SEGMENTATION BASED IMAGE MINING FOR AGRICULTURE

Dr.Hari Ramakrishna,
T.S.Praveena,
Prof. N.Rama Devi,
Department of Computer Science and Engineering,
Chaitanya Bharathi Institute of Technology Gandipet- Hyderabad, A.P., INDIA -500075

Abstract
The paper presents the use of Computer Aided techniques in agriculture. This paper reviews some of the important segmentation based algorithms and recent trends in image processing techniques which are applicable to crop quality evaluation and defect identification. Issues related to crop quality evaluation are discussed. A case study of medical image segmentation which adopts similar methodology for the detection of diseases is discussed. The results obtained on two sample medical images are enclosed. The paper strongly suggests the need to use similar Methodology of Information Technology for the agricultural domain.

I INTRODUCTION

Countries like India and Sri Lanka give prime priority to agriculture as most of the people depend more on agriculture. The communication and information technology is focusing more on support techniques to channelize the modern technology for meeting the agricultural challenges. Several domains like Agricultural Information systems and Web based Agricultural Information System can be seen in this context.

Web based information and advisory Systems for improving the quality and productivity for agricultural processes are emerging [9]. Intelligent Process Controlling Models are available to monitor process in a distributed environment. Such Models help us to monitor cultivation over distributed environment [7]. The process of cultivation for optimum yield and quality can be achieved with the support of such techniques.
Agricultural field needs computer based techniques for crop quality evaluation and defect identification. Image mining is one of such techniques suitable for agricultural applications.

The management of perennial fruit crops requires close monitoring especially for the management of diseases that can affect production significantly and subsequently the post-harvest life.

The image processing and image mining can be used in agricultural applications for following purposes:

i). Identification of diseased leaf, stem, fruit
ii). Identification and quantification of affected area by disease.
iii). Identification of intensity of diseases and their effect on productivity.

In case of plant, the disease is defined as any impairment of normal physiological function of plants, producing characteristic symptoms. A symptom is a phenomenon accompanying something and is regarded as evidence of its existence. Disease is caused by pathogen which is any agent causing disease.

In most of the cases, pests or diseases are seen on the leaves or stems of the plant. Therefore identification of plants' leaves stems and finding out the pests or diseases, percentage of the pest or disease incidence, symptoms of the pest or disease attack, play a key role in successful cultivation of crops. It is found that diseases cause heavy crop losses amounting to several billion dollars annually [1][2][3] .

II. IMAGE SEGMENTATION

Segmentation of images is an important technology for image processing. Many applications require image segmentation on synthesis of the objects based on computer graphic images. Preprocessing an Image appropriately for removing noise and artifacts is prerequisite for interpreting the image.

Image segmentation is a process in which regions or features sharing similar characteristics are identified and grouped together. The output of the segmentation is usually a set of classified elements. Image segmentation can be performed on statistical classification such as *thresholding edge detection*, *region detection*, or any combination of these techniques.

Edge-based techniques rely on discontinuities in image values between distinct regions, and the goal of the segmentation algorithm is to accurately demarcate the boundary separating these regions.

Region-based techniques rely on common patterns in intensity values within a cluster of neighboring pixels. The cluster is referred to as the region, and the goal of the segmentation algorithm is to group regions according to their anatomical or functional roles.

Segmentation is a process of extracting and representing information from an image through grouping pixels together into regions of similarity. Region-based segmentation methods attempt to partition or group regions according to common image properties. Sample image properties consist of the following:

i) Intensity values from original images, or computed values based on an image operator

ii) Textures or patterns those are unique to each type of region

Spectral profiles that provide multidimensional image data, elaborate systems that can use a combination of these properties to segment images, while simpler systems may be restricted to a minimal set of properties depending on the type of data available.

Some of the practical applications of image segmentation are:

i) Locate objects in satellite images (roads, forests, etc.)
ii) Face recognition
iii) Fingerprint recognition
iv) Traffic control systems
v) Brake light detection
vi) Machine vision

Discussion and review on different segmentation algorithms and its applications which suit Agricultural images for defect identification are discussed in this paper.

III SEGMENTATION ALGORITHMS

All image processing operations generally aim at a better recognition of objects of interest, i. e., at finding suitable local features that can be distinguished from other objects and from the back ground. The next step is to check each individual pixel to see whether it belongs to an object of interest or not. This operation is called *segmentation* and produces a *binary image*. A pixel has the value one if it belongs to the object; otherwise it is zero. Segmentation is the operation at the threshold between *low-level image processing* and *image analysis*. After segmentation, it is known that which pixel belongs to which object. The image is parted into regions and we know the discontinuities as the boundaries between the regions. The different types of segmentations are Pixel based, edge based and region based.

i) Pixel-Based Segmentation: This is also known as point based segmentation. It *is* conceptually the simplest approach used for segmentation. In this each pixel is labeled so that the image can be analyzed more precisely.

ii) Edge-Based Segmentation: An *edge based segmentation* approach can be used to avoid a bias in the size of the segmented object without using a complex thresholding scheme. Edge-based segmentation is based on the fact that the position of an edge is given by an extreme of the first-order derivative or a zero crossing in the second-order derivative.

iii) Region-based Segmentation: These methods focus on an important aspect of the segmentation process missed with point-based techniques. There a pixel is classified as an object pixel judging solely on its gray value independently of the context. This means that isolated points or small areas could be classified as object pixels, disregarding the fact that an important characteristic of an object is its *connectivity*. The procedure is to limit the mask size at the edge to points of either the object or the background. This can be achieved only if we can distinguish the object and the background after computation of the feature.

Obviously, this problem cannot be solved in one step, but only iteratively using a procedure in which feature computation and segmentation are performed alternately. In the first step, the features are computed disregarding any object boundaries. Then a preliminary segmentation is performed and the features are computed again, now using the segmentation results to limit the masks of the neighborhood operations at the object edges to either the object or the background pixels, depending on the location of the center pixel. To improve the results, feature computation and segmentation can be repeated until the procedure converges into a stable result [4][5].

IV MODEL-BASED SEGMENTATION

All segmentation techniques discussed so far in this paper utilize only local information. The human vision system has the ability to recognize objects even if they are not completely represented. It is obvious that the information that can be gathered from local neighborhood operators is not sufficient to perform this task. Instead specific knowledge about the geometrical shape of the objects is required, which can then be compared with the local information. This thought model leads to *model-based segmentation*. It can be applied if the exact shape of the objects contained in the image is known. Model-based segmentation has different types such as Color image segmentation, gray level image segmentation and text segmentation.

i) Color Image Segmentation algorithm

The human eyes have adjustability for the brightness by which we can identify dozens of gray-scale at any point of complex image. Accordingly, human eye can identify thousands of colors. In many cases, gray-Level information cannot extract the target from background, therefore, it should be done by means of color information. The color image segmentation is also widely used in many multimedia applications, for example; in order to effectively scan large number of images and video data in digital libraries, they all need to be in the compiled directory. Sorting and storage, the color and texture are two most important features of information retrieval based on the content of the images and video.

Therefore, the color and texture segmentation are often used for indexing and management of data; another example of multimedia applications is the dissemination of information in the network. Today, a large number of multimedia data streams sent on the Internet, However, due to the bandwidth limitations; we need to compress the data, and therefore it calls for image and video segmentation [6].

ii). Gray-scale Image Segmentation

The segmentation of image raster data into connected regions of common gray-scale has long been seen as a basic operation in image analysis. In texture analysis, just this type of segmentation is possible after individual pixels in an image have been labeled with a numeric classifier. In geographic information systems (GIS) this segmentation is usually followed by the production of a vector representation for each region.

The original algorithm for segmentation, developed by Rosenfeld-pfaltz, describes a two pass 'sequential algorithm' for the segmentation of binary images. The key feature of the Rosenfeld-pfaltz algorithm is that the image is raster-scanned, first the forward direction, from top left to bottom right, then backwards. During the forward pass, each pixel is located in a region label, based on information scanned through; the regions so demarcated may have pixels with more than one label therein. During the backwards pass, a unique label is assigned to each pixel. Hence this classic algorithm can be described as a two pass algorithm.

Cohen presented a one-pass algorithm for segmentation of gray-scale images. Cohen's single pass algorithm proceeds to label pixels on a forward pass, exactly as in Rosenfeld-Phltz, except that during the same forward pass, a lookup- table is prepared, that reveals the connectivity of pixels that have different labels. For most purposes, the labeling of the pixels found the first pass, combined with the look-up-table, provides a complete segmentation, but to actually output an image with unique pixel labels in each region, a pass through the image using the look-up table is required.

A new parallel region segmenting and labeling algorithm is available, that is applicable to gray-scale images, and is appropriate to coarse scale parallel programming. The key feature of this algorithm is the geometric splitting of the image into rectangular blocks, with one pixel overlap at joins. Then using Cohen's one pass algorithm, each region is separately labeled. Then by examination of the overlap regions, the connectivity of different region labels is determined, through connectivity tables, and finally the overall image is completely segmented into connected domains of common gray-scale. The parallelizable algorithm for the segmentation of gray-scale images involves performing the one-pass algorithm on rectangular sub-images, with a single row or column overlap. What is thus produced is a label for each image pixel, together with connectivity for each region. Then, from the overlap rows and columns produce overlap tables, showing how labels in each region are related.

iii) Text Segmentation: It is well known that text extraction, including text detection, localization, segmentation and recognition is very important for video auto-understanding. Text segmentation in video images is much more difficult than that in scanning images. Scanning images generally has clean and white background, while video images often have very complex background without prior knowledge about the text color. Although there have been several successful systems of video text extraction, few researchers specially study text segmentation in video images deeply. The used strategies could be classified into two categories:

(i) Difference (or top-down) and
(ii) Similarity based (or bottom-up) methods.

The first methods are based on the foreground background contrast. Some of the examples for this are fixed threshold value method, Otsu's adaptive thresholding method, global & local thresholding method, Niblack's method. In general, they are simple and fast, but fail when foreground and background are similar.

In the other category namely the similarity based method, clusters pixels with similar intensities together. For example, Lienhart uses the split & merge algorithm, Wang et al. combine edge detection, watershed transform and clustering however, these methods are unstable, since they exploit many intuitive rules about text shape.

As an alternative for these two methods is "Chen et al", which convert text pixel clustering to labeling problem using Gibbsian EM algorithm. This method is effective but more time consuming. The main problem of most existing methods is that they are sensitive to text color, size, font and background clutter, since they simply exploit either general segmentation method or some prior knowledge.

These problems are resolved in a new algorithm which is based on stroke filter. It discovers the intrinsic characteristics of text and designs a robust algorithm especially for text segmentation. In this algorithm stroke filter (SF), describes the intrinsic characteristics of text in terms of scale, orientation and response, a stroke feature based text polarity determination method is available and local region growing method for segmentation refinement based on stroke features and global & local statistic similarities.

V MEDICAL IMAGE PROCESSING

Image processing and mining algorithms are already in use for several applications which are similar to that of agriculture problems. For example these applications are successfully in use for identification of defects and faulty areas in Medical Imaging. Some of them are listed below:

i) Locate tumors and other pathologies
ii) Measure tissue volumes
iii) Computer-guided surgery
iv) Diagnosis
v) Treatment planning
vi) Study of anatomical structure

Different application domains in the medical imaging within the medical imaging domain, two extreme situations may be experimented and a large variety of intermediary cases which may be more or less interpolated from these cases. At one extremity, one family of applications covers the transmission of medical documents over public networks, like in telemedicine, remote or collaborative tele-diagnosis or tele-surgery, distance learning and several applications dealing with data base consulting. In this case, the demand is very close from the desiderata of image watermarking as expressed for e-commerce or multi-media applications over open networks. The images will face transmission errors and lossy compressions; the protocols will probably be heterogeneous; secure modalities (firewalls, accredited software) will be rare and the end-terminals will not be secured. Under these circumstances, we find watermarking solutions very similar to the many developed in other domains of image communication, the basic constraints of medical images being taken as additional guide-lines for selecting the watermarking method.

On the other extremity, within the hospital network and under the complete security system developed in the framework of HIS, very different problems will be faced. Transmissions are done most of the time without loss and specific workstations with adequate protocols and software may be available for handling local security problems. Under these conditions, security problems may only arise from either malicious attempts to break the security protocol, or human negligence or mistakes.

Medical images are usually stored for diagnosis, database or long term usage. They enclose information with different levels of importance. The most important information is located in areas called Regions of Importance (ROI). In the storing operation, the ROI must not be subjected to any distortion. For this purpose initially image segmentation is performed to differentiate between important regions and the other less important ones.

Methodology

Image pre-processing for medical images is generally done by converting the host images into binary images that separates Region of Importance and Non-important Regions as the problem is to identify the default region of the host image for example tumor affected in case of brain image .

The host images after performing morphological operations and pre-processing functions are subjected to segmentation with selected suitable segmentation algorithms. One of the famous algorithms used for such applications is Otsu Algorithm.

VI OTSU' ALGORITHM

Otsu Algorithm is a thresholding algorithm. Otsu's method is employed to perform histogram shape-based image thresholding. The algorithm assumes that the image to be thresholded contains two classes of pixels (e.g. foreground and background) then calculates the optimum threshold separating those two classes so that their combined spread (intra-class variance) is minimal. The process adopts the following steps:

i) Compute histogram and probabilities of each intensity level
ii) Set up initial $\omega_i(0)$ and $\mu_i(0)$
iii) Step through all possible thresholds t=1 maximum intensity
Update ω_i and μ_i
Compute $\sigma_b^2(t)$
iv) Desired threshold corresponds to the maximum $\sigma_b^2(t)$

The results of two case studies of two medical images, features of which are very close to agricultural applications are enclosed in Figure [1].

CONCLUSION

The basic segmentation algorithms are discussed in this paper. The most accurate and suitable segmentation algorithm for sample applications (medical images) is the Otsu's Algorithm. The related methodology is discussed. The results of Otsu's Algorithm for medical images are enclosed [8].

The paper suggests the need of customization of similar techniques of modern information technology for the purpose of building GIS base Agricultural information system with suitable knowledge base for the purpose of estimating and improving quality of cultivation.

REFERENCES

[1] Sankaran, Sindhuja / Mishra, Ashish / Ehsani, Reza / Davis, Cristina "A review of advanced techniques for detecting plant diseases ",Computers and Electronics in Agriculture, 72 (1), p.1-13, Jun 2010

[2] Jayamala K. Patil1 , Raj Kumar, "ADVANCES IN IMAGE PROCESSING FOR DETECTION OF PLANT DISEASES", Journal of Advanced Bioinformatics Applications and Research ISSN 0976-2604 Vol 2, Issue 2, June-2011, pp 135-141

[3] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. ALRahamneh "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Applications (0975 – 8887) Volume 17– No.1, March 2011

[4] Krishna Kant Singh1 , Akansha Singh, " Study Of Image Segmentation Algorithms For Different Types Of Images", IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 5,September 2010.

[5] B.Chandra Mohan, S.Srinivaskumar and B.N.Chatterji, A Robust Digital Image Watermarking Scheme using Singular Value Decomposition , Dither Quantization and Edge Detection, ICGST-GVIP Journal, ISSN: 1687-398X , Volume 8, Issue 1, June 2008.

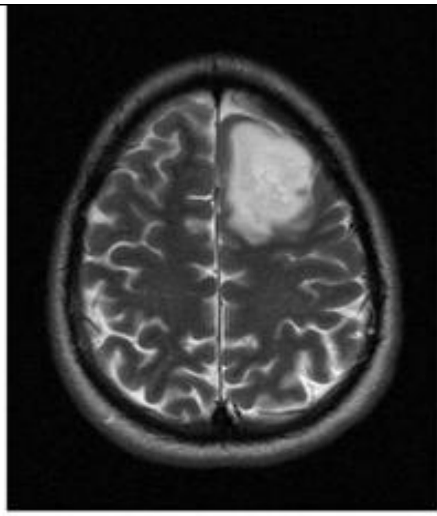
[6] H. Seddik and E. Ben Braiek, Color Medical Images Watermarking, Based Neural Network Segmentation, GVIP Special Issue on Medical Image Processing, March, 2006.

[7] Dr.Hari Ramakrishna ,"Architecture of Intelligent Process Controlling Model through Image Mining" CVR Journal of Science nad Technology Vol 2 , June 2012 ISSN 2277-3916

8] Chandra Mohan.B, Srinivaskumar.S and Chatterji.B.N., A Robust Digital Image Watermarking Scheme using Singular Value Decomposition , Dither Quantization and Edge Detection, ICGST-GVIP Journal, ISSN: 1687-398X , Volume 8, Issue 1, June 2008.

9] Shrikant G. Jadhav, G.N. Shinde, " A Web Based Information & Advisory System", International Journal of Innovative Technology & Creative Engineering (Issn:2045-8711), Vol.1 No.2 February 2011.

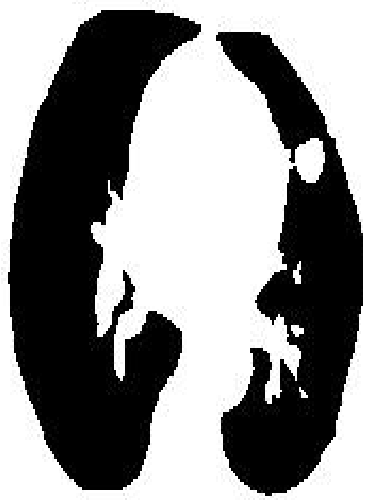
ISSN: 2278-0181
Vol. 1 Issue 6, August - 2012



Host Image A) Brain CT-scan Image

B) Otsu's Segmented Brain CT-scan image

Host Images A) Lung PET-scan Image



B) Otsu's Segmented Lung PET-scan image

Figure: 1 Results of Segmentation of two medical images using Otsu's segmentation algorithm