

DESIGN AND DEVELOPMENT OF AN IMAGE BASED PLANT IDENTIFICATION SYSTEM USING LEAF

KomalAsrani and Renu Jain

Head of Department(Information Technology),B.B.D.E.C., Lucknow
Head of Department(Computer Science Engineering), U.I.E.T. Kanpur

ABSTRACT

Because of huge diversity of plants existing on earth and large inter-species similarity, the manual process of identification of plants is difficult and at times, the results generated may be confusing. So, it is necessary to automate the process of identification of plants to generate faster and accurate results. Among various parts of plant, leaf is easily available in all seasons and can be easily scanned for understanding the shape details. So, in this paper, we have worked on plant identification using leaf image. In the process of identification of plants, the key challenge identified is to keep the size of feature vector reasonable and still achieve accurate and effective results. As the results are to be interpreted by the user, it was necessary to understand the theories behind human perception to reduce the semantic gap. We have designed and developed a tool 'SbLRS: Shape based Leaf Recognition System' which is a three-staged system that performs identification of plant using leaf shape. Our system can be helpful for following users- common layman person, as replacement for biologists in remote locations, as teaching aid for teachers and students, for a farmer to identify the plant species and to assist biologists in identification process. The parameters important for identifying leaf image according to human perception are identified and defined at three different levels. On the basis of results generated, the effectiveness of SbLRS is compared with existing contour based methods in terms of recall, precision. It was observed that our tool SbLRS showed satisfactory results for identification of plant using leaf at small feature vector size and simpler computations.

1. INTRODUCTION

Content based image retrieval(CBIR) is a rapidly expanding area which has wide applications in the field of matching images and image retrieval. These applications include utilization of content based image retrieval in medical diagnosis, handwriting recognition, museums, aquatic life, person's identification, military, retail catalogue and in plant identification. Here, in this paper, we are focusing our discussions on plants. Looking into the biodiversity, there are approximately above 400,000 species of plants existing on earth. However, due to a large number of plants and their close similarity with various species, the botanists sometimes ever face difficulty in the correct identification of a plant. Also, the presence of botanists cannot be ensured everywhere. Hence, there is a need for automating the process of plant identification. Various keyword based search engines are available for plant identification which includes My Nature Tree Guide, Tree Book, Botany Buddy, Plant Finder, Bachman's Landscaping Plant Library, Native Plant Information network (NPIN), Nature Gate and E-Flora. These engines use keywords for generating information about the plant like the nature of soil required for its growth or the typical characteristics of the plant etc and search the same in the database to extract the matching results. These engines can be considered a solution for remote locations where the availability of the botanists is difficult. But this keyword based procedure of matching and identification may require involvement of the user. And if the user does not have all the required details to define the minute characteristics of plant, then the keyword based approach would not provide effective results. So, a need for automating the process of plant identification using images is felt. For automating the process of plant identification, various parts such as root, bark,

leaves, fruit and flower are available to assist the process. Among these, leaves are available in all seasons and can be easily captured as a 2D object for studying its characteristics. However, leaves have various features like color, shape and texture. We know that the color of most of the leaves is generally green, so it cannot be considered as a good parameter for differentiating among different leaves. Leaves contain very fine textured surfaces, which can be captured by very high resolution scanner. However, shape of the leaf is the parameter which can be effectively captured by scanner of reasonable resolution and is also easily understood by human beings. Thus, in the paper, we would be focusing our discussion on plant identification using leaf shape.

2. LITERATURE REVIEW

Extensive work has been done in the field of plant identification using shapes of leaves. Some of the earlier works focused on the basic morphological characteristics of a leaf shape which included area, perimeter, eccentricity, maximum length and maximum width. But, it was observed that these basic morphological parameters are helpful for the broader classification of leaves and do not provide sufficient information about defining the minute details of the shape of a leaf. To reduce the search space and improve the retrieval efficiency, the concept of clustering and dimension reduction was initiated in 1997, where the authors Abbasi, Mokhtarian, & Kittle, 1997 used the curvature scale space technique and the k-NN classifier to chrysanthemum leaves. Neto, Meyer, Jones, & Samal, 2006 used elliptic Fourier harmonic function for defining the leaf shape. These Fourier values were reduced by using Principal component analysis. The clustering and dimension reduction approach caused reduction in the size of the feature vector but in turn resulted in the loss of minute details of the leaf shape, hence affecting the retrieval efficiency. Use of statistical discriminant analysis was done by Timmermans & Hulzebosch, 1996 along with the color based clustering and neural network was used for classification of a flowered plant and a cactus plant. Efforts for approximation of a leaf shape were done by Belongie, Malik, & Puzicha, 2002. He computed 2D histogram from shape content by computing inner distances and angles between sample points of leaf margin. Yahiaoui, 2006 used two directional fragment histogram to compute directions and relative lengths in a succession of elementary fragments of contour. The concept of histogram was helpful in reducing the representation of feature vector but it caused a major loss of individual data by merging data into bins. Im, Nishida, & Kunii, 1998 proposed a hierarchical technique for representing leaf shapes by using polygonal approximation of leaf contour. Experiments were also conducted by using neural networks. Saitoh, Aoki, & Kaneko, 2004 used automatic method for recognizing wild flowers by taking frontal image of a flower and a leaf using neural network. Qi & Yang, 2003 studied extraction of saw tooth features of a leaf using support vector machine. Wu, 2002 used probabilistic neural network for classification of leaf on the basis of leaf features. These features were reduced to five values using PCA for further processing. Though, elaborate work had been reported on different individual aspects of leaf shape, but a complete detailed description of the leaf (the type of leaf shape, the pattern of saw tooth, the nature of leaf apex and base) which maintains a balance between the feature vector size and retrieval accuracy is not exactly available.

During the study of various works, it was observed that most of the systems focused on computational logic involved in image representation. But, it was actually the perception of human vision that is important, which has not been effectively considered in these systems. Thus the main challenge identified was the semantic gap which occurs because of the difference in the representation of the digital image and that of human perception. So, it was necessary to understand the causes of the existing semantic gap and identify solutions to reduce the same. For this, human perception for defining an image was closely understood. As shape is considered as an important visual feature and is closest to human perception, various shape representation approaches in context to image retrieval were studied. According to classification as proposed by

Luciano, Coasta, & Marcond, 2004, a shape can be defined in the form of two categories: Contour based and Region based approach. In contour based approach, the outer boundaries of an image are considered whereas in region based approach, besides the boundary details, the interior details of the shape are also considered. Comparing contour based approach and region based approach, it is found that contour based approach is the preferred approach. Contour based representation is compact and provides an overall preview of image. Moreover, when dealing with huge and scaling database, a compact representation ensures quick and efficient computation whereas region based approach provided very fine details of the shape.

3. FEATURE VECTOR GENERATION AND IMPLEMENTATION DETAILS

Process of identification of important features of leaf: Initially, we worked with various existing contour based approaches to understand the extent to which they represented the leaf shape. The feature vector of each of the contour methods was analyzed for defining the leaf boundary and results were compared with other existing contour based approaches. The advantages and disadvantages of these approaches were understood in defining the leaf shape in terms of efficiency and size of feature vector. Two important issues were identified for retrieval: efficiency of retrieval results and time for generating results. These issues, directly or indirectly focused on the size of the feature vector, because if the size of feature vector is small, the time required for representation of image and comparison with database would be very much reduced. To regularize the size of feature vector and search space, we tried to perform clustering on the basis of eccentricity and then utilizing additional contour based approaches, but the results were not satisfactory as the finer details of the leaf shape like leaf margin, leaf curvature, leaf width proportionality etc. were not captured. Hence, it was concluded that additional methods should be explored which are capable to capture the required finer details of leaves. Here, in this paper, we propose a three-staged plant identification system referred as 'Shape based Leaf Recognition System: SbLRS'. This system takes the query in the form of leaf image from the user, processes the image by its shape and returns the best match or the closest match corresponding to the query. SbLRS software is a stand-alone application implemented in Java on Windows platform. It recognizes leaf images on the basis of shape parameters. The application was developed using NetBeans 7.3.1 and Oracle 10g. The user will scan the leaf against a plain background and then by using various contour based techniques, SbLRS provides closest match corresponding to the query image. Our system helps in identification of the plant.

4. DATABASES CONSIDERED FOR WORK

Two databases have been considered for our work. They are 'UC Irvine Machine Learning Repository: Leaf'. This dataset consists of a collection of shape and texture features for 40 different species. An average number of 10 leaf species are defined for each plant. A total number of 443 images are available. The second dataset are created by us and referred as 'User-created Database'. Only simple leaves have been considered for our work and testing. These leaves in User-created database are collected from various sources like NBRI, CIMAP and other herbariums and botanical garden. Besides these, leaves are also obtained from online dataset of LeafSnap. Approximate 10-15 leaves are collected from each plant and in all, there are 1000 leaves in User-created Database. Because we have considered multiple leaves from each different species, approximately 50% of the leaves of individual species are used for training and remaining for testing the output results.

5. EVALUATING THE FEATURES FOR GENERATING FEATURE VECTOR

A rich terminology has been used by biologists for describing leaf characteristics. This terminology is based upon different leaf patterns, shape and their arrangement. Further,

these specific leaves were further categorized on the basis of apex and base angle. Initially, the broad shape of the leaf is estimated by the basic geometrical interpretation and overall visual characteristics defining the leaf boundary. On the basis of broader aspect of leaf contour, the leaves are classified as linear, lanceolate, long, elliptical, ovate, cordate, oblanceolate, obovate, falcate, auriculate, hastate, detoid, reinform, peltate, obcordate, spatulate and cuneate. Looking into the shape characteristics of leaf shapes, it is observed that the classification is based on the overall proportionality of leaf height and width, the extent of broadness at different heights of leaf and the proportionality of curvature variation. The leaves are classified as Entire, Toothed, Lobed and Parted/Cleft on the basis of smoothness of leaf margin.

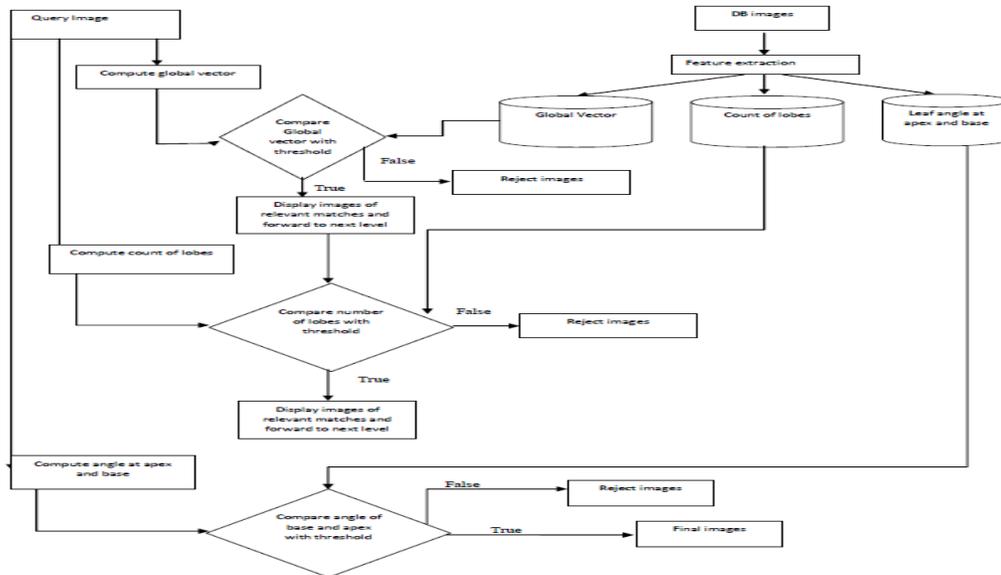


Figure 1. Architecture of Shape based Leaf Recognition System

On the basis of leaf apex angle, the classification is Acuminate, Acute, Obtuse and Truncate. Lastly, on the basis of leaf base angle, the leaves are classified as Cuneate, Acute, Cordate and Obtuse. The details of variation of leaf shape aspects as per the discussed classification are considered while designing the plant identification system using leaf.

6. ARCHITECTURE OF SBLRS

SbLRS is a three staged plant identification system which identifies plant on the basis of leaf shape. The details of the three stages of identification are discussed below.

First Stage of Identification

The first stage of SbLRS captures the complete details of the leaf boundary and thus these parameters are referred as ‘Global Vector’. The ‘Global Vector’ includes area, perimeter, aspect ratio, image length, image width, image top width, image bottom width and Quad Centroid Distance Variation (QCDV)[11]. The varying parameters are assigned weightage depending upon the importance in defining the shape of the leaf. For assigning weights, matrix is constructed based on the identification of critical factors. The process for identifying critical factors in context to plant identification is defined as follows:

- How much important is the variable for identification process?
- What is the datatype of the variable?

- What is the impact and scope of variable?
- What levels of computations are involved in calculating the value of the variable?
- Whether the variable is an independent or dependent variable?

Based on the above discussion, the critical factors affecting the retrieval process are dependent upon the size and nature of the feature vector and how much representative it is of the leaf shape. Accordingly, the critical factors identified in plant identification using leaf shape are as follows:

- Computational Process: It defines the complexity in computing the values and the number of variables involved in calculating the values (simple calculation means more weightage).
- Datatype of variable value: It identifies the data type of the variable: integer or decimal (integer means more weightage).
- Scope of variable: It defines the dependent nature of the variable: independent or dependent (independent means more weightage).
- Number of dependent variables: It identifies the number of variables dependent on a specific variable (more dependency of other variables on specific variable means more weightage).
- Single valued or vector variable: It determines the number of values required to define the variable (single-valued means more weightage).

Once the variables are identified, a matrix is constructed based on following rules:

- Identify critical success factors.
- Assign score to each critical factor.
- Identify the optimum value for each criteria.
- Assign each element with respective weight.
- Add values to generate the score of each success factor.

Table 1. Matrix construction for identifying weightage of parameters

Parameters Criteria	Height	Width	Aspect Ratio	Area	Perimeter	Top Width	Bottom Width	QCD V
Computation Process(5)	5	5	5	3	3	5	5	2
Datatype of variable (2)	7	7	3	3	3	7	7	3
Scope of variable(3)	6	6	4	6	6	4	4	4
No.of dependent variables(4)	5	2	0	2	0	0	0	0
Single valued or vector(1)	7	7	7	7	7	7	7	3
Summed Values	77	65	43	43	39	51	51	28
Normalized values	0.19	0.16	0.11	0.11	0.1	0.13	0.13	0.07

Table 2. Final weightage assigned to parameters in SBLRS

Parameter	Weightage
Height	0.19
Width	0.16
Aspect Ratio	0.11
Area	0.11
Perimeter	0.1
Top width	0.13
Bottom width	0.13
QCDV	0.07

Hence, depending on their contribution for defining the shape of leaf, these parameters are assigned relevant weightages. The results showed that the following leaves were classified to be identical inspite of their differences in shapes: Ovate and elliptical, linear leaves with and without lobes, rein form and peltate and leaves with entire and toothed margin features.

Second stage of identification

The second stage defines the smoothness of the shape and count the number of lobes in the shape. For computing the number of lobes, first the distance of boundary points from centroid is computed. The difference between successive distances is computed. The successive positive and negative values of distances are identified and standard deviation is computed to identify the lobe count. The results showed that leaves with different margin features like entire, toothed, lobed and cleft were clearly distinguished.

Input: Leaf image

Output: Identifies tooth and sinus in leaf boundary and computes count of lobes in leaf.

Functionality: Helps in identifying the number of lobes in leaf.

Algorithm details for leaf_margin_lobes:

1. Compute centroid of image and define it as (x_c, y_c) .
 2. Compute distances of boundary from centroid and define them as $(d_1, d_2, d_3, \dots, d_n)$.
 3. Compute differences of successive distances for all distances as $(d_1 - d_2, d_2 - d_3, \dots, d_{n-1} - d_n)$.
- These differences are defined as $(dd_{12}, dd_{23}, dd_{34}, \dots, dd_{n-1})$.*
4. Identify positive and negative values in $(dd_{12}, dd_{23}, dd_{34}, \dots, dd_{n-1})$. Sum up consecutive positive and negative values in $(dd_{12}, dd_{23}, dd_{34}, \dots, dd_{n-1})$, and define them as $(ds_1, ds_2, ds_3, \dots, ds_p)$. If there are no negative values, no lobes exist in the boundary of the leaf and exit.
 5. Compute standard deviation for pair of positive and negative values and define them as $(std_1, std_2, std_3, \dots, std_{p-1})$.
 6. Compare value of each std with threshold value . If condition for threshold is satisfied, consecutive pair of positive and negative values of ds defines a lobe.
 7. Compute such combination of positive and negative ds and calculate the total number of lobes.

Third stage of identification

Once the lobes of the shape are captured, leaf angle at the tip and base are used for refining the results obtained from second stage. For computing apex and base angle, leaf top and bottom width is calculated using leaf height. The apex and base point of the leaf are identified and the cosine of the angle gives the required value. Unlike the first stage, these leaves were clearly distinguished in this stage: Ovate and elliptical, Reinform and peltate.

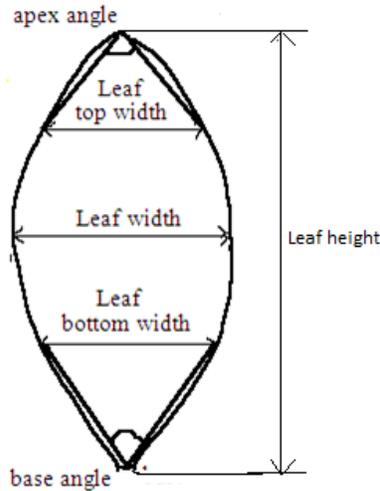


Figure 2. Length and width of leaf

Input: Leaf image

Output: Computation of apex and base angle of leaf.

Functionality: Leaf angle at base and apex help in identification of leaf type.

Algorithm details for leaf_angle:

1. Compute leaf width for top and bottom segment of leaf using Leaf Width computation algorithm.

2. For each apex/ base,

i) Identify three points that constitute apex/base point, left top/bottom extreme and right top/bottom extreme from boundary coordinates.

ii) Compute angle θ using

$$\theta = \cos^{-1}\left(\frac{\overline{AB} \cdot \overline{BC}}{\|\overline{AB}\| \cdot \|\overline{BC}\|}\right)$$

where $\overline{AB} = B - A$ and $\overline{BC} = C - B$;

We have implemented the Leaf Recognition System in NetBeans IDE using Java as frontend and Oracle 10g as backend. Two interfaces have been provided: 'User' and 'Admin'. 'Admin' button provides interface for the administrator (as shown in Figure 3) where he can load images to the database, view images from the database and generate feature vectors corresponding to a leaf using different boundary based approaches.

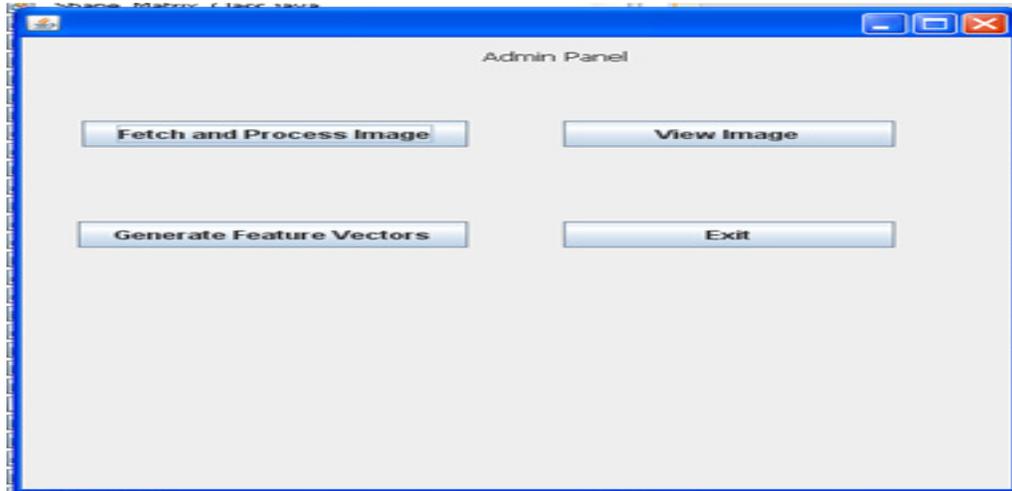


Figure 3. Snapshot for Admin interface

Various boundary based approaches for which feature vectors can be generated by the administrator. The feature vectors, corresponding to each approach are stored in different tables of database. The administrator interface for generating feature vectors using various contour based approaches is shown in Figure 4.

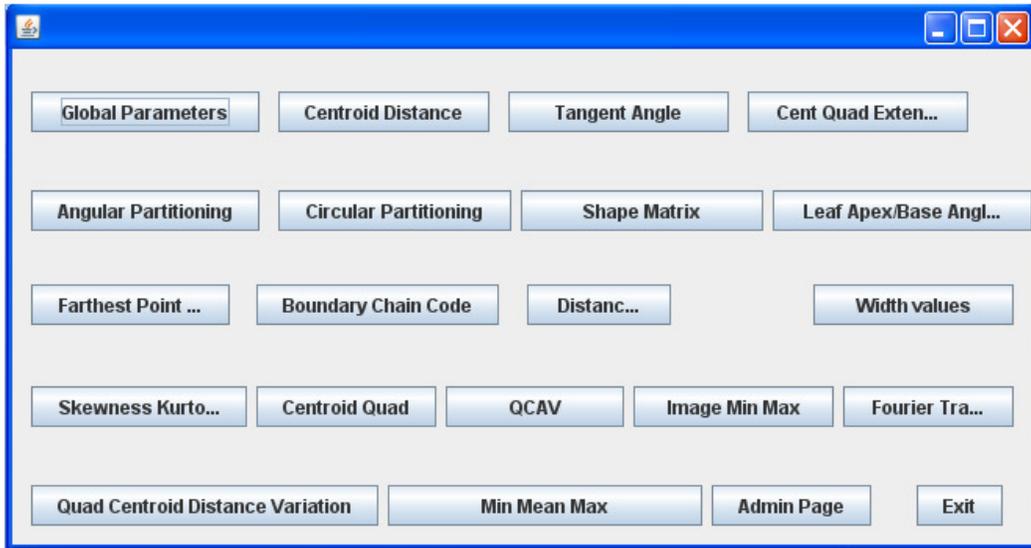


Figure 4. Snapshot for generating feature vectors using contour based approaches

The 'User' interface provides interface where the user can search for the query leaf using our approach SbLRS as shown in Figure 5. The feature vector generated using SbLRS Approach for the query leaf is shown in Table 3.

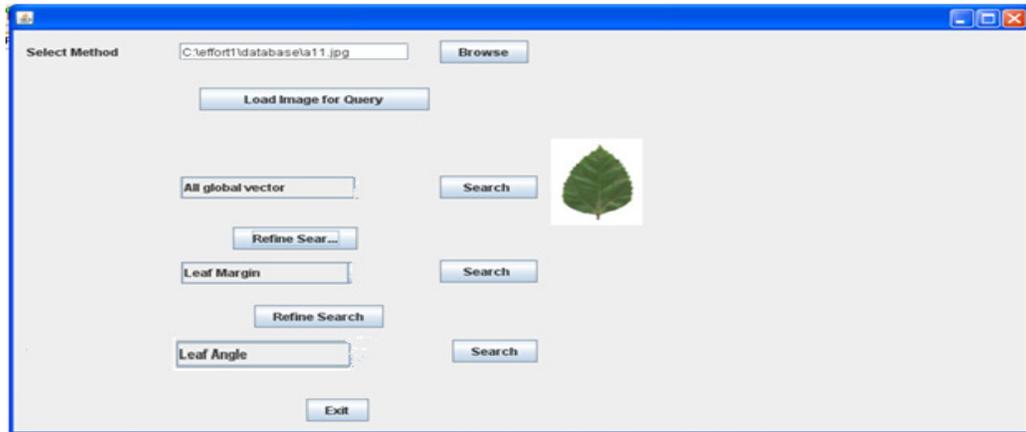


Figure 5. Snapshot for querying leaf using SbLRS

 <p data-bbox="451 1241 651 1272">Input Leaf Image</p>	<p data-bbox="829 804 1110 835">eccentricity=0.64705884</p> <p data-bbox="829 835 971 867">area=3224.0</p> <p data-bbox="829 867 1073 898">perimeter=318.51038</p> <p data-bbox="829 898 1089 930">Value in q1= 5.195099</p> <p data-bbox="829 930 1105 961">Value in q2= 11.334496</p> <p data-bbox="829 961 1105 993">Value in q3= 4.9252552</p> <p data-bbox="829 993 1105 1024">Value in q4= 10.753253</p> <p data-bbox="829 1024 943 1056">length=85</p> <p data-bbox="829 1056 984 1087">top width=29</p> <p data-bbox="829 1087 1032 1119">middle width= 55</p> <p data-bbox="829 1119 1024 1150">bottom width=43</p> <p data-bbox="829 1150 1198 1182">apex angle=s 54.7714080567232</p> <p data-bbox="829 1182 1198 1213">base angle= 74.97634826875132</p> <p data-bbox="829 1213 911 1245">lobe=0</p>
--	--

Table 3. Feature vectors generated for query leaf using SbLRS

Results

Screenshots of results generated using SbLRS are presented in Figure 6.



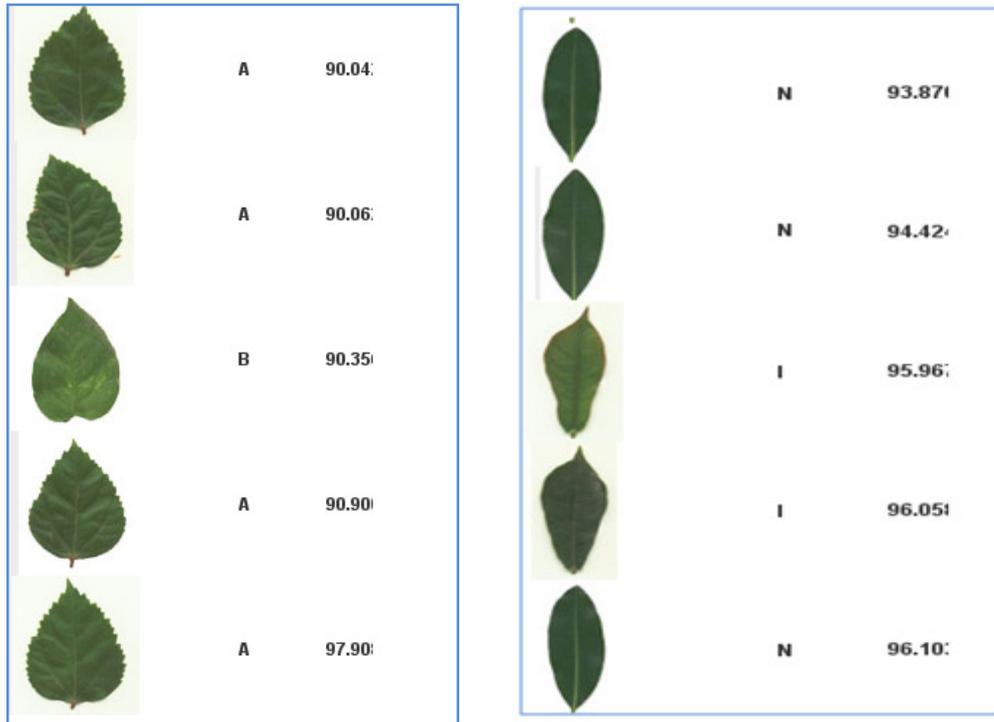


Figure 6. Screen shot of results generated using SbLRS

Table 4 presents the classification accuracy, recall and precision for UC Irvine Machine Learning Repository: LEAF Dataset. The results have been presented according to leaf type. Relevant categories of leaves have been identified from the database and calculations have been done using confusion matrix. Table 5 presents the classification accuracy, recall and precision for User-created Database. The results have been presented according to leaf type. Relevant categories of leaves have been identified from the database and calculations have been done using confusion matrix.

Table 4. Classification accuracy, recall and precision for UC Irvine Machine Learning Repository: LEAF Dataset

Leaf type	Leaf name	Classification accuracy	Recall	Precision
Linear	Nerium oleander	98%	85%	75%
Elliptical	Quercus suber	98%	83%	78%
Ovate	Celtis sp.	91%	81%	78%
Cordate	Tiliatomentosa	85%	79%	74%
Deltoid	Populus nigra	95%	84%	73%
Ob lanceolate	Salix atrocinera	89%	79%	71%
Peltate	Alnus sp.	94%	84%	76%
Lanceolate	Magnolia soulangeana	92%	82%	79%
Spatulate	Primula vulgaris	88%	80%	74%

Table 5. Classification accuracy, recall and precision for User-created database

Leaf type	Leaf name	Classification accuracy	Recall	Precision
Linear	Fern Pine	96%	85%	76%
Elliptical	Apple	91%	84%	77%
Ovate	Ashok Mango	92%	85%	75%
Cordate	Bean	86%	80%	74%
Deltoid	Catawba	94%	85%	76%
Obovate	Cashew Mahua	91%	85%	76%
Spatulate	Black Jack Oak	89%	84%	77%
Reinform	Ginkgo Biloba	92%	86%	80%
Oblong	Arjun	89%	83%	79%

Table 6 presents comparison of feature vector size for contour based approaches and our approach SbLRS. These values have been computed on the basis of the results generated by using multiple leaves for both the databases.

Table 6. Comparison of feature vector size for contour based approaches and our approach SbLRS

Method Name	Size of feature vector
Centroid Distance Approach	Depends on number of boundary points (200-250)
Tangential Angle Approach	Depends on number of boundary points(200-250)
Angular Partitioning	4
Circular Partitioning	4
Shape Matrix	Depends on size of grid
Farthest Point Distance	Depends on number of boundary points(200-250)
Fourier Approach	Twice the number of boundary points(400-500)
Boundary Chain Code	Depends on number of boundary points(200-250)
SbLRS Approach	14

Figure 7,8,9 presents comparison of classification accuracy, recall and precision for various contour based approaches and SbLRS for UC Irvine Machine Learning Repository: LEAF.

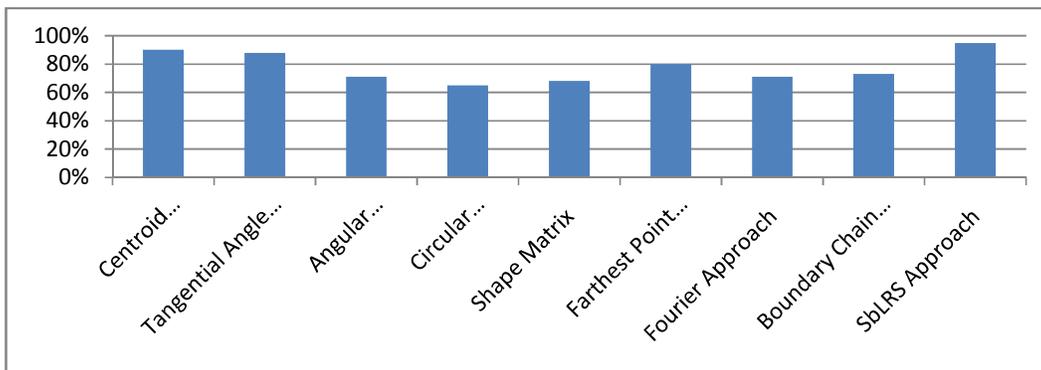


Figure 7. Comparison of classification accuracy for contour based approaches and SbLRS for UC IrvineMachine Learning Repository: LEAF

Concluding remarks and prospects

All the testing was done on SbLRS which was developed as a tool for plant identification based on a leaf shape. On the basis of results, average classification accuracy was found to be 95.14% on UC Irvine Machine Learning Repository: LEAF and 94% on SbLRS dataset. SbLRS tool provided results with relevant leaf images with their degree of matching corresponding to the query leaf. High classification

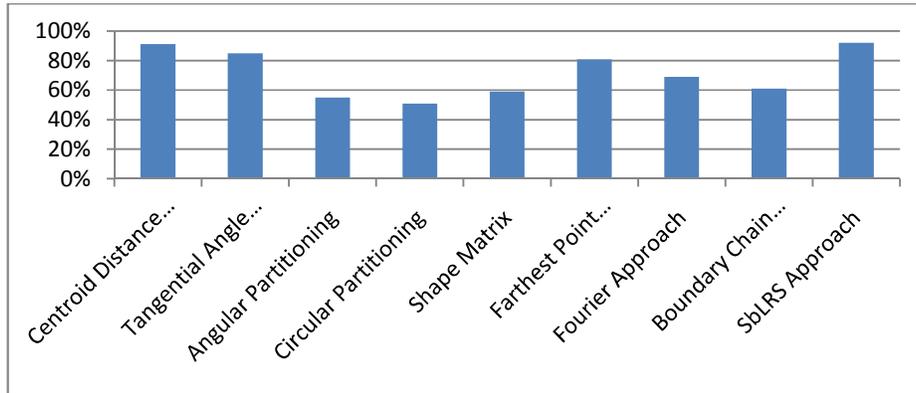


Figure 8. Comparison of recall for contour based approaches and SbLRS for UC IrvineMachine Learning Repository: LEAF

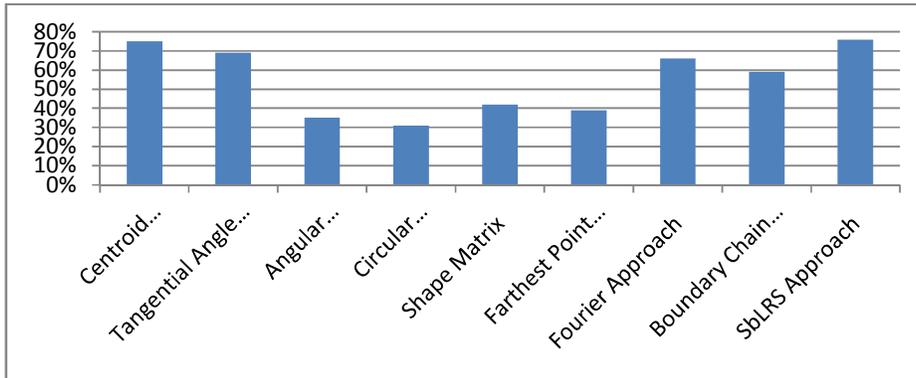


Figure 9. Comparison of precision for contour based approaches and SbLRS for UC IrvineMachine Learning Repository: LEAF

accuracy was achieved because all the parameters important according to human perception for recognition of a leaf were considered in SbLRS. Low computational load was achieved because all the feature vectors of database images were pre-generated and the feature vector generated for query image required minimum calculations. Keeping in view the small feature vector, the classification accuracy of the proposed tool was highly satisfactory. In case, when the leaf of the same plant was not available in the database, then the tool returned to the nearest matches available.

The average error rate was 4.86%. This was partially due to loss of data involved in pre-processing stage involved scaling, thresholding, edge detection, denoising and finally converting to binary image. As far the structure of the leaf is concerned, the base point from where the petiole extends and the tip of the leaf was difficult to be identified without human intervention.

Also, there exists large inter species similarity in different plants which makes the search process too difficult. Complexity also arises because over a period of time the leaves of the same plant show variability in shape. Depending upon the plant age, the shape and dimension of the leaf is also variable. A young plant has small and tender leaves, whereas a full grown plant has bigger and mature leaves. It is also observed that the shape of leaves is highly affected by the intensity of sunlight and climatic conditions. It was also observed that the size of the leaf is highly variable across plant species. Some plants have very large or very small leaves. All these factors, directly or indirectly, affect the retrieval effectiveness. For dealing with these issues, we have tried to consider multiple leaves of the same plant. But still, it was not possible to consider leaves of all stages.

Based on the study and extensive analysis in the area of plant identification, it was observed that it is impossible to attain 100% results generated for an image based search engine only using some leaf features. But efforts can be incorporated to improve the effectiveness by working on the mentioned areas: In SbLRS, we have not considered very fine margin details along the contour which assist in providing minute angular details of the leaf and can be very useful in identifying the correct match. Secondly, veins and venation are present inside the leaf boundary. The network pattern provides important information about the details of the primary, secondary and tertiary vein structure of a leaf. Moreover, leaves are attached to the stem in different patterns. So, instead of considering a single leaf, a small branch consisting of multiple leaves can be considered for representing a plant which may somehow help in improving the results.

REFERENCES

1. Abbasi, S., Mokhtarian, F., & Kittle, J. (1997). Reliable classification of chrysanthemum leaves through curvature scale space. *International Conference on Scale-Space Theory in Computer Vision*, (pp. 284-295). Utrecht.
2. Neto, J. C., Meyer, G. E., Jones, D. D., & Samal, A. K. (2013). Plant species identification using elliptic Fourier leaf shape analysis. Hal-00726785 version 1 .
3. Timmermans, A., & Hulzebosch, A. (1996). Computer vision system for online sorting of pot plants using an artificial neural network classifier. *Computers and Electronics in Agriculture*, Vol.15 , 41-55.
4. Belongie, S., Malik, J., & Puzicha, J. (2002). Shape MAtching and Object Recognition using Shape Contexts. *IEEE Trans. on Pattern Analysis and Machine Intelligence* Vol. 24, no. 4 , 509-522.
5. Mouine, S., Yahiaoui, I., & Blonde, A. V. (2013). A shape-based approach for leaf classification using multiscale triangular representation. *ACM International Conference on Multimedia Retrieval*. Dallas, United States.
6. Im, C., Nishida, H., & Kunii, T. L. (1998). Recognizing plant species by leaf shapes—a case study of the acer family. *International Conference on Pattern Recognition*, (pp. 1171-1173). Brisbane.
7. Saitoh, T., Aoki, K., & Kaneko, T. (2004). Automatic recognition of blooming flowers. *International Conference on Pattern Recognition*, vol. 1, (pp. 27-30). Cambridge.
8. Qi, H., & Yang, J. (2003). Saw tooth feature extraction of leaf edge based on support vector machine. *ICMLC*, vol. 5 , 3039-3044.
9. Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002). An efficient k-means clustering algorithm: Analysis and Implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, VOL. 24, NO. , 881-892.
10. Luciano, d. F., Coasta, R., & Marcond, C. (2004). Shape Classification and Analysis: Theory and Practice. In *Shape Classification and Analysis* (p. 15).
11. Asrani, K., & Jain, R. (2014). A Novel Shape based Descriptor for Plant Identification. *International Journal of Computer Applications* 95 (15) .