A Content Based Image Retrieval System for diagnosing Agricultural Plant Diseases

[1] Noor Hussain Ladkhan
Lead Software Developer,
Katwa Infotech, Belgaum, Karnataka.

[2] Shaheen Mujawar
Assistant Professor,
Dept of Computer Science & Engineering,
S G Balekundri Institute of Technology, Belgaum,
Karnataka.

[3] Kuldeep Sambrekar
Assistant Professor,
Dept of Computer Science & Engineering,
Gogte Institute of Technology, Belgaum, Karnataka.

***Abstract* - Agriculture sector provides gainful employment and livelihood for majority of the population of India and contributes significantly to the National Income. Farmers are the backbone of the Agriculture sector. Farmers are unable to cope up with the problems that arise due to agricultural plant diseases and they depend on Agriculture Plant Biologists to address these problems. Diagnosing the plant diseases manually by the Agriculture Plant Biologists by looking and examining the diseased plant is a time consuming process. If the Agriculture Plant Biologists are unable to identify the plant disease within the stipulated time the crops get spoiled. Since a single picture is worth a thousand words, it would be a good idea if the Agriculture Plant Biologists are provided with a system which can automatically extract and analyze significant features of the diseased plant by processing the uploaded diseased plant image. This will aid the Agriculture Plant Biologists to diagnose the plant disease easily and tell the farmers on the treatment measures to be taken at the earliest. Recognizing the importance and predominance of Agriculture sector, in this paper a system based on Content Based Image Retrieval for diagnosing the agricultural plant diseases is proposed. MPEG-7 CLD (Color layout descriptor) has been used to create the image descriptors and MPEG-7 CLD distance measure is used to compare the images.**

Index Terms — Content-based image retrieval (CBIR) systems, Color Layout Descriptor (CLD)

I. INTRODUCTION

Agriculture sector is vital for the food and nutritional security of India. The sector remains the principal source of livelihood for more than 58% of the population though its contribution to the national GDP has declined to 14.2% due to high growth experienced in industries and services sectors [1]. Compared to other countries, India faces a greater challenge, since with only 2.3% share in world's total land area, it has to ensure food security of its population which is about 17.5% of world population [1]. Agriculture along with fisheries and forestry accounts for one-third of the nation's Gross Domestic Product (GDP) and is its single largest contributor. Agricultural exports constitute a fifth of the total exports of the country.

Farmers are the main force behind the Agriculture sector. Agricultural plant diseases turn into dilemma for farmers and can cause significant reduction of the quality and quantity of the agriculture products. The diseases in plants are caused by several agents and affect different parts. Farmers depend on Agriculture Plant Biologists for addressing these problems. Sometimes Agriculture Plant Biologists are also unable to diagnose the disease that results in lack of identification of right type of disease and this leads to crop losses if not taken care of at right time. Hence there is a need to develop an expert system based on CBIR which can help the Agriculture Plant Biologists in diagnosing the plant disease at the earliest and tell the farmers of the corrective measures to be taken to prevent the crop or plant loss. Present work is an attempt towards building a prototype of an expert system based on CBIR using MPEG-7 CLD [3] for diagnosing Agricultural plant diseases by processing the uploaded diseased plant image.

Content based image retrieval (CBIR) has been used in [4][5][6][7][8] to detect various diseases related to agricultural plants. In content-based image retrieval (CBIR) system, images are indexed by their visual content such as color, texture, shape etc. and the desired images are retrieved from a large collection, on the basis of features that can be automatically extracted from the images themselves. Basically, most CBIR systems work in the same way: A feature vector is extracted from each image in the database and the set of all feature vectors is organized as a database index. At query time, a feature vector is extracted from the query image and it is matched against the feature vectors in the index. The crucial difference between the various systems lies in the features that they extract and in the algorithms that are used to compare feature vectors. The block diagram of a basic CBIR system is shown in Fig. 1.

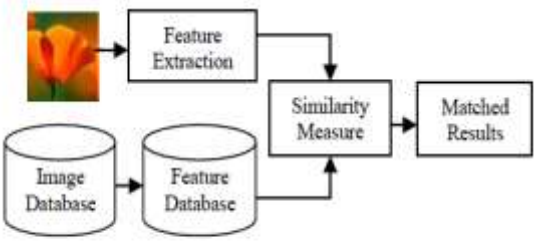

Fig 1. Block diagram of content-based image retrieval system

Because of high demand for searching image databases of ever-growing size, CBIR is becoming very popular. As speed and precision are important, it is needed to develop a system for retrieving images that is efficient. The remainder of the paper is organized as follows Section II discusses the related work. Section III provides the system design details. Section IV gives the implementation details. Section V gives the experimental results and Section VI gives the conclusion and future work.

II. RELATED WORK

MPEG-7 CLD (Color Layout descriptor).

The CLD is designed to capture the spatial distribution of color in an image or an arbitrary-shaped region [3]. The spatial distribution of color constitutes an effective descriptor for sketch based image retrieval, content filtering using image indexing, and visualization. The CLD is a compact descriptor that uses representative colors on an grid followed by a DCT and encoding of the resulting coefficients [3]. The feature extraction process consists of two parts; grid

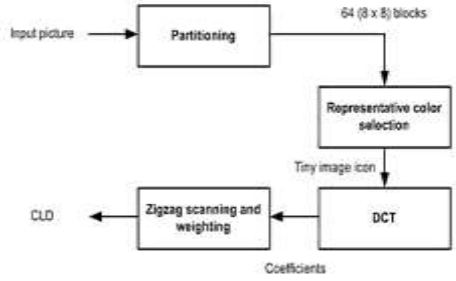

Fig 2: Extraction process of the Color Layout Descriptor

based representative color selection and DCT transform with quantization. An input picture is divided into 64 blocks and their average colors are derived [3]. Note that it is implicitly recommended that the average color be used as the representative color for each block. This partitioning process is important to guarantee the resolution or scale invariance. The derived average colors are transformed into a series of coefficients by performing DCT. A few low-frequency coefficients are selected using zigzag scanning and quantized to form a CLD. The color space adopted for CLD is YCrCb [3].

For matching two CLDs, {DY, DCb, DCr} and { DY', DCb', DCr' }, the following distance measure is used:

$$D = \sqrt{\sum_i w_{yi}(DY_i - DY_i')^2} + \sqrt{\sum_i w_{bi}(DCb_i - DCb_i')^2} + \sqrt{\sum_i w_{ri}(DCr_i - DCr_i')^2}$$

Here, DYi, DCri, and DCbi represent the ith DCT coefficient of the respective color components. The distances are weighted appropriately, with larger weights given to the lower frequency components.

III. SYSTEM DESIGN

In this section the architecture of the system is explained. The block diagram of the architecture is shown in figure 3.

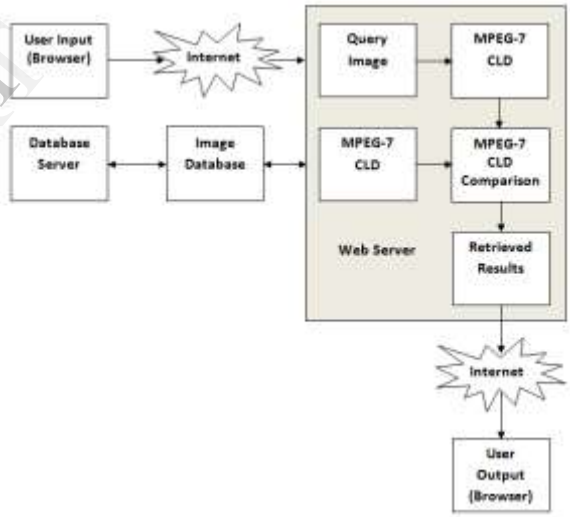

Fig 3: System Architecture

The system architecture enables its usage from a standard web browser. The user can upload the image after proper authentication using the provided GUI (browser).

The database layer provides the storage capability to the proposed system to store the image related details such as Image name, pathogen causing the disease, symptoms of the disease, treatment for the disease and various other details. The database also stores the user authorization and access rights details. The system also has a repository for images.

The system works in two phases. In the first phase the user uploads the desired image. As the image is uploaded a image query is built and the MPEG-7 Color layout descriptor (CLD) is generated using the system logic. In the

second phase the uploaded image CLD is compared with each of the MPEG-7 Color layout descriptors (CLD) of the Images stored in the Image database. Then the disease details are fetched from the database server depending on the matched image. Finally the image along with the disease details and other related information is displayed to the user.

IV. IMPLEMENTATION

For implementation of the proposed work Microsoft Visual Studio 2010 IDE and Microsoft Sql Server 2008 has been used. The application logic layer is developed using Microsoft C#.Net 4.0 and Asp.Net 4.0 (Active Server Pages) and can be deployed on any Web Server that supports Microsoft Dotnet Framework 4.0 and IIS 6.0 and higher hosting server. The application logic layer handles interaction among the users and the system. The system logic uses the classes provided by Microsoft C#.Net 4.0 to implement the Contend Based Image Retrieval method and Color Layout Descriptor (CLD). The database has been designed using the SQL Server 2008 Management studio.

V. EXPERIMENTAL RESULT

Interface for the application is shown below in Fig 5.

Fig 4. Interface for the application.

Wheat images dataset has been used to test the system. For the below given image input,

Fig 5. Diseased plant Image

The system produces the output results as shown in Fig 6. First the user logs in and uploads the diseased plant image using the GUI shown in Fig 6. The system then produces the CLD for the input image and stores it in the memory. Then the system produces CLD for all the images in the dataset and stores it in the memory. Then it compares the input image CLD with all the CLDs of dataset one by one. After the match is found the results are displayed as shown below in Fig 6.

Fig 6. Experimental Input & Results

Based on the experiments conducted, the results indicate that Color Layout Descriptor is very much effective in retrieving similar images. This technique can retrieve similar images with query image from the image database.

VI. CONCLUSIONS AND FUTURE WORK

This paper uses Color layout Descriptor (CLD) technique to retrieve agricultural plant images. This technique is effective based on experiment that has been done. This technique has a potential to be improved with added new function or algorithm. Also the retrieval speed can be increased if the CLDs of the images are stored in xml files.

REFRENCES

1. Ministry of Agriculture – India. Website www.agricoop.nic.in
2. M. Lux and S. A. Chatzichristofis, "Lire: Lucene Image Retrieval - An Extensible Java Cbir Library", ACM International Conference on Multimedia 2008, (ACM MM), Open Source Application Competition, pp.1085-1087, October 27 to 31, 2008,Vancouver, British Columbia
3. B. S. Manjunath, Jens-Rainer Ohm, Vinod V. Vasudevan and Akio Yamada "Color and Texture Descriptors". IEEE Transactions on Circuits and systems for video technology, vol 11, no. 6, June 2001.
4. Kamaljot Singh Kailey and Gurjinder Singh Sahdra, " Content-Based Image Retrieval (CBIR) For Identifying Image Based Plant Disease", International Journal of Computer Technology And Applications, Vol 3 (3), 1099-1104. 2012
5. Sabah and Navdeep, " Remote Area Plant Disease Detection Using Image Processing", IOSR Journal of Electronics and Communication Engineering (IOSRJECE) ISSN : 2278-2834 Volume 2, Issue 6 (Sep-Oct 2012), PP 31-34
6. Vinay S. Mandlik " Agricultural Plant Image Retrieval System Using CBIR", International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459,Volume 1, Issue 2, December 2011
7. Harini and Lalitha. "Identification of Leaf Diseases in Tomato Plant Based on Wavelets and PCA", IEEE, World Congress on Information and Communication Technologies. 2011
8. Viraj A. Gulhane and Ajay A. Gurjar. "Detection of Diseases on Cotton Leaves and Its Possible Diagnosis", International Journal of Image Processing (IJIP), Volume (5) : Issue (5) : 2011.