❏   36

Plant Leaf Disease Image Retrieval Using Color Moments

Jayamala K. Patil*, Raj Kumar**
* Ph.D. Scholar, Bharati Vidyapeeth (Deemed Univ.) Pune, India
** Defence Institute of Advanced Tech., Deemed University, Girinagar, Pune-25, India

Article Info	ABSTRACT
	This paper presents analysis of Content Based Image Retrieval (CBIR) method for plant leaf diseases. The retrieval uses color features to extract the contents of leaf images. The first three color moments i.e.mean, standard deviation and skewness are used to construct color feature matrix. For similarity measurement median value of moment feature vectors is used. The method is analysed for three different color spaces i.e. RGB, HSV & YCbCr to find the the suitability of these color spaces and color features for this problem of retrieval of diseased leaf images of plant.The experimental result shows that HSV color space provides better results for plant leaf disease retrieval. The average precision achived by HSV color space is 43%.

Corresponding Author:

Jayamala K. Patil,
Ph.D. Scholar, Bharati Vidyapeeth (Deemed Univ.) Pune, India

1.   INTRODUCTION

The   recent advancement and development in the internet and   multimedia included images and videos in the data. This enabled creation of large image datasets. Hence in current scenario retrieving useful and accurate information from such huge dataset is a hot research topic. One of the fundamental research area is content based image retrieval (CBIR) [1]. There are two different query categories associated with CBIR [2]:

i)   Query by example: In this images are compared based on low level features like color, shape and texture obtained from the whole image. Instead of image a sketch or a template can be used.

ii)   Semantic retrieval: It uses description of the search concept. It needs higher level of understanding of contents of the image. It uses local features like Scale Invariant Feature Transform descriptor. The different levels of abstraction of the query can be used in such retrieval.

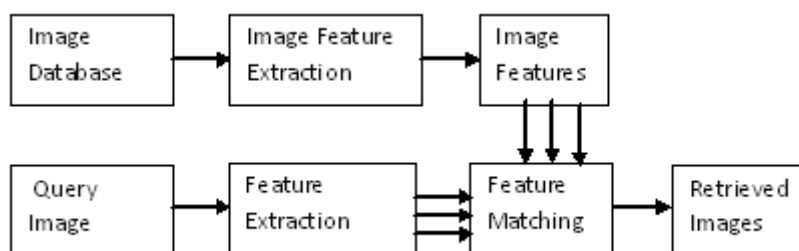The general architecture of CBIR is as shown in figure [1] below.

Figure 1. Architecture of CBIR

CBIR involves execution of two steps i.e. Feature Extraction and Feature Matching. The feature vectors from images are extracted and stored in another database for future use. When query image is given; its feature vector is extracted and compared with the feature vectors in the database. If the distance between the feature vector of the query image and database image is small enough compared with predefined threshold, then the corresponding image in the database is considered as a match to the query image. Then such images are retrieved.

CBIR has wide application area which includes internet, multimedia, medical image archives, and crime prevention. This paper propose the analysis of CBIR for agriculture domain. It is inspired by two aspect in this domain. The first is the difficulties faced by farmers during treatment of their diseased crops and another is the need of plant biologist for efficient computer software to automatically extract and analyze significant content.

As far as the farmer is considered; then for treatement of the crop regarding the disease following two experts are available to guide the farmers:

1. An Excellent Farmers : They precisely catch the change of the crops in the growing process and they manage the cultivation in proportion to the change in order to cultivate the agricultural products of high quality. Since sensing the delicate change of crops is acquired through the observation by the visual sense in their long cultivation experience, it is difficult for them to transmit the understood technique to future generations as a general cultivation one [3].

2. Agricultural Advisor: If farmers decide to take advice from agricultural expert regarding the treatment of incidence of pest /disease/trait to their crop/plant in order to increase the crop productivity then he may face following situations [4]:

i)   Sometimes they have to go long distances for approaching the expert.
ii)   Even though they go such distances expert may not be available at that time.
iii)   Sometimes, the expert whom a farmer contacts, may not be in a position to advise the farmer with the available information and knowledge.

In these cases seeking the expert advice is very expensive and time consuming. Hence electronic expert systems are needed. Electronic expert systems enables farmers in identifying type of diseases; making the right decision and selecting the proper treatment. The expert systems are intelligent computer programs that are capable of offering solutions or advices related to specific problems in given domain, both in a way and at a level comparable to that of human expert in a field. One of the advantages of using Electronic expert systems is its ability to reduce the information that human users need to process, reduce personnel costs and increase throughput. Another advantage of expert system is that it performs tasks more consistently than human experts [5].

In biological science, sometimes thousands of images are generated in a single experiment and can be required for further studies like classifying lesion, scoring quantitative traits, calculating area eaten by insects, etc. Almost all of these tasks are processed manually or with distinct software packages. It is not only tremendous amount of work but also suffers from two major issues excessive processing time and subjectiveness rising from different individuals. Hence to conduct high throughput experiments, plant biologist need efficient computer software to automatically extract and analyze significant content. Here image processing with CBIR (Content Based Image Retrieval) plays a vital role.

Hence we are proposing a development of the CBIR method which will help a remote farmer to treat his crop in the situation of disease according to the advice of the expert who may or may not available at the site of the crop. If the farmer sends the image of the disease to the expert; then expert using the proposed CBIR method can find the type of the disease in very short period (approximately 2-3 minutes) and can suggest the treatment immediately. This will save the time, efforts and money.

2.   PROPOSED METHOD:

This paper presents partial result of the proposed methodology where the features for the CBIR are extracted using color moments.

Color is one of the most widely used   feature in CBIR. [6]-[9]. Color image carries more information than the grey image. Only a limited number of gray levels can be perceived by Human beings but, our eyes are able to distinguish thousands of colors and a computer can represent even millions of distinguishable colors in practice. Color has been successfully applied to retrieve images, because it has very strong correlations with the underlying objects in an image. The color feature is robust to background complications, scaling, orientation, perspective, and size of an image.

The color of the image is represented through some color model. The commonly used color models are RGB (red, green, blue), HSV (hue, saturation, value) and YCbCr (luminance and chrominance); hence for any color image the color contents are characterized by 3-channels from some color model.

- RGB Color Model: The RGB color model is composed of the primary colors Red, Green, and Blue. it is used in color TV monitor and Camera. It has important shortcoming that it is sensitive to illumination changes [2].
- HSV Color Model: Due to its perceptual uniformity another widely used model is HSV color model. The HSV color space is a non-linear transform of the RGB-cube. It is widely used in the field of color vision. The chromatic components hue, saturation and value correspond closely with the categories of human color perception. The HSV values of a pixel can be transformed from its RGB representation according to the following equations [10]:

$$H = \cos^{-1}\left[(R-G) + (G-B)/2\sqrt{(R-G)^2 + (R-B)(G-B)}\right] \tag{1}$$

$$S = 1 - \frac{\min(R, G, B)}{V} \tag{2}$$

$$V = \frac{(R + G + B)}{3} \tag{3}$$

- YCbCr Color Model: The YCbCr color space is widely used for digital video. In this format, luminance information is stored as a single component (Y), and chrominance information is stored as two color-difference components (Cb and Cr). Cb represents the difference between the blue component and a reference value. Cr represents the difference between the red component and a reference value.

The color features can be well described by color histogram [6, 11], color correlogram [12], Color moment [1, 7, 13-15] etc.

2.1. Color moments:

Color moments are measures that can be used to differentiate images based on their features of color [16]. They provide a measurement for color similarity between images. These values of similarity can then be compared to the values of images indexed in a database for tasks like image retrieval.

The basis of color moments lays in the assumption that the distribution of color in an image can be interpreted as a probability distribution. Probability distributions are characterized by a number of unique moments. It therefore follows that if the color in an image follows a certain probability distribution, the moments of that distribution can then be used as features to identify that image based on color.

Stricker and Orengo [7] suggested three central moments of a image's color distribution. They are Mean, Standard deviation and Skewness. As color can be defined by 3 values or channels (e.g. RGB or HSV) moments are calculated for each of these channels in an image. An image therefore is characterized by 9 moments that is 3 moments for each of the 3 color channels.

Let P_{ij} is the ith Color channel at the jth image pixel. Then three Color moments can be defined as :
Moment 1: It is called Mean. It provides average Color value in the image. It is calculated using following statics:

$$Mean(\mu) = \sum_{N}^{j=1} \frac{1}{N} P_{ij} \tag{4}$$

Where N is the total number of pixels in the image.
Moment 2: It is called Standard Deviation. The standard deviation is the square root of the variance of the distribution. It is calculated using following statics:

$$Std.Deviation(\sigma) = \sqrt{(\frac{1}{N}\sum_{j=1}^{N}(P_{ij} - E_i)^2)} \tag{5}$$

Moment 3: It is called Skewness. It gives measure of the degree of asymmetry in the distribution. It is calculated using following statics:

$$Skewness(s) = \sqrt[3]{(\frac{1}{N}\sum_{j=1}^{N}(P_{ij} - E_i)^3)} \qquad (6)$$

2.2. Algorithm to be Executed
1. Find red, green and blue component of color image.
2. Find three moments i.e. mean, standard deviation and skewness for three components of image. And combine them to form nine dimensional feature vector.
3. For feature matching calculate absolute difference in feature vector of query image & feature vector of the database images by using equation (7)

$$Dist(q, d) = \sum_{i=0}^{l-1} (q_i - d_i) \qquad (7)$$

Where, q_i is the feature vector of query image, d_i is the feature vector of the database images and l is the dimension of the feature.

4. Then the distance feature vector is sorted in ascending order and median value is selected as a threshold. Then the most similar images are retrieved by comparison of distance vector with the median value.
Similar steps are executed for HSV & YCbCr Color space by converting RGB image to HSV and YCbCr color space.

3. EXPERIMENTAL RESULTS
The experiment is conducted in MATLAB R2009a using image and signal processing tool box.The database is created of three plant; hence we have three different classes i.e. diseased leaf images of three plants. It consist of 30 leaf images of Maize affected by diseases like gray leaf spots, rust; 50 images of potato affected by diseases like late blight, early blight and leaf roll; and 27 images of tomato affected by late blight and early blight. Figure 2 shows the sample of diseased leaf of potato from a created database.

Figure 2. Diseased Leaf of Potato

The proposed algorithm of CBIR is executed to observe that if query image is of e.g. class 1; then whether it is retrieving most similar images of class 1. Here for experimentation we are putting results of retrieval done for diseased leaf images of potato. The images are self explainatory.

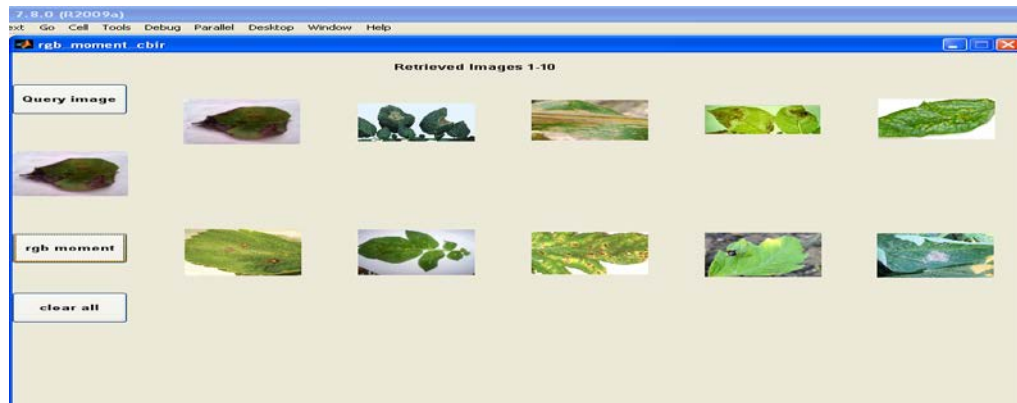- **Sample result of RGB Moment based retrieval:**

Figure 3. RGB Moment based retrieval

Figure 3 shows that in the retrieved images only one image matching to the query image is retrieved. Other retrieved images consist of diseased leaf images of the potato, tomato and maize which are irrelevant to the query. Though other retrieved images are of potato leaf diseases, but they are not relevant to the query image**.**

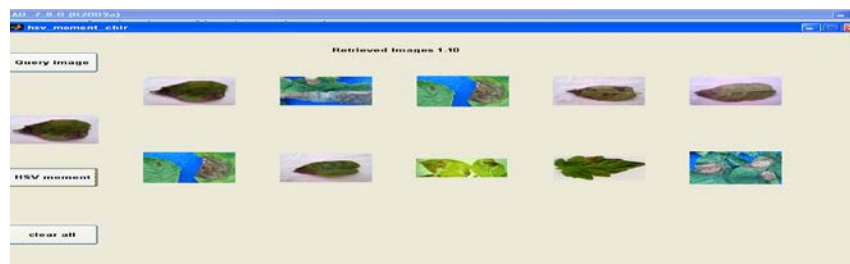- **Sample result of HSV Moment based retrieval:**

Figure 4. HSV Moment based retrieval

Figure 4 shows that in the retrieved images four images (i.e. almost all images of that disease) matching to the query image are retrieved. Other retrieved images consist of diseased leaf images of the potato only. Though other retrieved images are of potato leaf diseases, but they are not relevant to the query image**.** But one advantage of this method is that all retrieved images are of the potato diseased leaf.

- **Sample result of YCbCr Moment based retrieval:**

Figure 5. YCbCr Moment based retrieval

Figure 5 shows that in the retrieved images only two images matching to the query image are retrieved. Other retrieved images consist of diseased leaf images of the potato, tomato and maize which are irrelevant to the query.

Thus the observation of figures (3-5) shows that for moment based retrieval, HSV color space provides better results than RGB & YCbCr color space. The same experimentation is carried for diseased leaf images of tomato and maize. It is found that the observation for potato leaves is valid for tomato & maize leaves also.

Precision & Recall : The quality of CBIR is evaluated by Precision [17]. Precision provides measure of exactness. It is calculated by equation (8).

$$\text{Precision(p)} = \frac{\text{number of relevant images retrived } (R_i)}{\text{total number of images retrived } (T_{ri})} \tag{8}$$

Table 1 shows the comparison of average Precision of retrieval for three classes of diseased leaves. These figures are obtained by observing top 10 retrieved images. The experiment is conducted for randomly selected 10 images of maize, 21 images of potato and 10 images of tomato leaves.

Table 1. Average Retrieval Precision

Image class	Moment based retrieval		
	RGB	HSV	YCbCr
Potato	52.38%	**53.80%**	46%
Maize	34%	**41%**	39
Tomato	30%	**35%**	33%

While a direct comparison is not meaningful, the fact many CBIR problems report precision values around 50–60% [2], can give an idea of the difficulty of the CBIR problem. Moment -based retrieval which is the closest study to our problem, reports precision rates of 53.80% for potato leaves, 41% for Maize leaves and 35% for Tomato leaves at top 10 retrieval ranks.

4. CONCLUSION

The proposed methodology is tested for retrieval of diseased leaves of three plants i.e. Potato, Tomato and Maize and for three different color spaces which are mostly used to present the image. The experimental results shows that the methodogy provides better results for HSV color space as far as precision of retrieval is considered as compared to RGB and YCbCr color space.

Though there is clearly scope for improvement, the proposed approach got promising results for the plant diseased leaves retrieval problem. Hence for further study the color moments can be combined with other features like texture and shape to improve the system performance. Future studies will also include expanding the database in order to make it useful as a practical application. How to combine different features is also an important problem that we plan to study further. Future studies will also include the calculation of the severity of the diseases of the leaves.

REFERENCES

[1]   Zhi-Chun Huang, Patrick P.K. Chan, Wing WY Ng, Daniel S Yeung. "*Content-Based Image Retrieval Using Color Moment And Gabor Texture Feature*". Proceedings of Ninth International Conference on Machine Learning and Cybernetics, Qingdao. 2010: 719-724

[2]   H Kebapci, B Yanikogli, G Unal. "Plant image retrieval using color, shape and texture features". *The Computer Journal*. 2010.

[3]   Hiroya Kondou, Hatuyoshi Kitamura and Yutaka Nishikawa, Yoshitaka Motonaga, Atsushi Hashimoto. "Shape Evaluation by Digital Camera for Grape Leaf".

[4]   MS Prasad Babu and B Srinivasa Rao. Leaves Recognition Using Back Propagation Neural Network-Advice For Pest and Disease Control On Crops. *IndiaKisan.Net: Expert Advissory System.* 2007.

[5]   MT Maliappis, KP Ferentinos, HC Passam and AB Sideridis. "Gims: A Web based Greenhouse Intelligent Management System". *World Journal of Agricultural Sciences*. 2008; 4(5): 640-647.

[6]   M Swain and D Ballard. "Color indexing". *International Journal of Computer Vision.* 1991; 7: 11-32.

[7]   Stricker and M Orengo. "*Similarity of color images*". Proceedings of SPIE Storage and Retrieval for Image and Video Databases III, (San Jose, CA). 1995; 2185: 381-392.
[8]   HJ Zhang and D Shong. "*A scheme for visual feature-based image indexing*". Proceedings of SPIE Conference on Storage and Retrieval for Image and Video Databases III, (San Jose, CA). 1995; 2185: 36-46.
[9]   G Pass and R Zabith. "*Histogram refinement for content-based image retrieval*". Proceedings of of IEEE Workshop and Applications of Computer Vision. 1996: 96-102.
[10] B Chanda, D Dutta Majumder. "Digital Image Processing and Analysis". *PHI publication*. 56-57.
[11] R Brunelli and O Mich. "Histograms Analysis for Image Retrieval". *Pattern Recognition*. 2001; 34(8): 1625-1637.
[12] J Huang, SR Kumar, M Mitra, WJ Zhu, R Zabih. "*Image Indexing Using Color Correlograms*". Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, 1997: 762-768.
[13] A Del Bimbo, M Mugnaini, P Pala, and F Turco Picasso. "*Visual querying by color perceptive regions*". Proceedings of the 2nd International Conference on Visual Information Systems, San Diego. 1997: 125-131.
[14] Jau-Ling Shih & Ling-Hwei Chen. "*Color Image Retrieval Based on Primitives of Color Moments*". Proceeding, Visual'02 Proceedings of the International Conference On Recent Advances In Visual Information Systems. 88-94
[15] Priti Maheswary, Namita Srivastav. "Retrieving Similar Image Using Color Moment Feature Detector and K-Means Clustering Of Remote Sensing Images". *2008 International Conference on Computer and Electrical Engineering*. 2008: 821-824
[16] Noah Keen. "Color Moments". ID: 0341091 Date: February 10, 2005
[17] Shoujue Wang, Hong Qin. "A Study of Order- Based Block Color Feature Image Retrieval Compared With Cumulative Color Histogram Method". *2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery*, *IEEE Computer Society*. 2009: 81-84.

BIOGRAPHIES OF AUTHORS

	Jayamala K Patil was born in 1978 in Mumbai. She has completed post graduation (M.E.Electronics) in 2008 from T.K.I.E.T. Warananagar of Shivaji Universit,y Kolhapur. She is Ph.D. Scholar of Bhatati Vidyapeeth (Deemed Univ.) Pune. Her field of study are Digital Electronics &I mage Processing. She is Associate Professor in Bharati Vidyapeeth's College of Engineering, Kolhapur. She developed Digital Electronics and Video Engineering Laboratories. She has written several technical paper in reputed International Journal and Conferences. She is life member of ISTE.
	Raj Kumar was born on 14th May1963 in Muzaffarnagar U.P., India. He has completed his M.Sc. (Electronics) Degree in 1987 from University of Meerut, Meerut. He has been awarded M.Tech. and Ph.D degree in 1992 & 1997 respectively from University of Delhi, New Delhi. He worked at CEERI Pilani from 1993 to 1994 as a research associate. From May 1997 to June 1998, he worked as Assistant Professor in Department of Electronics and Communications Engg, Vellore College of Engg. (Now VIT), Vellore. He worked in DLRL (DRDO), Hyderabad as Scientist from June 1998 to August 2002 and later on came in DIAT (DU) in Sept. 2002. At present, he is Scientist 'E' in department of Electronics Engg., DIAT (Deemed University), Pune. He established a Microwave and Millimeter Wave Antenna Laboratory in DIAT (DU), Pune and formulated the M.Tech. Programme in the Department of Electronics Engg. In 2010. He has written several technical paper in reputed international Journal and Conferences.