A Farmer Helping System, Disease Detection on Cotton Leaves: CBIR Approach

Pragati Agarwal, Jyoti Chavan, Amruta Ainapure
Department of computer engineering,
Smt. Kashibai Navale College of Engineering,
Vadgoan (BK), Pune

Abstract- **In this paper, we have depict the design and development of a system based on CBIR i.e., content based image retrieval. As there is a explosive growth of the digital images, we will have to execute the competent and effectual retrieval techniques of the image. Within this domain (image processing)scientist has done many researches for identifying and diagnosing in biomedical field.So we have proposed a model in agriculture field which will help farmer for cultivation. This model will be time consuming for farmer. Detecting and diagnosing the pattern of the cotton leaf is one of the important part. There are different types of diseases in cotton leaf. So there is a great need of detecting it diseases with the help of different features such as color, texture, entropy etc. These features are extracted with the help of two different algorithms ie Feature Extraction of HSV color model and Feature Extraction by GLCM and CCM.These algorithms are mainly used to recognize the color and texture of image. The contents of the image are extracted from the pixels to the cotton leaf image, and the further analysis is based on the nature of the image.**

Keywords: ***CBIR, Feature extraction, MY SQL, HSV, GLCM and CCM***

1. INTRODUCTION

Image retrieval technique is helpful in many image processing application. Advances in computer and network technologies coupled with relatively cheap high volume data storage devices have brought tremendous growth in the amount of digital images. Media,Business, government agencies and even individuals all need to organise their images in some way. Because of the amount of collection of digital image increases, the problem finding a desired image in the web application becomes a hard task. There is a great need to expand an efficient method to retrieve digital images. CBIR is known as the hotspot of digital image processing techniques.

Content Based Image Retrieval (CBIR) system works with whole image contents and searching is based on comparison of the contents of the query image. Various algorithms such as HSV, GLCM and CCM for image retrieval features are colour, texture, shape,homogeneity etc. These algorithms are applied to get an image from the image database. They are not concerned about various resolutions of the images, size and spatial colour distribution.

2. CBIR APPROACH

CBIR- Content-based image retrieval has drawn substantial research attention in the last decade. CBIR usually indexes images by low-level visual features such as colour, texture and shape. The visual features cannot completely characterize semantic content, but they are easier to integrate into mathematical formulations. One of the important task in CBIR is extraction of good visual features which efficiently represent a query image.

The term has since been widely used to describe the process of retrieving desired images from a large collection on the basis of features (such as colour, texture and shape) that can be automatically extracted from the images themselves. This approach of CBIR can be used to retrieve images in detection of disease in cotton leaves.

2. DISEASE DETECTION ON COTTON LEAVES-CBIR APPROACH

Within a research of identifying and diagnosing cotton disease in the leaves, the pattern of disease is important part. A variety of features of the images are extracted viz. the colour of actual infected image, texture of the leaf. There are different diseases which occur on the cotton leaf subsequently the leaf colour for different diseases is different, Besides there are various other features related to shape of image, also there are different shapes of holes which are present on the leaf of the image. Generally the leaf of infected image have elliptical shape of holes, thus calculating the major and minor axis is the major task. This information is used to segment cotton leaf pixels within the image, now image which is under consideration is well analyzed and depending upon this software perform further analysis based on the nature of this image.

3.1 Categorization of diseases on cotton leaves

The diseases on the cotton leaves are classified as follow,

a) Bacterial disease: e.g. Bacterial Blight, Crown Gall, Lint Degradation.

b) Fungal diseases: e.g. Anthracnose, Leaf Spot.

c) Viral disease: e.g. Leaf Curl, Leaf Crumple, Leaf Roll.

d) Diseases Due To insects: e.g. White flies , Leaf insects.

There are many diseases which are found on cotton leaves out of them. Some of the major diseases which are observed in large scale have been discussed below:

3.1.1 Foliar leaf spots

As shown in figure, foliar disease[5] is arise due to the potassium deficiency in the plant. In the early stage of foliar disease, leaf shows light spots on it. Above figure shows the final stage of foliar disease, in this stage leaves get plunge on the floor.

3.1.2 Curl Gemini virus

As shown in figure, the curl Gemini virus causes major disease of cotton leaves. Because of the attack of this virus, upward curling, downward curling and thickening of the veins are found in cotton leaves. Due to attack of curl Gemini virus, the yield of cotton drastically reduces.

3.1.3 Bacterial blight

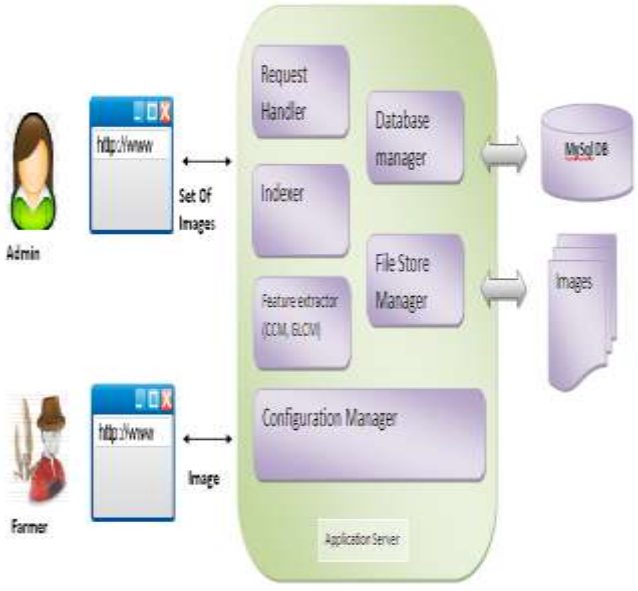
Xanthomonas campestis pv. Malvacearum Bacterial blight starts out as angular leaf spot with a red to brown border. The angular appearance is due to restriction of the lesion by fine veins of the cotton leaf. Spots on infected leaves may spread along the major leaf veins as disease progresses, leaf petioles[5] as shown in above figure. The angular leaf spot, results in premature defoliation and stems may become infected resulting in premature defoliation.

4. How do we extract the feature of image?

In this paper, experimental data set contains 100 images which are being searched and clicked by our own. These images consist of cotton leaves in which different diseases are been shown. By analyzing and giving it complete description about diseases and how to cure it. The two different feature extraction techniques are been used: feature extraction based on the HSV colour space and texture feature extraction technology. At texture feature extraction techniques, we introduce two different extraction methods the gray co-occurrence matrix and CCM.Colour and texture gives us accurate feature extraction of images which helps us to calculate different extraction values.Barely simple features of image information cannot get comprehensive description of image content. We consider the colour and texture features combining not only be able to express more image information ,but also to describe image from the different aspects for more detailed information in order to obtain better results.

5. SYSTEM ARCHITECTURE

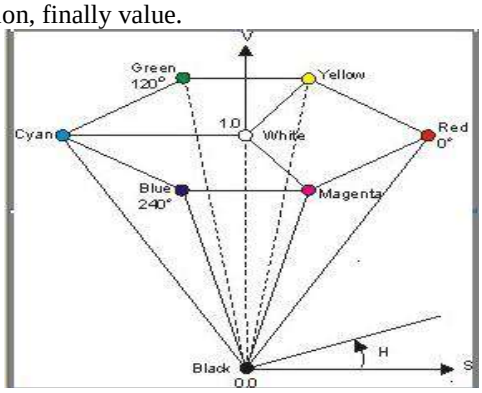
In the above diagram,Farmer uploads a image of cotton leaves ie Query image.The query image is searched in database by the



admin.With the help of different feature extraction method different features and different values are extracted,the required result ie image with disease description is shown to the farmer.

6. HSV

Firstly image was coberted with the help of RGB but for better results ,amny researches are done for finding appropriate results.Generally the image content is consist of colour, texture, shape, homogeneity etc. The different images, shape and texture consist of different angles .HSV colour space must first be quantified.HSV– Hue Saturation Value. HSV colour model forms a uniform colour space, which uses a linear gauge. Hue value is quantized into 16 bits. Saturation is quan tized into 4 bits. Value is quantized into 4 bits. Among those three components, human cognitive about colour is mainly based on hue, and then saturation, finally value.



HSV colour model forms a uniform colour space, which uses a linear gauge. The perceived distance between colours is in proportion to Euclidean distance between corresponding pixels in HSV colour model. This makes suitable for colour based image. For calculating colour histogram in HSV colour space, HSV colour space must be quantified. According to research about colour, three components of HSV must be quantified in a non-uniform manner. Quantized results are coded as

$$F=16H+4S+V$$

F=integer between 0 and 255.

7. TEXTURE FEATURE EXTRACTION BASED ON GLCM

GLCM creates a matrix with the directions and distances between pixels, and then extracts meaningful statistics from the matrix as texture features. GLCM texture features commonly used are shown in the following GLCM is composed of the probability value, it is defined by P (i, jd, θ) which expresses the probability of the couple pixels at θ direction and d interval. When θ and d is determined, P (i, j d,θ) is showed by Pi, j . Distinctly GLCM is a symmetry matrix; its level is determined by the image gray-level. Elements in the matrix are computed by the equation showed as follow GLCM expresses the texture feature according the correlation of the couple pixels gray-level at different positions. It quantification ally describes the texture feature. In this paper, four features is selected, include energy, contrast, entropy, inverse difference

Energy E= $p(x, y)$ ^2

It is a gray-scale image texture measure of homogeneity changing, reflecting the distribution of image gray-scale uniformity of weight and texture

Contrast I = ΣΣ(x − y) 2 p(x, y)

Contrast is the main diagonal near the moment of inertia, which measure the value of the matrix is distributed and images of local changes in number, reflecting the image clarity and texture of shadow depth. Contrast is large means texture is deeper.

Entropy S= ΣΣ p(x, y) log p(x, y)

Entropy measures image texture randomness, when the space co-occurrence matrixes for all values are equal, it achieved the minimum value; on the other hand, if the value of co-occurrence matrix is very uneven, its value is greater. Therefore, the maximum entropy implied by the image gray distribution is random.

Inverse difference H= ΣΣ (1/(1+(x-y)^2)*P(x,y)

It measures local changes in image texture number. Its value in large is illustrated that image texture between the different regions of the lack of change and partial very evenly. Here p(x, y) is the gray-level value at the Coordinate (x, y)

8. FEATURE EXTRACTION BASED ON CCM

Assuming colour image is divided into $N * N$ image subblock, for anyone image sub-block T(i, j) ($1 \le i \le N, 1 \le j \le N$) , using the main colour image extraction algorithm to calculate the main colour C(i, j) . For any two 4-connected image sub-block T(i, j) and T(k ,l) (i − k =1 and j = l ; or j− l = 1 and i = k) , if its corresponds to the main colour and in the HSV space to meet the following condition.

a. Cj And Ci belong to the same colour of magnitude, that is, its HSV components hi = hj , si = s j , vi = v j ;

b. Cj And Ci don't belong to the same colour of Magnitude, but satisfy si ∗ 3 + vi = s j ∗ 3 + v j , and hi − hj = 1; or satisfy hi = hj , si = s j and vi , v j ∈{0,1} . We can say image sub-block T (i, j) and T (k, l) are colour connected. According to the concept of colour-connected regions, we can make each sub-block of the entire image into a unique colour of connected set S = {Ri} ($1 \le i \le M$) in accordance with guidelines 4-connected the set S corresponds to the colour-connected region. For each colour-connected region{Ri}($1 \le i \le M$) , the colour components R, G in RGB colour space and H in HSV colour space are respectively extracted the CCM at the direction δ = 1;θ = 0□ ,45□ ,90□ ,135□ . The same operation is done with I (intensity of the image). The statistic features extracted from CCM are as follows

Energy E=Σ Σ m (i, j) ^2

Contrast I = (i -j) ^2 m (i, j)

Entropy S = −ΣΣ m (i, j) log [m (i, j)]

If m (i, j) = 0, log [m (i, j)] = 0

Inverse difference H = m (i, j)/ (1+ (i-j) ^2)

Through this method, we can get a 16 dimensional texture feature for component R, G, H and I, each component correspond to four statistic values E, I, S and H. F=[FR,FG,FH,FI]=[fRE,fRI,fRS,fRI,…,fIE,fII,fIS,fIH]

9. ALGORITHM OF GLCM AND CCM

1. Take an input image.

2. Resize the input image as per the database defined pixel resolution.

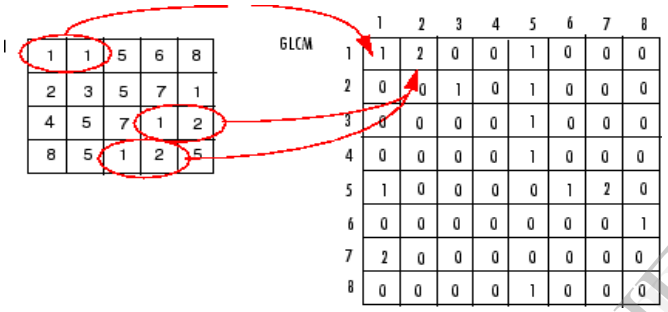
3. Convert input image RGB to Gray Scale,

4. Now, quantize 256 x 256 the matrix given by colour histogram.

5. After Quantization we obtain a matrix of 8X8 (by doing quantization, computation power will be less as we are quantizing 256X256 to8X8)

6. Create a GLCM empty Matrix of (8X8).

7. The above created quantize matrices can be taken in 8 different angels to form a GLCM matrix



10. FUTURE SCOPE

In this paper we have introduced a farmer helping system by detecting diseases on cotton leaves.Mainly researches are done on biomedical fields.So we were inspired doing some work in field of agriculture. In future farmer can also work on different types of leaves such as fruits, vegetables,plants etc. We have implemented at a small scale by making our own database.In future it can take help from large databases which are available on site ie coral draw database.

11. CONCLUSION

In this paper,we have presented a detection of diseases on cotton leaves by content based image retrieval.We have demonstrated by using a feature extraction methods ie HSV,GLCM and CCM.A CBIR is well known in image processing.With the help of these algorithms we can find a effective and efficient result.

12. REFERENCES

1. Hatice Cinar Akakin and Metin N. Gurcan "Content-Based Microscopic Image Retrieval System for Multi-Image Queries" IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, VOL. 16, NO. 4, JULY 2012.
2. Lining Zhang, *Student Member, IEEE*, LipoWang, *Senior Member, IEEE*, and Weisi Lin, *Senior Member, IEEE"* Semi supervised Biased Maximum Margin Analysis for Interactive Image Retrieval*"* IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 4, APRIL 2012.
3. Jianlin Zhang, Wensheng Zou, "Content-Based Image Retrieval Using Color and Edge Direction Features", IEEE Conference 2011.
4. Xiaoyi Song, Yongjie Li, Wufan Chen,"A Textural Feature-Based Image Retrieval Algorithm", Fourth International Conference on Natural Computation.
5. P.Revathi, M.Hemalatha, "Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques", 2012 - International Conference on Emerging Trends in Science, Engineering and Technology.
6. Chih-Chin Lai, *Member, IEEE*, and Ying-Chuan Chen, "A User-Oriented Image Retrieval System Based on Interactive Genetic Algorithm", IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 60, NO. 10, OCTOBER 2011.