CHAPTER 2

REVIEW OF LITERATURE

The main goal of this research is to build a computer aided classifier that classifies leaves and assigns the names of the plant they belong to, when supplied only with the leaf image as input. This goal is met through a series of sequential steps, namely, image enhancement techniques, extraction of ROI region, extraction of features and finally classification. Several researchers have contributed and proposed various algorithms in each of these steps. This section presents a review of some important publications made in these steps to understand the current research status.

2.1. IMAGE ENHANCEMENT

Most of the reviewed solutions consider denoising as an important step in image enhancement. Various other preprocessing techniques can also be used to enhance the leaf image obtained. Examples include boundary or edge enhancement, smoothening and contrast adjustment. These techniques can also be applied to improve the quality of the leaf image (Tzionas *et al.*, 2005). In this study, the input leaf image is enhanced through three operations, namely, denoising, contrast adjustment and edge enhancement. This section presents studies related to these three areas.

2.1.1. Noise Removal

Plant leaf image is normally affected by three types of noises, namely, Fixed-Valued Impulse Noise (Salt-and-Pepper Impulse Noise), Random-Valued Impulse Noise (Uniform Impulse Noise) and Gaussian Noise. Solutions were proposed that denoises any one of the above mentioned noise using various techniques. Some of these techniques are presented in this section.

Rui *et al.* (1996) proposed a Modified Fourier Descriptor (MFD) method to achieve translation, scaling and rotation invariance by considering the distance between the FD (Fourier Descriptor) magnitude and the phase angle separately so as to decrease the discrimination noises. A Minimum Noise Fraction (MNF) transformation was performed to control the noise in the imagery by Green *et al.* (1988).

According to El-Helly *et al.* (2003), image enhancement is a sub-field of image processing and consists of techniques to improve the appearance of an image, to highlight important features of an image and to make the image more suitable for use in subsequent tasks of leaf classification. They proposed a three-step algorithm for enhancing an image. The first step used HSI transformation to take advantage of their capability of separating color information from its intensity information. The second step used histograms to analyze the intensity channel and used threshold method to enhance the contrast of the image. The third step again used threshold methods to adjust the intensity of the image.

A histogram equalization method was used by Pan and He (2008) to enhance the leaf image. A coupling method was adopted by Li *et al.* (2010) who combined adaptive local smoothing method and wavelet to cope with noisy leaf image. The method was able to preserve edges while removing noise while maintaining the contrast and visual effect of the leaf image. Ma *et al.* (2010) used an image preprocessing technique to reduce noise from source leaf image and enhanced areas of interest using minimum error threshold method. Sathyabama *et al.* (2011) used the difference of Gaussians to increase the visibility of edges and other details present in the digital leaf images.

According to Wang and Wu (2009), noise in an image can be detected at a high noise level. Using this information Zhang (2010), introduced a two-phase method for removing noise using Adaptive Center-Weighted Median Filter (ACWMF) and variation method. This method successfully removed the noise, but was sensitive to the size and shape of the filter window. Bigger size window resulted in over smoothening of the image, and smaller size did not remove the noise efficiently. The window size problem is a common problem that is shared by many filtering algorithms.

Solutions to these weaknesses in noise removal algorithms were provided. They proposed a method based on Interval-Valued Fuzzy Sets (IVFS) entropy application to denoise an image. The method combined image histogram information and spatial information about pixels of different gray levels by using an IVFS multi-thresholding technique. The advantage of this method was that it was able to remove both impulse and Guassian noise.

Rubio (2010) extended this method to remove impulse and Guassian noise present in the same image by using Iteratively Reweighted Norm (IRN) method to yield predictions of the original pixel values and compute the corresponding predicted errors and train the noise model by using an Expectation-Maximization (EM) algorithm. The proposed algorithm can effectively remove the impulse noise with a wide range of noise density and produce better results in terms of the qualitative and quantitative measures of the images.

2.1.2. Contrast and Edge Enhancement

Improving the quality of edges in leaf images is a critical factor that can improve the performance of recognition and identification. While considering digital images a maximum number of reported works concentrated on edge detection (Yu and Acton, 2004) but only a few have been reported with edge enhancement.

According to Li *et al.* (2007), edge enhancement is an important operation which helps in detecting object boundaries and in subsequent steps of recognition and classification. It helps in differentiating the features by improving the visual quality perception of the image and provides insight into the shape and outline of objects and offers vital information to the Human Visual System (HVS).

One of the early methods to address the problem of edge quality enhancement uses anti-aliasing, wherein, two adjacent pixels in the oblique direction are detected and a corrective pixel is inserted to smooth the line (Yonezawa *et al.*, 1978).

Techniques that (slightly) manipulate image content for better edge quality was studied by Gupta (1981) where the intensity of a pixel is chosen depending on the distance between the center of the pixel and the edge of the image. Ort (1981) proposed a technique for shifting pixels by half a position while printing diagonal elements. Shirasaka (1998) used a similar technique to detect the staircase regions. This is not a very effective approach for regions containing complex contours.

Template matching is another widely used technique for edge quality enhancement, wherein the image region is scanned in the piece-wise order (windows) and compared against a defined set of patterns to be rectified (Yao *et al.*, 2006; Lund, 1997; Tung, 1989). However, this approach requires predefined regions and conditions that can add to the processing time depending on the image complexity. Clayton (2006) first converted image to binary form (black and white), enhanced edges, removed noise and then reconverted to its original colour domain.

Braica (2006) states a method of enhancement by increasing the contrast of the image at the edges. Chen *et al.* (1997) globally shifted the intensity value in the approximation-coefficients to achieve contrast enhancement. But they did not provide an efficient way to decide the size of the shifting step. They also established a zero-crossing tree, which consists of zero-crossings of each component in multiple-resolution levels, to represent multiple-resolution edges that are used to suppress noise.

Fu *et al.* (2000) analyzed the drawback of the HEQ procedure in the spatial domain. Then they proposed a wavelet-based contrast enhancement method. In their method, after performing the HEQ procedure in the spatial domain, the output image was transformed into the wavelet domain. Then, all approximation-coefficients were squared. They claimed that the proposed process could compensate the information that was lost during the HEQ process.

Reeves *et al.* (1997) investigated a wavelet transform domain filter, based on the LLMMSE filter (Kaun *et al.*, 1985) to suppress noise and enhance edges. They also applied global HEQ to the wavelet approximation-coefficients at the coarsest decomposition level to enhance contrast. But further investigation was required in order to understand how the selection of the approximation-coefficients' range and histogram bin values affect the reconstructed image.

Xu *et al.* (1997) combine a wavelet phase filter at finer scales in the wavelet domain to reduce noise, and a semi-soft wavelet shrinkage technique was proposed by Bruce and Gao (1996) at coarse scales in the wavelet domain to further reduce noise and enhance edges. But the proposed method still could not automatically adjust its parameters to achieve optimal result.

There are some other enhancement attempts. For example, Gong *et al.* (2000) rationally enlarged coefficients on multiple scales in the wavelet domain. Xu *et al.* (2000) altered the amplitude of coefficients in the wavelet domain. Peng *et al.* (2000) used a non-linear enhancement operator on coefficients at multi-scale in the wavelet domain. However, these attempts aimed at improving only image contrast and ignored edges.

2.2. SEGMENTATION

Segmentation, a subtask in image processing, is an extensive task which is applied in many areas other than computer vision (Jain and Dubes, 1988). Recently there has been a considerable amount of work on image segmentation (Karvelis *et al.*, 2008; Withey *et al.*, 2007; Peters and Kerdels, 2007; Zhang *et al.*, 2010). With the increasing size and number of digital images, the use of computers in facilitating their processing and analysis has become necessary. In particular, computer procedures for the delineation of objects and other regions of interest are a key component in assisting and automating recognition and identification tasks. These algorithms, called image segmentation algorithms, play a vital role in numerous applications like medical (Sharma and Aggarwal, 2010), Content based image retrieval (Frigui and Caudill, 2006) and biometric recognition (Mathivanan *et al.*, 2011) systems. Methods for performing segmentations vary widely and depend on various factors like application and modality.

General imaging artifacts such as noise, contrast variation, partial volume effects and motion artifacts can also have significant consequences on the performance of leaf extraction algorithms. Furthermore, each imaging modality has its own idiosyncrasies with which to contend. In this era of modern technology, no single segmentation algorithm that provides acceptable results for every image processing application is available. Methods that are more general do exist and can be applied to a variety of data. Nevertheless, methods that are tailor-made to specific applications and which take into account prior knowledge often attain better performance. Selection of an appropriate approach to a segmentation problem may therefore be a difficult task.

2.2.1. General Segmentation Methods

A fundamental aspect of segmentation algorithms is to represent the chosen objects. A detailed review is proposed by Pham *et al.* (2005), who reviews the various methods available for medical image segmentation and also reviews them under various biometric parts of the body.

The first attempt at automatic segmentation of images was simple techniques such as global or local thresholding (Lang *et al.*, 1996). The major advantage with this representation is that it is very easy to implement. It is also easy to calculate object quantities such as volume; it is simply to count the voxels. Disadvantages are that the representation is not capable of subvoxel resolution, and direct visualization with volume or surface rendering gives poor results due to the difficulty in calculating surface normal. Furthermore, the main concern of this representation is that the local methods can generate infeasible object boundaries due to spuriously detected edges.

Seed growing algorithms were early introduced (Cline *et al.*, 1987) and have been combined with multi-spectral approaches where a set of images, acquired using different imaging parameters, was used forming a feature vector in each voxel (Wigstrom and Svensson, 2002, Bijnens *et al.*, 2005).

Edge based or boundary tracing algorithms(Ma *et al.*, 2000, from a user selected point traces a boundary, often using dynamical programming have been developed (Maes *et al.*, 1993). This work described a semi-automatic system for the extraction of the endocardial border on a sequence of two-dimensional short axis echocardiograms. The delineation in the images was based on a dynamic programming technique which takes into consideration the radial and tangent gradient, the smoothness, the intensity of the extracted contour along with a rough estimation of the location of the contour, extracted from the previous frame of the time sequence. The user is asked to specify a region of interest for the first frame and the developed system defines a circle as a rough estimation of the location of the endocardium within this region of interest. The system iterates this process to acquire a more detailed outline of the endocardium, which them from successive echo-frames are verified as the system compares their areas.

Boundary tracing approaches were also combined with algorithms based on knowledge based algorithms (Suh *et al.*, 2003) or regional information (Chakraborty *et al.*, 2006). Most of these algorithms were combined with different types of methods to include a priori information. Examples include self learning approaches such as neural networks (Cheng *et al.*, 2008), mathematical morphology (McEachen and Duncan, 1997) and Bayesian reasoning (Klingler *et al.*, 1988).

One of the most popular methods to include a priori information has been by using deformable curves or models. The deformable model approach that has attracted the most attention is popularly termed 'snakes' model (Kass *et al.*, 1988).

Ranganath (2005) investigated automatic extraction of left ventricular contours from cardiac Magnetic Resonance Imaging (MRI) studies. The contour extraction algorithms were based on active contour models or snakes. Based on cardiac MR image characteristics, the algorithm proposed extract contours from these large data sets. The algorithm considers contour propagation methods to make the contours reliable in the presence of artifacts, noise and poor resolution with minimum of user interaction. The study considered both spin and gradient echo. The result after contour extraction was used to determine quantitative measures for the heart along with graphically rendered cardiac surfaces.

Snakes are inherently a 2D approach. The concept of deformable models, can, however be extended to multidimensional images, 2D+T (Kucera and Martin, 2007) 3D or 3D+T (Bardinet *et al.*, 1996). The main difficulty for deformable models in 3D is the object representation. For numerical stability, the object representation may need to be refined, and this is a non-trivial task in 3D. Commonly used representations are triangles (Kaus *et al.*, 2004; McInerney and Terzopoulus, 1995).

Another segmentation approach in digital images is level set algorithms. The novelty with this approach lies in the object representation with the introduction of a level set function $\phi(x, y,...)$. The object boundaries are found where $\phi(x, y, ...) = 0$, where ϕ is the level set function, and x, y, ... are spatial coordinates. An analogy of the level set function is the isobar curves on the weather forecasts. The function ϕ is usually simply used for sampling over a Cartesian grid. The level set function is calculated as a solution to a differential equation. From this boundary condition, a solution ϕn is found by solving a partial differential equation where zero-level set has propagated with a speed determined by a speed image, and locally estimated parameters such as curvature of the level set surface. The process is repeated until some sort of convergence or stopping criteria is met. The calculation of the level set is rather computationally expensive, and therefore, several approximations and fast schemes have been proposed (Sethian, 2009).

An interesting approach to further speed the process up by carrying out the calculations on the modern hardware accelerated graphics card, and was proposed by Aaron *et al.* (2003). An overview of level set methods and their applications is given by Sethian (1999). The advantage with level sets is the ability to handle complex geometries, and multiple objects. The disadvantage is that it is difficult to include a priori information, even though there have been attempts (Leventon *et al.*, 2000, Paragios, 2003).

Leventon *et al.* (2000) presented a novel method of incorporating shape information into the image segmentation process. They introduced a representation for deformable shapes and define a probability distribution over the variances of a set of training shapes. The proposed segmentation process embeds an initial curve as the zero level set at a higher dimensional surface and evolves the surface such that the zero level set converges on the boundary of the object to be segmented. The MAP (Maximum a Posteriori) and Shape of the object in the image was estimated at each step of the surface evolution. The shape of the object is based on the prior shape information and the image information. They then evolved the surface globally and locally based on image gradients and curvature. The proposed system was analyzed using a synthetic data and 2D and 3D medical imagery.

Paragios(2003) introduced knowledge-based constraints for segmentation while preserving the ability to deal with local deformations. He proposed a variation level set framework that can account for global shape consistency as well as for local deformations. In order to improve the performance, the problems of segmentation and tracking of the structure of interest were dealt with simultaneously, introducing the notion of time in the process and looking for a solution that satisfies the prior constraints while being consistent along consecutive frames. Experimental results in magnetic resonance and ultrasonic cardiac images were used during performance evaluation.

Another popular segmentation algorithm is the active appearance models. In active appearance models the behavior and basic shape is learned from a manually delineated training set. A three dimensional (3D/2D+T) implementation for the left ventricle is given by Mitchell *et al.* (2002). In their work, a model-based method for three-dimensional image segmentation was developed.

An excellent overview and review of active appearance models is given by Stegmann (2000) and Stegmann (2004). They proposed and evaluated methods for automated analysis and quantification of digital images. A common theme was the usage of generative methods, which drew inference from unknown images by synthesizing new images having shape, pose and appearance similar to the analyzed images. The theoretical framework for fulfilling these goals was based on the class of Active Appearance Models.

Chaabane *et al.* (2010) a novel method of colour image segmentation, based on fuzzy homogeneity and a data fusion technique, was proposed. The general idea of mass function estimation in the Dempster-Shafer evidence theory of the histogram was extended to the homogeneity domain. The fuzzy homogeneity vector was used to determine the fuzzy region in each primitive colour, whereas, the evidence theory was employed to merge different data sources in order to increase the quality of the information and to obtain an optimal segmented image. Experimental results from the proposed method were validated and evaluated and were compared with existing techniques. The experimental results demonstrated the superiority of introducing the fuzzy homogeneity method in evidence theory for image segmentation.

2.2.2. Plant and Leaf Related Methods

Segmentation for plant identification can be performed in two manners. The first is to separate an entire plant from its background and the second is to separate the leaf from other objects of the image. Separating the entire plant may be done using a number of spectral and colour based approaches. Usage of near-infrared (NIR) information was proposed by Guyer *et al.* (1993) who used a camera with visible wavelength blocking filter. As digital RGB cameras have become readily available (and NIR bands have not), segmentation approaches using colour indices were more common. A modified excess green measure, using the normalized chromacity values, was used to segment plants (Woebbecke *et al.* 1995; Tang *et al.* 2003). A study by Philipp and Rath (2002) worked on using colour space transformations for plant discrimination.

Until plants reach a point where individual plants are starting to overlap each other, the entire plant approach may be appropriate. However, as scene complexity increases and individual leaves are required, these approaches are hindered by the similarity of leaf colour. Several works have focused on this problem. Lee and Slaughter (2004) developed a watershed-style algorithm for defining boundaries between multiple-leaf blobs that had been segmented from a colour image. This approach attained upto 57% separation performance on tomato seedlings, with this measure defined as the number of properly segmented leaves divided by the sum of leaves that were fragmented, not separated, and separated. It was, however, computationally demanding.

Deformable templates have also been investigated for leaf segmentation (Manh *et al.* 2001). Green foxtail (*Setaria viridis (L.) Beauv*.) and background were segmented using colour information. Starting points for template fitting were defined as leaf tips, which were found using a global search by a small window, followed by regional assessment when the small window was covering plant pixels exclusively. An ellipse was rotated around the tip and the position that generates the highest number of plant pixels covered by the ellipse was taken as the template starting point. The template was iteratively deformed outward until a stopping criterion was met. This approach was dependent on leaves having definable tips, the appropriate definition of the template skeleton and was not able to account for shared leaf edges within the term driving the deformation, making it somewhat susceptible to occlusion.

An edge-tracing algorithm was developed by Franz *et al.* (1995), where after determining gradient slope and magnitude using a set of 3x3 kernels, edges were traced using a series of algorithms that were similar to those of the Canny edge detector to produce a single-pixel wide edge definition using upper and lower edge thresholds. A considerable amount of additional logic for excluding petioles and stems and for linking edges was included beyond the classical Canny detector. This algorithm required user input for deciding where to start traces and locating areas of interest.

A majority of the methods proposed are based on two basic properties of the pixels in relation to their local neighbourhood: discontinuity and similarity (Cheng *et al.*, 2001). The approaches based on discontinuity partition an image by detecting lines, isolated points and edges, which were known as edge detection techniques (Maire *et al.*, 2008; Dollar *et al.*, 2006). Another class of segmentation techniques are based on multiresolution analysis. Such methods are typically based on multiresolution transformations. Small details, in general, are detected in higher resolution images, while larger objects are segmented in coarser images.

Kim and Kim (2002, 2003) proposed a multiresolution wavelet-based watershed image segmentation technique, using markers and a region merging procedure to reduce over-segmentation. Jung and Scharcanski (2005) used a combined image denoising/enhancement technique based on a redundant wavelet transform for segmentation.

Ma and Manjunath (1997, 2000) proposed the Edge Flow segmentation technique, which consists of computing and updating changes in color and texture in a pre-defined scale. Deng and Manjunath (2001) proposed the JSEG method for multiscale segmentation of color and texture, based on color quantization and region growing.

Wu *et al.* (2000) proposed a multiscale wavelet-based directional image force as external force for snake segmentation. Comanicui and Meer (2002) used a kernel in the joint spatial-range domain to filter image pixels and a clustering method to retrieve segmented regions. Ozden and Polat (2007) proposed a color image segmentation method based on low-level features including color, texture and spatial information. Chen *et al.* (2003) proposed a color texture image segmentation algorithm based on wavelet transform and adaptive clustering algorithm.

Arbelaez and Cohen (2008) proposed an algorithm for constrained segmentation. The proposed method is a front propagation algorithm on the ultra-metric contour map that constructs Voronoi tessellations with respect to collections of subsets of the image domain. The algorithm is parameter-free, computationally efficient and robust. However, it needs to place the seed point inside each object of interest in the image by a human user.

Ning *et al.* (2010) presented a new region merging based interactive image segmentation method. It needs to roughly indicate the location and region of the object and background by using strokes, which are called markers. Moreover, a novel maximal-similarity based region merging mechanism was proposed to guide the merging process with the help of markers. The method is efficient but it is human user dependent.

Zhang *et al.* (2010) proposed a novel region-based ACM for image segmentation which was implemented with a new level set method named Selective Binary and Gaussian Filtering Regularized Level Set (SBGFRLS) method. The SBGFRLS method reduces the expensive re-initialization of the traditional level set method to make it more efficient.

2.3. FEATURE EXTRACTION

Feature selection is an important task that allows the determination of the most relevant features for pattern recognition. Selecting suitable features is a critical step for successfully implementing an image classification. Many potential variables may be used in image classification, including vegetation indices, spectral signatures, textural or contextual information, transformed images, multisensor images, ancillary data and multitemporal images. Due to different capabilities in land-cover separability, the use of too many variables in a classification procedure may decrease classification accuracy (Hughes 1968, Price *et al.* 2003).

Many approaches, such as principal component analysis, discriminant analysis, minimum noise fraction transform, non-parametric weighted feature extraction, decision boundary feature extraction spectral mixture analysis and wavelet transform analysis (Myint 2001, Okin *et al.* 2001, Rashed *et al.* 2001, Lobell *et al.* 2002, Asner and Heidebrecht 2002, Neville *et al.* 2003, Landgrebe 2003, Platt and Goetz 2004) may be used for feature extraction, in order to reduce the data redundancy.

These methods can be divided into filter and wrapper models. Filter models investigate indirect performance measures, mostly based on distance and information measures (Liu and Motoda, 1998). These techniques study the feature selection task independently of a classifier, on the other hand, wrapper models use classification as a subtask, testing the classification for different subsets of features, until an optimum was found by Fukunaga (1990). Wrapper methods are computationally feasible only for small feature vectors because they are much more time-consuming as each iteration of the method requires classifier execution and testing.

Feature selection methods use search strategies, which can be categorized as exhaustive (greedy algorithms), heuristic (sequential) and non-deterministic (randomized). Both Euclidean and Mahalanobis distances can be calculated. However the information of Euclidean distance is limited as it is used for uncorrelated variables (Sabino *et al.*, 2004). In statistical analysis, forward and backward stepwise multiple regression (SMR) are widely used to select features. The output here is the smallest subset of features resulting in a correlation coefficient value that explains a significantly large amount of the variance.

Rough sets theory was also used to determine the degree of dependency of sets of attributes for selecting binary features. The most popular feature selection methods in machine learning literature are variations of sequential forward search (SFS) and sequential backward search (SBS) (Piramuthu, 2004). SFS (SBS) obtains a chain of nested subsets of features by adding (subtracting) the locally best (worst) feature in the set. Lin and Cuningham (1995) proposed a very fast method for input selection introducing the fuzzy curve concept.

Optimal selection of spectral bands for classifications has been extensively analyzed in previous literature (Mausel *et al.*, 1990, Landgrebe, 2003). Graphic analysis (e.g. bar graph spectral plots, co-spectral mean vector plots, two dimensional feature space plot and ellipse plots) and statistical methods (e.g. average divergence, transformed divergence, Bhattacharyya distance and Jeffreys-Matusita distance) have been used to identify an optimal subset of bands (Jensen, 1996).

Penaloza and Welch (1996) explored the fuzzy-logic expert system for feature selection. Peddle and Ferguson (2002) examined three approaches (exhaustive search by recursion, sequential dependent search and isolated independent search) for optimizing the selection of multisource data, and found that these approaches were applicable to a variety of data analyses. In practice, a comparison of different combinations of selected variables is often implemented, and a good reference dataset is important. Specifically, a good representative dataset for each class is a key for implementing a supervised classification. The divergence-related algorithms are often used to evaluate the class separability and then to refine the training samples for each class.

As an intrinsic characteristic feature, the leaf vein certainly comprises of the important information for plant species recognition in spite of its complex modality. An effective two-stage technique for leaf vein extraction is presented by Fu and Chi (2003). At the initial stage, a preliminary segmentation based on the intensity histogram of the leaf image is executed to evaluate the rough regions of vein pixels. Then, at the second stage, a fine checking is carried out by means of a trained Artificial Neural Network (ANN) classifier. Ten features refined from a window centered at the pixel are utilized as the input to train the ANN classifier.

Leaf classification is a significant constituent of computerized living plant recognition(Zhang,2004). The leaf includes significant information for plant species recognition in spite of its complication. The major purpose of the system proposed by Zulkifli *et al.* (2011) is to evaluate the efficiency of Zernike Moment Invariant (ZMI), Legendre Moment Invariant (LMI) and Tchebichef Moment Invariant (TMI) characteristics in extracting features from leaf images. Subsequently, the features obtained from the most effective moment invariant approach are classified using the General Regression Neural Network (GRNN).

A new technique for feature extraction from a natural image like plant leaf was developed by Prasad *et al.* (2011) for automated living plant species identification which would be helpful for botanical students to carry out their research for plant species identification. A multi-resolution, multidirectional Curvelet transform is executed on subsegmented leaf images to obtain leaf information. Precisely, the orientation of the object that enhances the accuracy rate in the image is not taken into account. These coefficients are given as the input to a trained SVM classifier to categorize the result.

Chi *et al.* (2003) developed a novel approach of Gabor filter banks specifically designed for plant species recognition by using their bark texture characteristics. In this approach, texture is modeled as numerous narrowband signals that are distinguished by their central frequencies and normalized ratios of amplitudes. The normalized ratio of amplitudes is utilized as an energy weight for integrating narrowband signals. In accordance with this texture model, a collection of texture features can be obtained from each kind of plant bark that is helpful to differentiate the plant and to design the equivalent Gabor filter bank. A classifier is built by these Gabor filter banks.

2.4. PLANT IDENTIFICATION AND LEAF RECOGNITION

Plants play a considerable part in both human life and other lives that are present on the earth. Plant recognition, depending on images of leaf, flower and fruit is an extremely demanding task in the area of computer vision and pattern recognition. Identification of plants according to the images of leaf is a very challenging task. As a result of the corrosion in the environmental and inadequate awareness, numerous rare plant species are at the margins of death. Despite the enormous development in botany, there are several plants yet to be exposed, classified and exploited; unidentified plants are resources waiting to be established.

Leaf classification and recognition for plant identification play a major role in all these activities. There has been a modest work concerned on leaves, flower and fruit image processing and recognition. At present, a lot of researchers have committed their work on leaf recognition. As an inherent characteristic, leaf vein certainly contains significant information for plant species recognition regardless of its complex modality.

Many of the techniques proposed for leaf recognition fall under two headings:-

1. Shape and contour-based classification techniques and
2. region-based classification techniques.

2.4.1. Shape and Contour Feature Based Classification

Leaf identification is a process resulting in the assignment of each individual plant to a descending series of groups of related plants, as judged by common characteristics. So far, this time-consuming process has mainly been carried out by botanists. Plant identification has had a very long history, from the dawn of human existence. Currently, automatic (machine) plant recognition from color images is one the most difficult tasks in computer vision due to several factors such as

- Lack of proper models or representations
- A great number of biological variations that a species of plant can take and
- Imprecise image preprocessing techniques such as edge detection and contour extraction, thus resulting in possible missing features.

As the shape of leaves is one the most important features for characterizing various plants, the study of leaf image classification or retrieval will be an important step for plant identification. In this section, leaf image classification based on shape features is to be addressed. A number of shape representations such as chain codes, Fourier descriptors, moment invariants and deformable template as well as various matching strategies have been proposed for shape based image classification. This section analyzes some works based on shape and contour based features.

Automatic recognition of wild flowers using shape features of leaves and flowers (Saitoh and Kaneko, 2000), leaf image retrieval with combination of different shape based features of leaves (Wang *et al.*, 2000), feature extraction of leaves using image processing techniques (Cunha, 2003) and recognizing plant species of Acer family by leaf shapes are examples of studies pursuing computer based plant biometrics. Most of the studies were based on global shape descriptors (area, perimeter, width and length, compactness, eccentricity) (Wang *et al.*, 2002), shape signatures (centroid-contour distance) (Wang *et al.*, 2003), global representations of leaf peripheral such as polygonal approximations (Nam *et al.*, 2005a; Nam *et al.*, 2005b) and the local feature extraction techniques (angle code histogram) (Wang *et al.*, 2002). Others were based on leaf vein extraction using intensity histograms and trained artificial neural network classifiers (Fu and Chi, 2003). Some of them (Wang *et al.*, 2003) achieved relatively low accuracy due to the fact that the techniques they applied extracted only global information.

Wang *et al.* (2000) presented an efficient two-step approach of using a shape characterization function called centroid-contour distance curve (CCD) and the object eccentricity (or elongation) for leaf image classification and retrieval. Furthermore, they have proposed a thinning-based starting-point locating algorithm (closest point on the contour for each end-point on the skeleton) for CCD, which is effective in identifying starting-point(s) and reducing the rotation-and-matching time. In the first step, the eccentricity was used to rank leaf images, and the top scored images were further ranked using the centroid-contour distance curve together with the eccentricity in the second step. Two data sets, 135 leaf images from one plant and 233 images from ten plants were used during experimentation and the results showed that the proposed starting-point locating algorithm is more efficient than the Fourier transformation and the correlation methods.

A further modification to the above method has been done by Wang *et al.* (2003). In this study, they have added another feature, angle code histogram (ACH), for the above two-step leaf image retrieval approach. In the second step, in addition to CCD and ECC, ACH was also used to rank the top scored images resulted from the first step described in Wang *et al.* (2000). For locating the starting points, a further improved algorithm (by removing the very short skeleton branches) was also proposed. A database containing 1400 colour leaf images from 140 species have been used and the results proves that this method is computationally more efficient compared to the existing two methods (curvature scale space method and modified Fourier descriptor method) in feature extraction and feature matching time. However, they obtained relatively low recall rates (defined as a percentage of number of returned images) which have the same class to the number of database images and it reflects the overall fact that the leaf shape alone is not sufficient to distinguish different plant species because different species of plants may have very similar leaf shapes. Leaf-features such as shape of the leaf apex and the base, leaf margin, colour, venation and the texture of the leaf surface, leaf arrangement are also very important in plant identification.

Continuing in the same line of contour based leaf classification, Mokhtarian and Abbasi (2004) addressed the problem of two-dimensional (2-D) shape representation and matching in the presence of self-intersection for large image databases. Self-intersection occurs when part of a leaf object is hidden behind another part and results in a darker section in the gray level image of the leaf. For many classes of leaves, self-intersection is inevitable during the scanning of the image. The boundary contour of the object must include the boundary of this part which is entirely inside the outline of the object. The curvature scale space (CSS) image of a shape is a multiscale organization of its inflection points as it is smoothed. The CSS-based shape representation method has been selected for MPEG-7 standardization. The authors studied the effects of contour self-intersection on the curvature scale space leaf image. When there is no self-intersection, the CSS image contains several arch shape contours, each related to a concavity or a convexity of the shape. Self intersections create contours with minima as well as maxima in the CSS image. An efficient shape representation method that describes a shape using the maxima as well as the minima of its CSS contours was introduced. This is a natural generalization of the conventional method which only includes the maxima of the CSS image contours. The conventional matching algorithm was also modified to accommodate the new information about the minima. The method was successfully used in a real world application to find similar classes from a database of classified leaf images representing different varieties of chrysanthemum.

Following the work of Mokhtarian and Abbasi (2004), Park *et al.* (2008) proposed a new and effective leaf image categorization scheme. In this scheme, leaf venation was analyzed for leaf categorization. The leaf shape features were then extracted and utilized to find similar leaves from the already categorized group in a leaf database. The venation of a leaf corresponds to the blood vessels in organisms. Leaf venations are represented using points selected by a curvature scale scope corner detection method on the venation image. The selected points were then categorized by calculating the density of feature points using a non-parametric estimation density. The effectiveness of the system was proved using several experiments on a prototype system.

Using 1032 leaf images, Nam and Hwang (2005) implemented a prototype shape-based leaf image retrieval system. They used a hybrid-search scheme that uses the leaf shape and the leaf arrangement on the stem. Experimental results proved that the system was efficient than the methods involving Fourier descriptors, centroid-contour distance curve, moment invariant, curvature scale space descriptor and minimum perimeter polygon. Using the same findings of the study, Nam and Hwang (2005) ,with a new hybrid-search scheme that uses leaf shape, leaf arrangement and venation and Nam *et al.* (2005b) presented a leaf image retrieval system (CLOVER) for mobile devices and Nam *et al.* (2005a) presented an efficient leaf retrieval system called ELIS.

In yet another work, Nam *et al.* (2008) proposed a novel system that used similarity-based measurement leaf identification. For the effective measurement of leaf similarity, a combination of shape and venation features were considered together. In the shape domain, a matrix of interest points to model the similarity between two leaf images was constructed. In order to improve the retrieval performance, an adaptive grid-based matching algorithm was implemented. This algorithm computes a minimum weight from the constructed matrix and uses it as similarity degree between two leaf images, based on the Nearest Neighbor (NN) search scheme. The main purpose behind using NN search scheme is to reduce the search space during matching process. In the venation domain, an adjacency matrix was constructed from the intersection and end points of a venation to model similarity between two leaf images. Based on these features, a prototype mobile leaf image identification system was developed. The experimental results published showed that the scheme proposed a great performance when compared with existing methods.

Most of these studies have been carried out to apply image classification techniques for a range of leaves without any prior categorization. However, leaves are very different in shape. Studies show that different shape signatures have different effects on shape retrieval and classification. For example, with character recognition, the shape signature cumulative angular function has been used most successfully, whereas in discriminating general shapes, centroid distance is more robust (Zhang and Lu, 2001). On the other hand, very limited techniques have been applied to a small range of shapes to ensure successful results. In this case, polygonal approximation is very appropriate for discriminating the images of maple leaves (Im *et al.*, 1998). They tried to recognize species in the maple-family (Acer) and were able to identify Nine maple species successfully (Figure 2.1). According to the overall leaf shape of these species, first the leaves were classified into the two groups,

(i) Species which have more than five apices (Figures 2.1a to 2.1f)

(ii) Species which have three apices (Figures 2.1g to 2.1i).

In the next step, within these two groups, the leaves were further classified into individual species using leaf shape features.

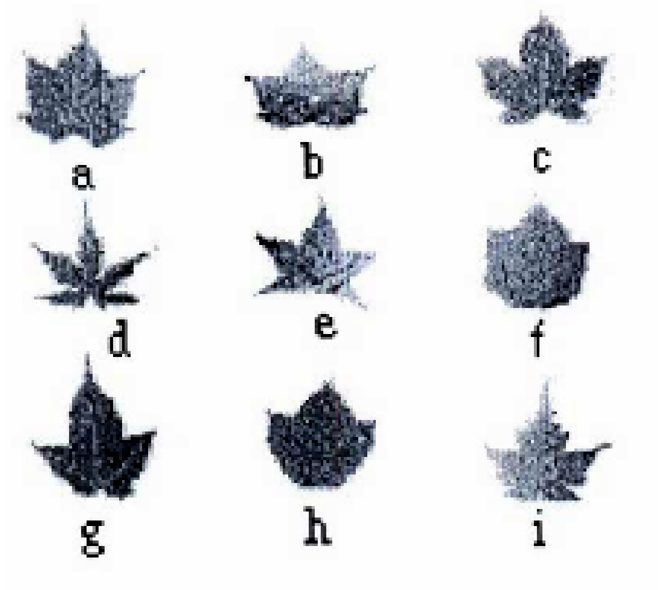

Figure 2.1 : Nine maple leaf shapes of the study

With 400 images from 40 different Chrysanthemum varieties, Abbasi *et al.* (1997) introduced a semi-automatic method for leaf classification, based on leaf shape. Their method finds the most similar class to an input image and the final decision was done by the user manually. Four classes of images, each with 5 sample images are shown in Figure 2.2. They had 40 different classes and each class contained 10 sample images.

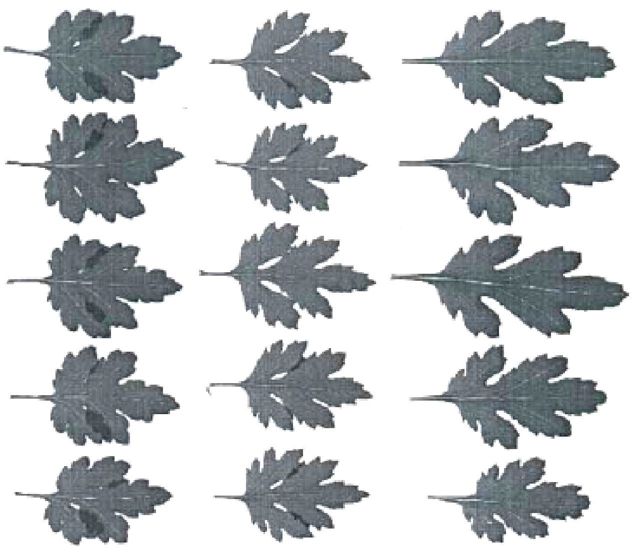

Figure 2.2: Four classes of Chrysanthemum leaf images

Mündermann *et al.* (2003) worked with lobed leaves (such as oak) and proposed a method to model those types of leaves using 2D leaf silhouettes as inputs to their system. Using shape as the main biometric, Wu *et al.* (2006), Zhang *et al.* (2004), Wang *et al.* (2003), Wang *et al.* (2002) are the other authors who have tried to develop leaf retrieval systems for a whole range of leaves (without any prior categorization). One of the difficulties in getting leaf shape as a biometric is the presence of self intersection leaf parts. Mokhtarian and Abbasi (2004) addressed this issue which is a common problem in 2-dimensional shape representation analysis. The following Figure 2.3 illustrates how this can be affected in shape based leaf image retrievals. Figure 2.3(a) represents the gray level image, Figure 2.3(b) shows the boundary of object without considering self-interaction and Figure 2.3(c) shows the defined new boundary of the object.

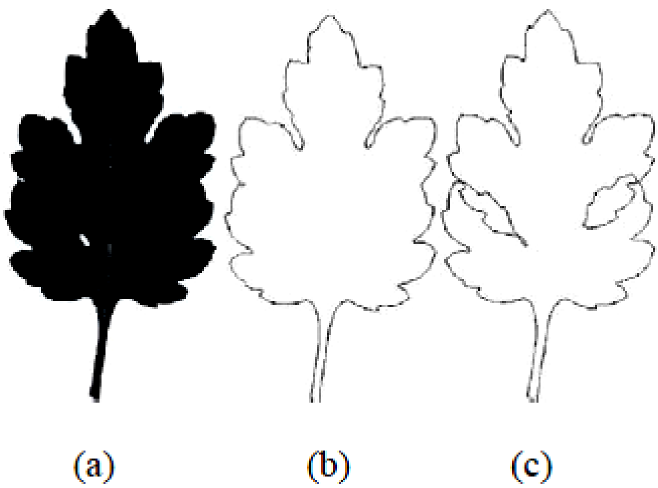

Figure 2.3: An example of self interaction

2.4.2. Region based classification

Lee and Chen (2003) proposed region-based classification method for leaves, where different features are extracted from leaves and are used for leaf recognition and classification. The features used are the aspect ratio, the horizontal and vertical projection. In addition, a feature valuation was proposed to analyze and compare the discrimination ability of various features. Therefore an appropriate set of feature can be selected for leaf recognition and classification. The system was compared with a contour based method and it was found that the region-based method was much superior to the contour based system. The region based system produced a 82.33 percent classification accuracy whereas it was only 37.67 percent with contour based system. The recall rate with 10 returned images was around 48.27 percent as opposed with a 21.7 percent from contour based system. The results clearly shows that region based algorithms have high prospects with leaf classification.

In another work proposed by Wang *et al.* (2003), a region-based binary tree representation incorporating adaptive processing of data structures was proposed to address this problem. After segmentation, to characterize its contents by using region merging method a binary tree was established. Next, with binary tree representation, an adaptive processing of data structure algorithm was used to perform the classification task. Experimental results on seven categories of images showed that this region-based structural representation is superior to other methods.

A new image classification approach through a tree-structured feature set is proposed. In this approach, the image content is organized in a two-le`v`1el tree, where the root node at the top level represents the whole image, and the child nodes at the bottom level represent the homogeneous regions of the image. The tree-structured representation combines both the global and the local features through the root and the child nodes. Features from the tree-structured representation was then processed using a two-level self-organizing map (SOM), which used an unsupervised SOM and supervising concurrent SOM (CSOM) classifier for processing image regions and overall classification of images respectively. The proposed method extracted both global (image-based features) and local (region-based features) to enhance the operation classification. The results showed that the proposed approach performed better than traditional algorithms.

2.5. CHAPTER SUMMARY

This chapter presented the various studies focusing the four steps of CAP-LR. From the literature survey, it can be understood that while many proposals have been published to tackle the problem, noise, contrast and weak edges, the research still is immature. Most of the proposed solutions concentrate on only one of these enhancement operations. Applying each enhancement operations separately consumes time and there is a possibility of missing one or more operations during enhancement. To solve these drawbacks, a single model that can automatically remove noise, enhance edges and adjust contrast in a leaf image is very much desired.

Although several techniques have reported good segmentation results, several disadvantages of the reviewed methods still exist. For example, several techniques require user intervention to define the region of interest and most of them are parameter dependent. Thus, to segment leaf from its background, automatic methods that are non-parametric and does not rely on user intervention is required.

While considering leaf extraction and selection, though several features like wavelet features and edge features have been used to classify leaf images, the optimal feature set which can accurately identify plants and recognize leaves is still sparse. Hence, more research is required to be developed in a innovative manner for feature extraction and selection.

From the survey, it can be understood that, according to features utilized in object recognition, past research can be broadly classified into two categories: contour-based and region-based approaches. Most plant recognition methods used contour-based features. The disadvantage of the contour-based features is that it is hard to find the correct curvature points. Region-based approaches, on the other hand, are more robust. From the survey it is seen that not much work has been performed using region based features, but has great potential in the field of leaf recognition for plant classification.

Similarly, inspite of various machine learning algorithms available, the search for a fast and accurate classifier is still an on-going process. This research work proposes algorithms that enhance the operation of various processes of leaf recognition for plant recognition. The methodology used for designing such a system is presented in the next chapter, Chapter 3, **Methodology**.