same normalized second central moments as the leaf region.
- Perimeter: Perimeter is the distance around the boundary of leaf region.
- Convex Hull: Convex hull represents the smallest convex polygon that encapsulates the leaf region.
- Minor Axis Length Ratio of Major Axis Length: This feature is denoted as ratio of minor axis length to major axis length. It is reverse of the aspect ratio that is used in the literature.

5. RESULTS:

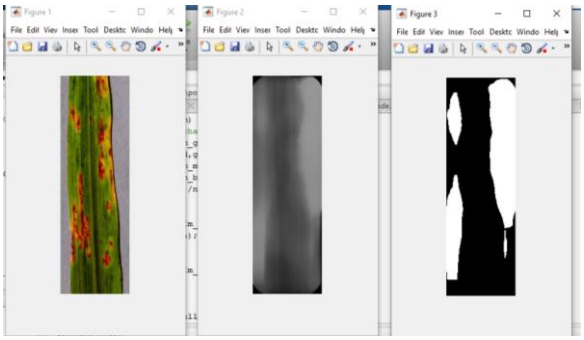

Fig. 5. Image Pre processing of a leaf.

Ar	premiter	convex_hull	major	minor	major_minor
7869	489	8225	245.54	42.3	5.8047

Fig. 6. Feature Extracted.

6. CONCLUSION

The paper discusses about the simple technique to extract features from leaves of plants by just capturing the images of that specific leaf. Firstly the pre-processing techniques are applied to the leaf image and then the features are extracted. So, Image processing in Computer Vision plays an important role in this project. Pattern recognition is a very important field within computer vision, and the aim of pattern recognition/classification is to classify or recognize the patterns based on extracted features from them. Thus this is a very simple method applied using a simple and advantageous platform MATLAB.

REFERENCES

[1] Z. Wang, Z. Chi, and D. Feng. Shape based leaf image retrieval. IEEE P-Vis Image Sign. 150:34–43, 2003.

[2] M. D. Dassanayake, editor. A revised handbook to the flora of Ceylon, volume 4. CRC Press, Boca Raton, 2003

[3] S. Wu, F. Bao, E. Xu, Y. Wang, Y. Chang, and Q. Xiang. A leaf recognition algorithm for plant classification using probabilistic neural network. In 7th IEEE International Symposium on Signal Processing and Information Technology, Cairo, Egypt, 2007.

[4] H. Fu, Z. Chi, D. Feng, and J. Song. Machine learning techniques for ontology-based leaf classification. In 8th IEEE International Conference on Control, Automation, Robotics and Vision, Kunming, China, 2004.

[5] R. C. Gonzalez and R. E. Woods. Digital image

processing, 2nd edition. Prentice Hall, Upper Saddle River, NJ, 2004.

[6] L. K. Senarathne. A checklist of the flowering plants of Sri Lanka. National Science Foundation Publishers, Sri Lanka, 2001.

[7] Otsu, N. A threshold selection method from gray-level histograms. IEEE T. Syst. Man. Cyb. 9:62–66, 1979.

[8] R. C. Gonzalez and R. E. Woods. Digital image processing, 2nd edition. Prentice Hall, Upper Saddle River, NJ, 2004.

[9] P.-E. Danielsson and O. Seger. Generalized and separable Sobel operators. In H. Freeman, editor, Machine vision for three–dimensional scenes, pages 347-380. Academic Press, CA, 1990.

[10] D. S. Huang. The local minima free condition of feedforward neural networks for outer-supervised learning. IEEE T. Syst. Man. Cy. B. 28:477–480, 1998.

[11] B. V. Dasarathy. Nearest neighbor NN norms: NN pattern classification techniques. IEEE Computer Society Press, Washington, DC, 1991.

[12] D. Warren, "Automated leaf shape description for variety testing in chrysanthemums," in Proceedings of IEEE 6th International Conference Image Processing and Its Applications, 1997.

[13] D. S. Huang. Systematic theory of neural networks for pattern recognition. Publishing House of Electronic Industry, Beijing, 1996.

[14] J.-X. Du, X.-F. Wang, and G.-J. Zhang, "Leaf shape based plant species recognition," Applied Mathematics and Computation, vol. 185, 2012

[15] M. T. Hagan, H. B. Demut, and M. H. Beale, Neural Network Design, 2002.

[16] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, Digital Image Processing Using MATLAB. Prentice Hall, 2004.

[17] Math Works Inc. Matlab R2012a documentation, 2014.http://www.mathworks.com/help/index.html, accessed May 2013