A study of Feature Extraction of leaf shape and Texture Quality Surface

Anil Kumar[1*], Dr Ajay Agarwal[2]

[1] Research Scholar, Shri Venkateshwara University Gajraula (U.P)

[2] Professor Department of Physics, Shri Venkateshwara University Gajraula (U.P)

Abstract- The aim of this paper is to the field of leaf Disorder recognition for plant identification has experienced an increased need for fast and efficient classification algorithms to aid in keeping track of the most precious plants on earth. This requirement resulted in a number of techniques revolutionizing the automatic classification area. The increasing number of techniques has led to a dilemma in deciding which of these methods have the best qualities and potential to efficiently classify. In the botanical industry where the information distortion can produce inaccurate diagnosis, this problem is particularly important. Thus, the urgent necessity of the botanical field is automated tools that help identify factories. The main process of the machine's training includes: building a leaf database, image enhancement, segmentation (leaf extraction) features, extraction and classification. Leaf disorder recognition (CAP-LDR) consists of four stages. This research aims primarily at proposing techniques for improving every plant identification operation through leaf disease. A system for improving the leaf image was proposed, called 'Enhanced wavelet-based demoizing with built-in edge enhancement and automatic contrast adjustment algorithm. This 197 method combines the wavelets, CLAHE (contrast adjustment), corner enhances and a relaxed middle filter (noise removal), with a single procedure to increase the visual quality of the leaf image.

Keywords- Leaf Shape, Texture Quality Surface, leaf image, Disorder, segmentation

- X -

1. INTRODUCTION

Three different forms of extraction from herbal leaves are presented in this chapter. The methods proposed take advantage of the idea of the least inertia axis of the leaf form to extract characteristics. The first method is to identify the 'n' number of feature points on the form curve by rotating the least inertia axis by means of an angle, three different characteristics are extracted at every feature point. The features extracted take the shape region and form boundary information into account. The form is therefore characterized by the trivalued n-dimensional characteristic vector in general. The concept of fluid inference was explored in the second method for the extraction of the features of form. A n-dimensional feature vector of fluid membership values is considered to extract fused features of a form and is characterised thereby. The third method is to explore the technique of the centroid profile to remove form-dependent features. The numbers of characteristics extracted vary greatly from one species to another depending on the plant leaf structure. Thus one species has a different dimension of the feature vector, which characterizes the leaves. All three proposed techniques for geometric transformation, such as scaling, rotation and translation, are invariant.A couple of novel methods are presented for the extraction of texture from the herbal leaves for grading. The proposed method of extraction of features takes advantage of the concept of LBP as the LBP method is regarded as a popular method for describing features in terms of texture due to its simplicity and effects. Five various methods have been suggested to extract LBP characteristics. The first process uses numerically approximated LBP Histogram for extracting more robust features and for reducing the vector's dimensionality. The second and third methods change the calculation of basic LBP code by using the neighboring adaptive threshold. Applicability of triangular and trapezoidal fuzzy functions to extract LBP texture is exploited respectively by the fourth and fifth methods. Compared with basic LBP techniques, the proposed variants of LBP texture extraction methods were more robust and efficient.

Plant leaf shape plays an important role in the identification and classification of plant species. Various methods to characterize the form and experimented on leaf datasets were proposed in literature. This section discusses some of the interesting methods. The description and classification techniques of shape in Zhang and Lu

(2004) are divided into two techniques: contour-based and regional-based. Belongie et al. (2002) introduced a method of form description that describes the distribution in spatial relationship of points of reference in relation to orientation and distance from the centroid in form contour. A descriptor of shape that extracts the topological structure of the boundary from the beam angles statistics in the third order is proposed in Arica and Vural (2003). The new form descriptor by using the internal-distance form context was introduced by Ling and Jacobs (2007) (IDSC). Its method is an insensitive and effective articulation process for complex forms with component structures. A symbolic representation of form with multi-value interval type characteristics in Guru and Nagendraswamy (2007) was proposed with the axis of the least form inertia. The clustering of forms by extracting fuszy symbolic features from two dimensional forms using a fuszy inference technology has been proposed in Nagendraswamy and Guru (2007).

2. SHAPE FEATURE EXTRACTION

In order to extract shape features, colour images of plant leaves are first converted into grey scale images. A suitable contour extraction algorithm is used to obtain the leaf contour and the axis of least inertia is computed for the contour. The proposed feature extraction techniques are based on the axis of least inertia of a shape, which preserves the orientation of a shape curve. It is very important to preserve the orientation of the shape curve to extract features, which are invariant to rotation. Also the proposed feature extraction techniques are invariant to translation and scaling. The details regarding the computation of the axis of least inertia of a shape curve can be found in (Tsai and Chen 1995).

1. Tri-Valued Shape Features

Once the axis of the lowest inertia is calculated, the two farthest points of the boundary are projected in an axis. The curve of form with a minimum inertia axis and two extreme points is shown in the Fig. 1. The distance Euclidean D between the two points defines the axis length of a shape with the least inertia. The characteristics are extracted through the shape contour in clockwise direction, with the starting point E1 or E2. The distance from E1 to the shape centroid and the distance from E2 to the shape centroid is measured in order to identify this starting point. The distance from both is regarded as a point of departure. In some cases, these two distances may be identical and the selection of the starting point may result in ambiguity. In such cases we resolve the conflict by looking at the horizontal width of the shape from the two extreme points at subsequent points of the axis.

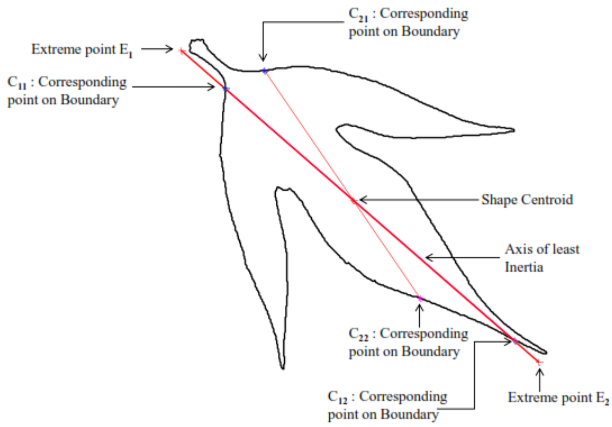


Fig. 1: After rotating the axis of least inertia of a leaf shape by an angle θ

The two extreme E1 and E2 points calculated for a shape curve might not lie on the shape boundary exactly as in Fig. 1. So we pass the axis of least inertia from two ends to cross the borders of the form curve in order to achieve the corresponding boundary points. The respective limits are regarded as characteristics. For each axis rotation the process is repeated by an angle of axis. Let C11 and C12 be the two extreme limits before the axis of least inertia is rotated. Let C21 and C22, after turning the axis of lesser inertia, be the two extreme end points of a form. The curved segment lengths, l1 and the euclidean distance, say d1 from the C11 and C21 boundaries, are calculated and the R1 ratio from d1 to l1 is calculated as feature value. Similarly, curve segment lengths are l2 and Euclideans have length d2 between C12 and C22 and R2 is calculated as another feature value. The R2 ratio is l2 and l2. In addition, Euclidean distance D is computed and taken to be another feature of two endpoints C21 and C22. R1 and R2 values are invariant by default to scale. However, for an invariant functional value (D) to be scaled, the length of the least inertian axis is divided (L).

$$D = \text{Euclidean distance between } C_{21} \text{ and } C_{22} \quad (1)$$

$$R_1 = d_1/l_1 \quad (2)$$

$$R_2 = d_2/l_2 \quad (3)$$

In eqn. (1) to eqn. (3) above, the three function values D, R1, R2 are calculated. The three D, R1 and R2 characteristic values are therefore extracted in an angle at a certain angle from the shape limit. This process is repeated by rotating the lesser inertia axis for different values of AC in the same direction in a clockwise direction. The extracted features are used to distinguish a shape. Figure2 shows a tri-evaluated feature extraction by an angle (2* a leaf sample) to rotate the less inertia axis. Fig. 3 shows the three-value leaf sample extraction by angle − an angle − after 'n' axis rotations of less inertia. The feature values obtained for a given sample of leafs are illustrated in Table 1.

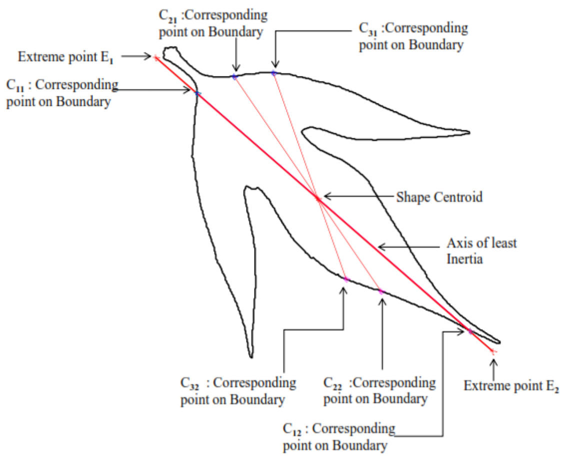


Fig. 2: After rotating the axis of least inertia by an angle (2*θ)

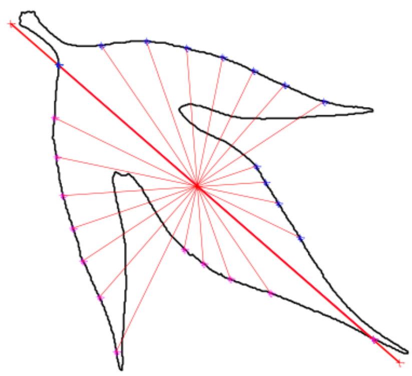


Fig. 3: After n rotations of axis of least inertia by an angle θ

Table 1: An example tri-valued crisp feature extracted from a leaf shape

Feature	Rotation	
1	θ+15°	[0.5811 0.9997 0.9203]
2	θ+30°	[0.4879 0.9994 0.9426]
3	θ+45°	[0.4190 0.9688 0.9923]
4	θ+60°	[0.3797 0.9838 0.9153]
5	θ+75°	[0.6063 0.9845 0.3947]
6	θ+90°	[0.5529 0.9397 0.9907]
7	θ+105°	[0.5884 0.5890 0.9919]
8	θ+120°	[0.3757 0.9565 0.9996]
9	θ+135°	[0.4037 0.9261 0.9803]
10	θ+150°	[0.4832 0.9823 0.9939]
11	θ+165°	[0.5854 0.6927 0.9812]
12	θ+180°	[0.8099 0.9825 0.9810]

2. Fuzzy Inference Features

The inference technique is referred to as characterizing the shape of an object as regards known geometric forms. Any general shape cannot be mapped to a known form accurately. But it is possible to measure the degree to which the general form resembles a geometric form. The value measured can vary from 0 to 1. The value closer to 1 shows that the shape is very similar to the geometrical shape and that it is nearer to 0. This is known as fuzzy inference technique in terms of membership values for a certain geometrical form and the membership values are known as fluid inference. The axis of the least inertia of a form curve is calculated and the limit points corresponding to the two extremities of the least inertia axis, as described above, are identified to extract fluid inferior characteristics from a form. With one of the boundary points, which as a starting point is the farthest away to the clockwise, the boundary of shape is divided into K segments. The center of each segment of the curve is calculated and considered a feature. The idea behind obtaining the local center of a shape curve is, instead of considering an angle or dominant points on the shape curve, to consider all the pixel-information to set a feature point, which are not consistent and solid due to noise and form transformation. Once we have reached the points on the form curve, we are moving the shape curve clockwise and at each FPI function point I = {1, 2, 3, ..., K}), forming a triangle taking into account the two projected ports PE1 and PE2 and the FPi function point. In order to approximate the triangle as equilateral, we use the fluff inference technique (Ross (2009)) and calculate the corresponding membership value as following:

Let θ1, θ2 and θ3 be the inner angles of the triangle in the order θ1 > θ2 > θ3. Let U be the universe of the triangle.

$$U = \{(\theta_1, \theta_2, \theta_3) / \theta_1 \geq \theta_2 \geq \theta_3 \geq 0; \theta_1 + \theta_2 + \theta_3 = 180 \quad (4)$$

Let d1, d2, d3 be the Euclidean distances between PE1 and FPi , PE2 and FPi and PE1 and PE2, respectively as shown in the Fig. 1.5(a). The inner angles of a triangle are computed as follows:

$$\theta_1 = \cos^{-1}\left(\frac{d_2^2 + d_3^2 - d_1^2}{2 * d_2 * d_3}\right) \quad (5)$$

$$\theta_2 = \cos^{-1}\left(\frac{d_1^2 + d_3^2 - d_2^2}{2 * d_1 * d_3}\right) \quad (6)$$

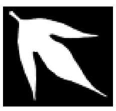
$$\theta_3 = \cos^{-1}\left(\frac{d_1^2 + d_2^2 - d_3^2}{2 * d_1 * d_2}\right) \tag{7}$$

The membership value of the triangle, which approximates equilateral triangle, is computed as

$$\mu(\theta_1, \theta_2, \theta_3,) = 1 - \frac{1}{180}(\theta_1 - \theta_3) \tag{8}$$

The membership values μi (for i = 1 to K) are computed and considered as feature values to describe a shape. Fig 4 shows the feature points obtained for a shape curve. Fig 5(a) shows an illustration of fuzzy equilateral triangle and Fig. 5(b) shows fuzzy equilateral triangulation using all the feature points. Table 2 shows an illustration of fuzzy inference feature values obtained for a given leaf sample.

Table 2: Fuzzy inference features extracted from a leaf shape (K=60)

	Fuzzy inference based features						
	1	2	3	…	…	…	*K*
	0.2204	0.1793	0.5321	…	…	…	0.1542

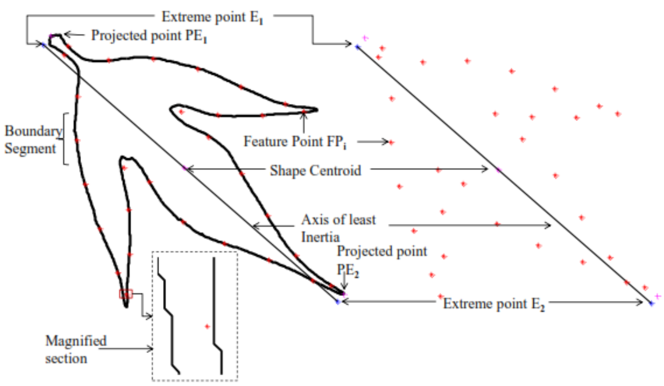


Fig. 4: Illustration of feature point extraction of a leaf shape

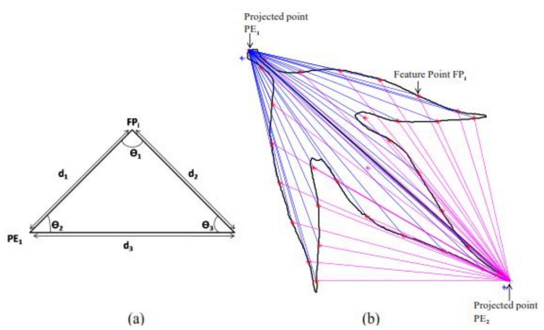


Fig. 5: (a) Illustration of fuzzy equilateral triangle (b) feature extraction of a leaf shape

3. TEXTURE FEATURE EXTRACTION

In many applications such as texture segmentation, image recovery and object detection and identification, textures analysis is an important issue. Different approaches to describe texture images, texture segmentation and classification have been offered in recent years. The texture analysis approaches can be divided into three categories, namely statistical methods, model-based methods and structural methods, according to Zhang and Tan (2002). A collection of statistics on the selected characteristics describes texture in statistical methods. A texture image is designed in model methods as a linear combination of a range of basic functions and the resulting texture picture image coefficients are used. A popular example of this method is the multichannel Gabor filter. Texture is regarded as the composition of many textured elements known as texels in structural methods. For structural methods the average intensity of elements and number of Eulers is a popular example. Statistical methods are popular for the characterization of texture among these categories. The method of feature distribution is one of the most common statistical methods that is widely considered because of its ability to characterize texture effectively. The popular texture description method for distribution of the feature models (1996) is Local Binary Patterns (LBPs) established by Ojala et al. Due to its easy computing, LBP is a popular method. In addition Ojala et al. (2002) proposed the multi-resolution analysis of LBP for every robust grayscale spatial resolution and quantization that is invariant in any monotonic grayscale transformation. In recent years, numerous LBP extensions have been proposed to overcome some of the LBP's restrictions. Guo et al. (2010a) proposed an LBP CLBP scheme that would address complementary components such as signs and magnitudes to make up the local differences in the center pixel. Guo et al. also (2010b) proposed to characterize local contrast information along with local gray scale information by providing a new descriptor called LBP Variance (LBPV).

1. Numerical Approximation of LBP Histogram

The LBP histogram obtained through basic LBP operator to capture texture can be further refined to reduce the feature set through numerical approximation. In the proposed work, we divide the histogram data bins into equal sized intervals and fit the data with an appropriate polynomial and the polynomial coefficients are evaluated. A polynomial is a function expressed as follows:

$$Y = p(x) = a_0 + a_1 x + a_2 x + \ldots + a_n x^n \tag{9}$$

for some coefficients a0, a1, a2,. . ., an. If an ≠ 0, the polynomial is an equation of order n. Based on the values of n, the equation represents a straight line (n = 1), quadratic (n = 2) simulating parabola, cubic (n = 3) and so on with (n + 1) unknowns, x being an independent variable and y as a dependent variable and constant coefficients ci . Constructing a polynomial arises from one data point (x, y) and the hypothesis about the order of the polynomial shall govern the process. Further, the values for those coefficients ci are computed. Fig. 6 shows the LBP histogram of a leaf sample.

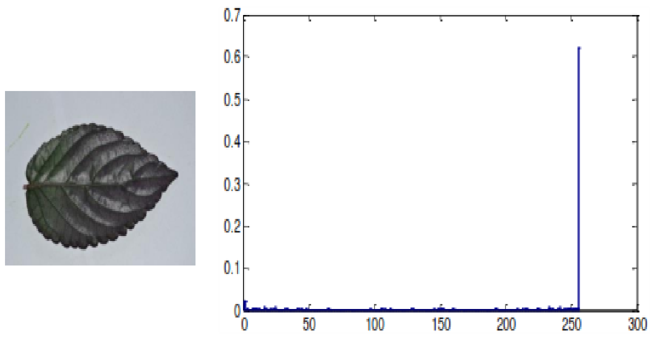


Fig. 6: LBP histogram of a leaf with 256 bins

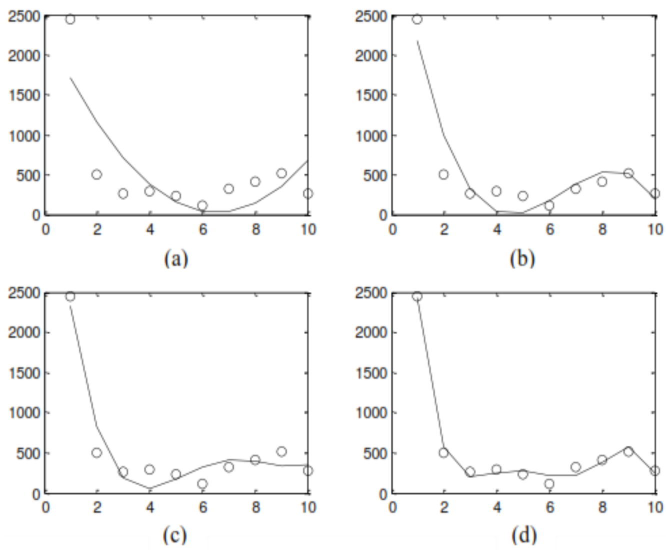


Fig. 7: Polynomial curve fitting generated for a sample of 10 data bins

Fig. 7 shows the polynomial curves fitted for a sample of 10 histogram data bins representing LBP histogram features. Fig.7 (a) to Fig.7 (d) shows the polynomial fit for 2nd order to 5th order respectively. From the plots it can be observed that the 5th order polynomial fits best for the data and hence, we have used 5th order polynomial for approximation of LBP in this method.

4. CONCLUSION

This research focuses on the automation of leaf disorder plant identification. The plant identification can be automated with the aid of different parts of a plant anatomy such as stem, flowers, petals, seeds and leaves in the addition of the plant. The leaf section of the plant is used for identifying a plant. Biodiversity and the facility to create digital images have increased the need for a machine-learning processing power and economical methods. The plant identification using computers has become a topic of interesting research in order to collect information. The global shortage of expert taxonomists further enabled non-botanical individuals to performatize valuable field work of plant identification and characterisation. In many fields, including agriculture, forestry and pharmacology, these tools are important. The first step in designing and developing such instruments begins with leaf disorder. The identification of plants based on leaf images is the most successful and proven method compared to other methods like cell and molecule biology methods. Due to the availability of low-cost digital cameras, sampling leaves and photography are convenient and feasible.Currently, leaf disorder identification requires the search for information on a plant which mostly matches the name (species) known in advance. While it is a time consuming task to identify plants using such a key, correct use of the key plays a direct role in plant search success. The alternative way to provide a picture is very practical, easy to use and eliminates key requirements. The job combines the challenges in various areas such as picture processing, machine learning and pattern disorder. The main focus of this study is the identification of the most favorable algorithms and techniques for the identification of plants by leaf disorders.

5. REFERENCES

1. Abdel-Aal et al., (2006) Abdel-Aal, R. E., Abdel-Halim, M. R., and Abdel-Aal, S. (2006). Improving the classification of multiple disorders with problem decomposition. Journal of biomedical informatics, 39(6):612–625.
2. Agarwal et al., (2006) Agarwal, G., Belhumeur, P., Feiner, S., Jacobs, D., Kress, W. J., Ramamoorthi, R., Bourg, N. A., Dixit, N., Ling, H., Mahajan, D., et al. (2006). First steps toward an electronic field guide for plants. Taxon, pages 597–610.
3. Alaei et al., (2010) Alaei, A., Nagabhushan, P., and Pal, U. (2010). A new two-stage scheme for the recognition of persian handwritten characters. In International Conference on Frontiers in Handwriting Recognition, ICFHR 2010, Kolkata, India, 16-18 November 2010, pages 130–135.
4. Anami et al., (2010) Anami, B. S., Nandyal, S. S., and Govardhan, A. (2010). A combined color, texture and edge features based approach for identification and classification of indian medicinal plants. International Journal of Computer Applications, 6(12):45–51.
5. Arica and Vural, (2003) Arica, N. and Vural, F. T. Y. (2003). Bas: a perceptual shape

descriptor based on the beam angle statistics. Pattern Recognition Letters, 24(9):1627–1639.

6. Backes and Bruno, (2010a) Backes, A. R. and Bruno, O. M. (2010a). Plant leaf identification using color and multi-scale fractal dimension. In ICISP, pages 463–470.
7. Backes and Bruno, (2010b) Backes, A. R. and Bruno, O. M. (2010b). Shape classification using complex network and multi-scale fractal dimension. Pattern Recognition Letters, 31(1):44–51.

Corresponding Author

Anil Kumar*

Research Scholar, Shri Venkateshwara University Gajraula (U.P)