AUTOMATED FEATURE EXTRACTION IN COTTON LEAF IMAGES BY GRAPH CUT APPROACH

[1]PRASHANT R. ROTHE, [2]RAVINDRA. V. KSHIRSAGAR

[1,2]Department of Electronics Engineering, Priyadarshini College of Engineering, Nagpur, India
E-mail: p_rrothe@rediffmail.com

Abstract- Feature extraction and representation is a decisive step for pattern recognition system. How to extract ideal features that can reflect the inherent content of the images as complete as possible is still a challenging problem in computer vision. Therefore goal of feature extraction is to improve the effectiveness and efficiency of analysis and classification.In this work we present a graph cut based approach for the segmentation of images of diseased cotton leaves.We extract color layout descriptors which can be used for classification. The diseases that have been considered are Bacterial Blight, Myrothecium and Alternaria. The testing samples of the images are gathered from CICR Nagpur, cotton fields in Buldhana and Nagpur district.

Keywords- Cotton Leaf Diseases, Gaussian Filter, Graph Cut, Color Layout Descriptors.

I. INTRODUCTION

Cotton is a principal cash crop in India. Millions of people depend for their living on cotton by way of its farming and processing. With the increasing demand the productivity of cotton become more important. Cotton fabrics constitute a 22% of export from India thus forming a significant part of total export. The naked eye observation is the method used for detection and identification of leaf diseases. The farmers make diagnosis of diseases by their experiences however any mistake leads to incorrect control measurments and excess use of pesticides. Therefore automatic detection of leaf diseases becomes important and it may prove beneficial in monitoring and supervising large crop fields. Also it helps in early detection of disease symptoms as soon as they appear on plant leaves. Such systems include image enhancement, image segmentation, feature extraction and training blocks

II. COTTON LEAF DISEASES

Leaf spots are considerd as important factor indicating the existence of disease and considered as indicator of crop disease (EI-Helly M.et al, 2003). In this work the leaf diseases under study are Bacterial Blight, Myrothecium and Alternaria. The symptoms of these diseases are as:

A. Bacterial Blight

Bacterial blight is also called as angular leaf spot and is highly destructive bacterial disease. It is caused by Xanthomonas campestrispv. malvacearum.

Symptoms: It starts as angular leaf spot with red to brown border. These angular spots appear as water-soaked areas which later on turn dark brown to black. The spots on the infected leaves may spread along the major veins of the leaf. In the later stages of disease leaf petioles and stems get infected and premature fall off of the leaves occur.



Figure 1 Bacterial leaf blight infected leaf

B. Alternaria

It occurs in the wet weather and when the temperature is about 27^0C. It was carried by a fungal pathogen that is present on infected cotton residues of the previous season.

Symptoms

The disease is most severe on lower leaves as compare to upper leaves and may get puzzled with the spots of bacterial leaf blight. Initially small circular brown, grey-brown to tan colored spots of size varing from 1-10mm appear on leaves. Mature spots have dry, dead, grey centres which crack and fall out. Sometimes the old spots combine together to produce irregular dead areas.



Figure 2 Alternaria infected leaf

C. Myrothecium

It is caused by the fungus Myrothecium roridum that survives in soil or plant debris. It occurs in the weather conditions with temprature between 70^0 and 81^0F and moisture on the leaves.

Symptoms

Circular to semicircular light brown to tan colored spots with violet to reddish brown margins appear on the leaves. Later on shield shaped small fruiting bodies are produced in the central portion of the spot. Big shot holes appear in leaves and the center becomes dry and drops off.



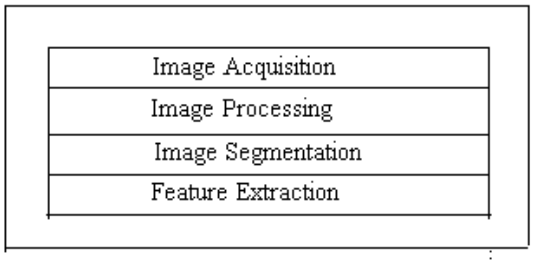
Figure 3 Myrothecium infected leaf.

III. LITERATURE REVIEW

Yan-chang zhang, Han-ping mao, Bo hu and Ming xi-li extracted the morphological features such as area, major and minor axis,orientation, eccentricity, solidity, moment of inertia, hydraulic radius, equiv diameter, ratio of principal axis etc.from images ofinfected disease leaves. They used fuzy curve for selection of significant features and fuzzy surface model to determine the best subset of features. These features are used for identification of Anthracnose, Black spot and Red leaf blight diseases on cotton leaf. P.ravathi and M. hemalatha used edge features and RGB pixel counting features for classification and identification of grey mildew, bacterial leaf blight, leaf curl, alternaria etc. Alexandre A. Bernardes and Jonathan G. Rogeri proposed a system for identification of Ramularia disease, Bacterial Blight, Ascochyta Blight on cotton crop.The input image is decomposed in various color channels like R, G, B, H, S, V, I3a, I3b, and grey levels then DWT is applied to each color channel and the wavelet energy is computed for each sub-band and compose the feature vectors. Support Vector Machine is used for classification. H. Al-Hiary, S. Bani-Ahmad, M. Reylat proposed automatic system for detection and classification of plant diseases. K-means clustering technique is used for segmentation and back propagation algorithm for classification to get the average efficiency of 94.67%.

IV. THE PROPOSED APPROACH

For image classification using vision related approach the basic approach remains the same. The digital images are acquired using a digital camera. Then image processing techniques are applied to these images for segmentation and feature extraction.The discriminating features are used to train the network which performs the classification.



IV. IMAGE ACQUISITION

The images requisite for the experimentation are acquired by using Cannon A460 digital camera and Leica Wild M3C stereo zoom microscope at CICR Nagpur and from the fields in Buldhana and Nagpur district in the state of Maharashtra.

VI. IMAGE ENHANCEMENT

The Gaussian operator is used to blur images and remove noise. It is used for filtering because of the fact that its frequency response is similar to low pass filter therefore it can be used to remove high frequency components from an image.The Gaussian operator is represented by

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

and has the distribution as shown below.

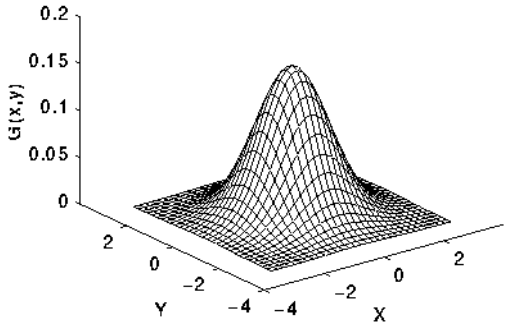


Figure 4 Gaussian distribution

This 2-D distribution is used as a 'point-spread'function for smoothing and this is performed using convolution. However since the image is stored as matrix of discrete values a discrete approximation of Gaussian function is to be performed before the convolution. Theoritically the Gaussian distribution is non-zero everywhere. But in practice it is in fact zero for than about three standard deviations from the mean, hence we can cut the kernel at this point.

The output of Gaussian filter is a 'weighted average' of each pixel's neighbourhood and the average is weighted more towards the value of central pixel.

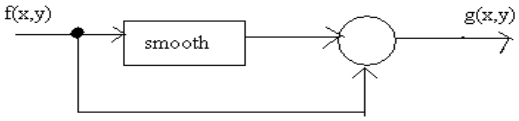


Figure 5 Image filtering

VII. IMAGE SEGMENTATION

Segmentation involves separating an image into regions corresponding to objects by identifying common properties. The graph cut method is used for image segmentation. In this process a directed or undirected graph is partitioned into disjoint sets. A graph G = (V, E) is a set of nodes or vertices V and a set of edges E connecting "neighboring" nodes. We concentrate on undirected graphs where each pair of connected nodes is described by a single edge e = {p, q} ∈ E. The nodes of graph represent image pixels. In addition to these, there are two special terminal nodes refered as S (source) and T (sink) that represent "object" and "background" labels.

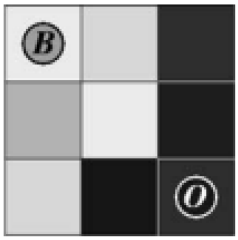
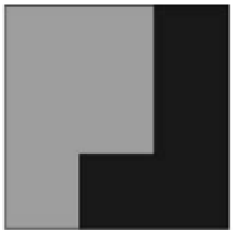
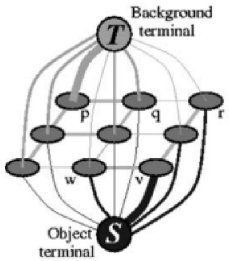


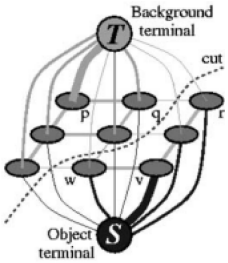
Image amid seeds



Segmentation result



Graph



Cut

The neighboring pixels are interconnected by edges in a regular grid-like fashion. Edges between pixels are called n-links. The edges used to connect pixels to terminals are called as t-links. All graph edges including n-links and t-links are assigned non-negative weight. In the above fig the edge weights (costs) are shown by the thickness of edges.The edge weight between pixel i and j will be denoted by W_{ij}^{I} and the terminal weights between pixel i and the source (s) and terminal (t) as W_i^s and W_i^t respectively and are given by

$$W_{ij}^{I} = e^{\left(-\frac{r(i,j)}{\sigma_R}\right)} e^{\left(-\frac{\|w(i)-w(j)\|}{\sigma_W}\right)} \qquad (1)$$

$$W_i^s = \frac{p(w(i)|i\epsilon s)}{p(w(i)|i\epsilon s) + p(w(i)|i\epsilon t)} \qquad (2)$$

$$W_i^t = \frac{p(w(i)|i\epsilon t)}{p(w(i)|i\epsilon s) + p(w(i)|i\epsilon t)} \qquad (3)$$

An s-t cut is defined as a subset of edges C ⊂ E such that the terminals S and T become completely separated by the induced graph G(C) = (V,E\C). A cut divides the nodes between the terminals. It performs binary partitioning of an underlying image into "object" and "background" segments. Here ‖ . ‖ denotes the euclidian norm, r (i; j) the distance between pixel i and j and λ, σ_R, and σ_W are tuning parameters weighing the importance of the different features. Hence, W_{ij}^{I} contains the inter-pixel similarity, that ensures that the segmentation more coherent. W_i^s and W_i^t describes how likely a pixel belongs to background and foreground respectively. The results of implementing graph cut on the sample of Myrothecium infected cotton leaf are shown below.

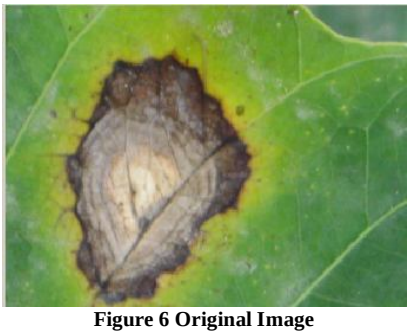


Figure 6 Original Image

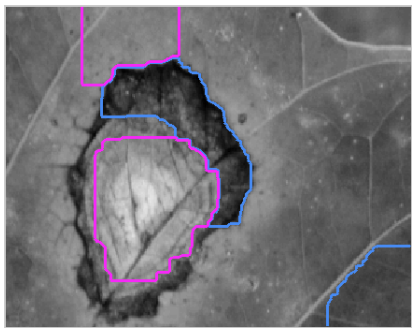


Figure 7 Clustered Image

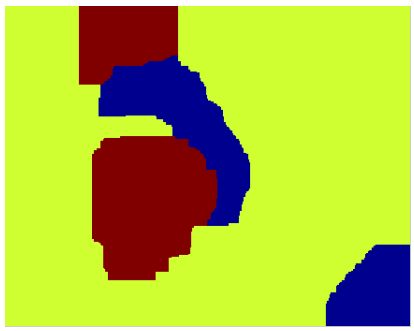


Figure 8 Clustered Image



Figure 9 Segmented Image

VII. FEATURE EXTRACTION

The most basic quality of visual content is color, therefore it is used to describe and represent an image.The MPEG7 standard has suggested procedures to describe the color. One of the methods used to describe the color is the CLD that allows the description between color relation among the sequences or group of images.

The Color layout descriptor is a very compact and resolution-invariant representation of color. It efficiently represents the spatial distribution of colors and can be used for a variety of similarity-based retrieval, content filtering and visualization. It is especially useful for spatial structure-based image retrieval applications. The extraction process of CLD consists of four stages:

- Image partitioning
- Representative color selection
- DCT transformation
- Zigzag scanning

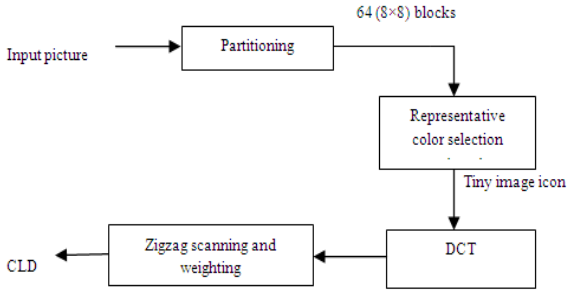


Figure 10 Feature extraction process

In the image partitioning, the input image in RGB color space is divided into 64 blocks so that there is invariance to resolution. After the image partitioning, a single representative color is selected from each block. Then the average of the pixel color in the block is selected as the representative color for that block. The result is an image of size 8×8 to which color space conversion is applied to convert the image from RGB to YcbCr color space. Next 8×8 DCT is applied to obtain three sets of 64 coefficient for luminenceY, crominance for blue and red color each. To calculate DCT the following formulae are used.

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad 0 \le p \le M-1,\ 0 \le q \le N-1$$

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \le p \le M-1 \end{cases} \qquad \alpha_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \le q \le N-1 \end{cases}$$

The result is 3 matrices of size 8×8 representing 64 coefficients for DCTY, DCTCb, DCTCr. A zigzag scanning is performed on these three sets of coefficients, to group the low frequency coefficients of the 8x8 matrix.

These three set of zigzag scanned matrices are the color layout descriptor of the input image.

VIII. RESULT

The tables show the sample of CLDs obtained for the Myrothecium, alternaria and Bacterial leaf blight infected leaves. These descriptors can be used as input for the classifier for identification of diseases.

Myrothecium:

Table 1 Color layout descriptors for Myrothecium

Columns 1 through 11	Columns 12 through 22	Columns 23 through 33		Columns 133 through 143		Columns 182 through 192
-4.1752	-3.7447	-2.9393		-0.1265	⋯	0.0121
-1.7539	-4.1752	1.9714	⋯	-0.3004	⋯	0.0003
-4.1752	-3.3623	-1.8297	⋯	-0.0741	⋯	0
-4.1752	-4.1752	-4.1752	⋯⋯	0.0130	⋯	0
-1.7451	-4.1752	-4.1752	⋯	-0.0010	⋯⋯	0
0.6139	-4.1752	-4.1752		0	⋱	0
-1.9078	-4.1752	-4.1752		0		0
-1.3569	-0.9682	-4.1752		0.0120		0
-4.0936	-2.9187	-4.1752		0		0
-4.1752	-4.1752	-4.1675		0.0162		0
-4.175	-4.1752	1.3072		0		0

Alternaria:

Table 2 Color layout descriptors for Alternaria

Columns 1 through 11	Columns 12 through 22	Columns 23 through 33		Columns 133 through 143		Columns 182 through 192
-0.7166	-0.3631	-0.8524		-0.2757	⋯	-0.4066
-0.6275	-0.3647	-0.7057	⋯	-0.1980	⋯	-0.3308
-0.9391	-0.7439	-0.1880	⋯	-0.1909	⋯	-0.3040
-1.3453	-0.0932	-0.1744	⋯⋯	-0.2686	⋯	-0.1690
-0.9192	0.0482	-0.0265	⋯	-0.3835	⋯⋯	-0.2740
-1.7652	0.0831	-0.6447		-0.1866	⋱	-0.3234
-1.3091	-0.5029	-2.8991		0		-0.4027
-1.3405	-0.4932	-0.4378		-0.3789		-0.1632
-0.3558	-1.0452	-0.0479		-0.4167		-0.0564
-1.6480	-2.8991	-0.4798		-0.3134		-0.1862
-2.8991	-2.7913	-0.6081		-0.3700		-0.0018

Bacterial Leaf Blight:

Table 3 Color layout descriptors for Bacterial Leaf Blight

Columns 1 through 11	Columns 12 through 22	Columns 23 through 33		Columns 133 through 143		Columns 182 through 192
-2.0957	-2.5099	-2.3122	…	0.0663	…	0.0567
-2.2143	-4.1238	-4.1752	…	0.0952	…	0.0775
-2.3134	-3.9343	-4.1752	…..	0.0999	…	0.1527
-2.3591	-4.1493	-4.1752	...	0.0611	….	0
-2.2390	-3.6911	-3.7251		0.0526	.	0
-2.1144	-4.1600	-1.7968		0.0429		0
-2.1904	-4.1752	-1.8677		0.0164		0.0488
-2.8184	-4.1752	-2.3653		0.0612		0.0597
-2.4517	-2.2767	-4.0588		0.0007		0.0606
-2.3595	-3.0484	-4.1752		0.0170		0
-3.8128	-2.7449	-4.1752		0.0002		0

REFERENCES

[1]. Yan-Cheng Zhang, Han-Ping Mao, Bo Hu, and Ming- Xi Li, "Features Selection of Cotton Disease Leaves Image based on Fuzzy Feature Selection Techniques", Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China, 2007.

[2]. P.Revathi, M.Hemalatha, "Classification of Cotton Leaf Spot Diseases using Image Processing Edge Detection Techniques" International Conference on Emerging trends in Science, Engineering and Technology, 2012.

[3] Alexandre A. Bernardes, Jonathan G. Rogeri, Roberta B. Oliveira, Norian Marranghello and Aledir S. Pereira, "Identification of Foliar Diseases in Cotton Crop " Springer Link Lecture notes in computational Vision and Biomechanics, Volume 8, pp67-85,2013.

[4]. P. R. Rothe, R.V. Kshirsagar, "A Study and Implementation of Active Contour Model For Feature Extraction: With Diseased Cotton Leaf as Example" International Journal of Current Engineering and Technology, Vol.4, No.2, April 2014.

[5]. Al-Bashish, D.M.Braik and S.Bani-Ahmed, Detection and classification of leaf diseases using k-means-based segmentation and neural-networks-based classification, Inform.Technol J., 265-275,2011.

[6] Bleschke M.; Madonski R.; Rudnicki R. "Image Retrieval System based on combined MPEG-7 texture and colour Descriptors" International conference on Mixed Design of integrated Circuits & Systems2009, pp 635-639

[7] Jalab H.A., "Image Retrival System based on Color Layout Descriptor and Gabor Filters " ICOS 2011, pp 32-36, 25 Sept.2011.

[8] Khin Hninn Phyu; Kutics A.; Nakagawa A. "Self Adaptive Feature Extraction Scheme for Image Retrieval of Flowers"; International Conference on Signal Image Technology and Internet Based Systems SITIS-2012, pp 366-373, 2012.

[9] Weizheng S.,Yachun W., Zhanliang C., and Hongda W., "Grading Method of Leaf Spot Disease Based on Image Processing" in Proceedings of the 2008 international conference on Computer Science and Software Engineering, Volume 06, December 12-14,2008

[10]. Rampf T., A. K. Mahlein, U. Steiner, E. C. Oerke, H.W.Dehne, L. Plumer, "Early detection and classification of plant diseases with Support Vector Machine based on Hyperspectral reflectance", Computers and Electronics in Agriculture, Volume 74, Issue 1, October2010, Pages 91-99, ISSN 0168-1699.

[11]. Hillnhuetter C., A. K. Mahlein, Early detection and localization of sugar beet diseases: new approaches, Gesunde Pfianzen 60 (4), pp. 143-149, 2008.

[12] Santanu Phadikar and Jaya Sil : " Rice Disease Identification using Pattern Recognition Techn iques", Proceedings of 11th International Conference on Computer and Information Technology (ICCIT 2008)25-27 December, 2008, Khulna, Bangladesh ,pp.420-423,2008.

[13] 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) pp 31-38,Volume 17–No.1, March 2011.

[14] Jinghui Li, Lingwang Gao, Zuorui Shen: "Extraction andanalysis of digital images feature of three kinds of wheat diseases", 3rd International Congress on Image and Signal Processing (CISP2010), pp2543-2548, 2010

★★★