Diseases Classification on Cotton leaves by Advance Digital Image Processing Approach

Viraj A. Gulhane
Lecturer
Department Of Electronics And Telecommunication
Sipna C.O.E.T. Amravati

Ajay A. Gurjar
Professor
Department Of Electronics And Telecommunication
Sipna C.O.E.T Amravati

ABSTRACT
In identifying and diagnosing cotton disease the pattern of disease is the important part. Various features of the image can be extracted viz. color of the infected part and by applying various color windows to the disease image and after that we obtained the vector value for this image, also similar procedure is applied for the normal cotton leaf image and that values compared with one another and vector distance is calculated and depending upon that vector distance the disease is identified and based upon this diagnosis is possible.

Keywords
Texture, eigen feature, classifier, feature extraction, detection.

1. INTRODUCTION
The world textile industries are being ruled by "King Cotton". The antiquity of cotton has been traced to the fourth millennium BC. For over three thousand years (1500 BC to 1700 AD), India was recognized as cradle of cotton industry [1]. India thus enjoys the distinction of being the earliest country in the world to domesticate cotton and utilize its fiber to manufacture fabric. India is India accounts for approximately 25 per cent of world's cotton area and 16 per cent of total cotton production. Maharashtra is the important cotton growing state in India with 31.33 lack hector area and production of 62.00 lack bales (2008-09). The 2nd largest producer of cotton in the world. About 3 million farmers are engaged in cotton cultivation in the state mostly in backward region of Marathwada and Vidarbha[1].

In Vidarbha(Maharashtra) region, cotton is the most important cash crop grown on an area of 13.00 lacks hectors with production of 27 lack bales of cotton (2008-09). Disease on the cotton is the main problem that decreases the productivity of the cotton. The main source for the disease is the leaf of the cotton plant. About 80 to 90 % of disease on the cotton plant is on its leaves. So for that our study of interest is the leaf of the cotton tree rather than whole cotton plant the cotton leaf is mainly suffered from diseases like fungus, Foliar leaf spot of cotton, Alternaria leaf spot of cotton. The machine vision system now a day is normally consists of computer, digital camera and application software. Various kinds of algorithms are integrated in the application software.

Image recognition has attracted many researchers in the area of pattern recognition, similar flow of ideas are applied to the field of pattern recognition of plant leaves, that is used in diagnosing the cotton leaf diseases. There are numerous methods have been proposed in the last two decades [2] which are not fully solved. However these are challenging problems. The critical issue is how to extract the discriminative and stable feature for classification. It is found that Linear subspace analysis has been extensively studied and becomes a popular feature extraction method. We can implement Bayesian maximum likelihood (BML) [4],[5],[6], and linear discriminate analysis (LDA) [7], [8] were introduced into cotton leaf disease classification.

PCA maximizes the variances of the extracted features and hence, minimizes the reconstruction error and removes noise present discarded dimensions. Now to differentiate the two diseased and non diseased image of the cotton leaf the discrimination of the feature is the most important for that LDA is an efficient way to extract the discriminative feature as it handles the within and between class variations separately. But this is the problematic method. So numerous methods have been proposed to solve this problem, a popular approach towards this is Fisher face (FLDA)[9] applies PCA first, in order to reduce the dimensionality and so as to make the within-class scatter matrix non-singular before the application of LDA. But use of PCA discriminative information loss is occurred [10], [11], [12].Direct LDA (DLDA) method [13], [14] removes null space of the between class scatter matrix and extracts the eigenvectors corresponding to smallest Eigen values of the within class scatter matrix. Here modified LDA approach is replaces[15] the within class scatter matrix by total scatter matrix. Subsequent work [16] extracts features separately from the principal and null spaces of the within class scatter matrix.

In this paper we present a approach for cotton leaf image Eigen feature regularization extraction. Image space spanned by the eigenvectors of the within class scatter matrix is decomposed into three subspaces. Eigen features are regularized differentially in these subspaces. Eigen features are regularized differently in these subspaces. Feature extraction and classification and classification could be the last stage.

2. Review of Prior Feature Technique
Various papers are suggesting to diagnosis the cotton leaves using various approach suggesting the various implementation ways as illustrated and discussed below. In the research of identifying and diagnosing cotton disease using computer vision intellectively in the agriculture, feature selection is a key question in pattern recognition and affects the design and performance of the classifier. In previous paper [17], the fuzzy feature selection approach fuzzy curves (FC) and surfaces (FS) - is proposed to select features of cotton disease leaves image. In order to get best information for diagnosing and identifying, a subset of independent significant features is identified exploiting the fuzzy feature selection approach. Firstly, utilize FC to automatically and quickly isolate a small set of significant features from the set of original features according to their significance and eliminate spurious features; then, use FS to get rid of the features dependent on

the significant features. This approach reduces the dimensionality of the feature space so that lead to a simplified classification scheme appropriate for practical classification applications. The results show that the effectiveness of features selected by the FC and FS method is much better than that selected by human randomly or other methods. Also another approach is used to diagnosis the grape leaf disease identification or diagnosis, i.e. paper explaining the grape leaf disease detection from color imaginary using hybrid intelligent system, in that automatic plant disease diagnosis using multiple artificial intelligent techniques. The system can diagnose plant leaf disease without maintaining any expertise once the system is trained. Mainly, the cotton leaves disease is focused in this work. The proposed system consists of concept of Eigen feature regularization and feature extraction technique as explained in introductory part. , in this we try to collect the information of disease leaf by moving RGB color space windows and extracting the values of color space, similar approach is used to apply on the no disease leaf of the image.

3. DISEASE CLASSIFICATION

The diseases on the cotton leaves are classified as

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.

Above disease are dramatically affect the leaf of cotton plant and its leaves. We go through the selective type of diseases on the cotton leaves.

3.1 Red Spots

Fig 1: Red Spots

Red Spots 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 as shown in Fig 1

3.2 Leaf Crumpel

Fig 2: Leaf of Infected Cotton curl upward

Fig 3: of Infected Cotton vain thickening

Cotton leaf curumple (CLCuV) causes a major disease of cotton in Asia and Africa. Leaves of infected cotton curl upward Fig 2. and bear leaf-like enations on the underside along with vein thickening Fig 3. Plants infected early in the season are stunted and yield is reduced drastically. Severe epidemics of CLCuV have occurred in Pakistan in the past few years, with yield losses as high as 100% in fields where infection occurred early in the growing season. Another cotton Gemini virus, cotton leaf crumple virus (CLCrV), occurs in Arizona, California, and Mexico**.**

4. EIGEN SPECTRUM MODELING

From a given set of w-by-h images of leaves, we can from a training set of column image vectors$\{X_{ij}\}$ where $X_{ij} \in IR^{n=wh}$, the ordering of pixel elements of image j of test image i. Now the number of total training sample is $l=\sum_{i=1}^{p} q_i$. For image recogination, each disease leaf image is a class with prior probability c_i. The within class scatter matrix[18] is defined by

$$S^w=\sum_{i=1}^{p} c_i/q_i \sum_{j=1}^{q_i} (X_{ij}-X_i^{'})(X_{ij}-X_i^{'})^T \quad (1)$$

The between class scatter matrix S^b and the total scatter matrix S^t [18]are defined by

$$S^b=\sum_{i=1}^{p} c_i(X_i^{'}-X^{'})(X_i^{'}-X^{'})^T \quad (2)$$

$$S^t=\sum_{i=1}^{p} c_i/q_i \sum_{j=1}^{q_i} (X_{ij}-X^{'})(X_{ij}-X^{'})^T \quad (3)$$

Now where, $X_i^{'}=q_i^{-1}\sum_{j=1}^{q_i} X_{ij}$ and $X_i=\sum_{i=1}^{p} c_i X_i^{'}$
Let S^g, g$\in${t,w,b} represents one of the above scatter matrices. If we regard the elements of the image vector and the class mean vector as a features decorrelated by solving the eigenvalue problem[18].

$$\wedge^g=\Phi^{gT} S^g \Phi^g \quad (4)$$

$\wedge^g$ is the diagonal matrix of eigenvalues $\lambda_1^g,\ldots, \lambda_n^g$ corresponding to the eigenvectors, now we have to stored this eigenvectors in descending order.

5. Eigenfeature regularization and extraction Algorithm

In general, it is desired to extract features that have the smallest within class variations and largest between class variations. That variations can be estimated by finite number of training samples.
Now as we solve the equation 4, which is related to eigenvalue problem.

The proposed eigenfeature regularization and extraction (ERP) approach[18] is summarized below
at the train stage

1. Given a set of cotton leaf images $\{X_{ij}\}$ computing S_w and solving eigen value problem
2. Decomposing the eigenspace of leaf image.
3. Transform the training set X_{ij} into $Y_{ij}^{'}$ and solve the eigenvalue problem as equation (4)
4. Computing $S^{t'}$ to solve eigen value problem.
5. Final feature regularization and extraction matrix obtained[18].

Now at the recognition[18] stage

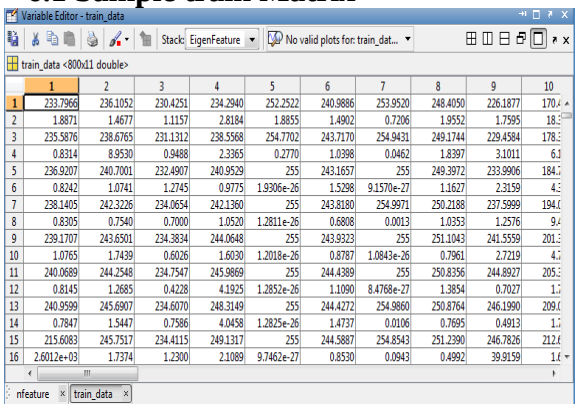
1. Transform each n-D sample images vector X into d-D feature vector F using Feature regularization and extraction matrix obtained in the training stage.

2. Using classifier method[18], Set of images are recognized.

This is eigen feature extraction and regularization algorithm.

6. ANALYSIS AND COMPARISION

6.1 Sample train Matrix



	1	2	3	4	5	6	7	8	9	10
1	233.7966	236.1052	230.4251	234.2940	252.2522	240.9886	253.9520	248.4050	226.1877	170.4
2	1.8871	1.4677	1.1157	2.8184	1.8855	1.4902	0.7206	1.9552	1.7595	18.3
3	235.5876	238.6765	231.1312	238.5568	254.7702	243.7170	254.9431	249.1744	229.4584	178.3
4	0.8314	8.9530	0.9488	2.3365	0.2770	1.0398	0.0462	1.8397	3.1011	6.0
5	236.9207	240.7001	232.4907	240.9529	255	243.1657	255	249.3972	233.9906	184.7
6	0.8242	1.0741	1.2745	0.9775	1.9306e-26	1.5298	9.1570e-27	1.1627	2.3159	4.3
7	238.1405	242.3226	234.0654	242.1360	255	243.8180	254.9971	250.2188	237.5999	194.0
8	0.8305	0.7540	0.7000	1.0520	1.2811e-26	0.6808	0.0013	1.0353	1.2576	9.4
9	239.1707	243.6501	234.3834	244.0648	255	243.9323	255	251.1043	241.5559	201.3
10	1.0765	1.7439	0.6026	1.6030	1.2018e-26	0.8787	1.0843e-26	0.7961	2.7219	4.7
11	240.0689	244.2548	234.7547	245.9869	255	244.4389	255	250.8356	244.8927	205.3
12	0.8145	1.2685	0.4228	4.1925	1.2852e-26	1.1090	8.4768e-27	1.3854	0.7027	1.7
13	240.9599	245.6907	234.6070	248.3149	255	244.4272	254.9860	250.8764	246.1990	209.0
14	0.7847	1.5447	0.7586	4.0458	1.2825e-26	1.4737	0.0106	0.7695	0.4913	1.7
15	215.6083	245.7517	234.4115	249.1317	255	244.5887	254.8543	251.2390	246.7826	212.6
16	2.6012e+03	1.7374	1.2300	2.1089	9.7462e-27	0.8530	0.0943	0.4992	39.9159	1.6

Fig. 4 Sample train Matrix

As shown in figure 4 it is the sample train matrices containg the training values of disease detection and based on this the image which is to be analysed is compare for its disease as shown in figure.

After that values obtained in the train data matrix is converted to 800x100 double format values. That values obtaind are called as modified N sample trained matrix value. Here the database is trained for two disease detection primeraly viz. fungal disease i.e. red spot and leaf crumple.

Here we try to put N=100 image for training so during this training results obtaind is as shown in figure below.

6.2 Disease Detection

Now Here we consider an image as shown in figure 5, these images we choose randomly for analysis point of view from the database of 100 images it is majority found that the cotton leaf is mostly affected by the disease called red spots. as shown fig below during the period of August to December 2011, is about 90% means out of 100 samples, about approx. 90 out of them are highly affected by Red Spots i.e. fungal disease.

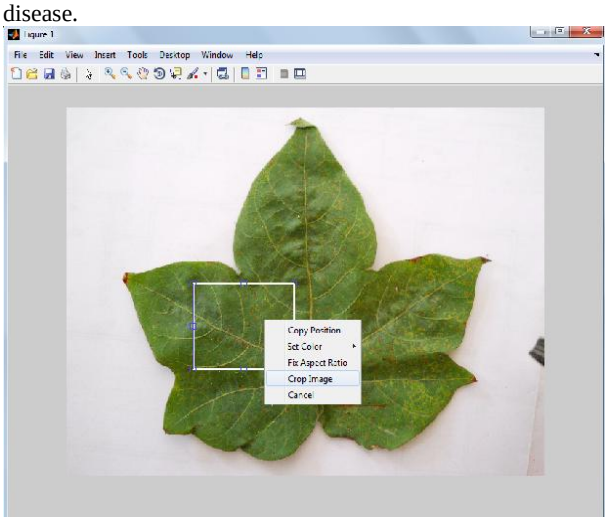


Fig 5. Sample of image

The detection of The red spot Disease is as shown in figure 6 given below.

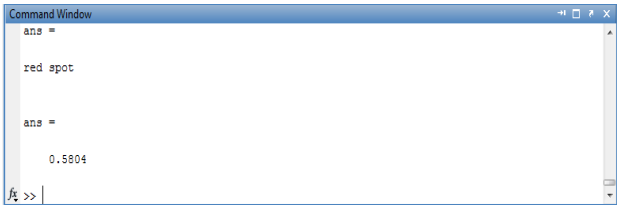


Fig 6. Disease Detection (Red Spot)

Also the another Disease called as leaf crumple is detected on the same leaf is as shown in figure 7 For that we select region of interest as shown in figure 8.

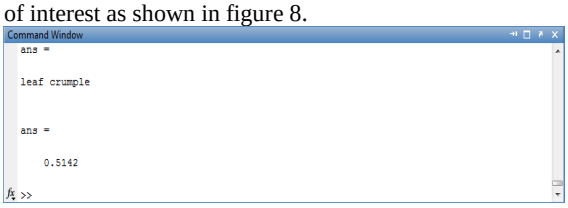


Fig 7: Leaf crumple detection.

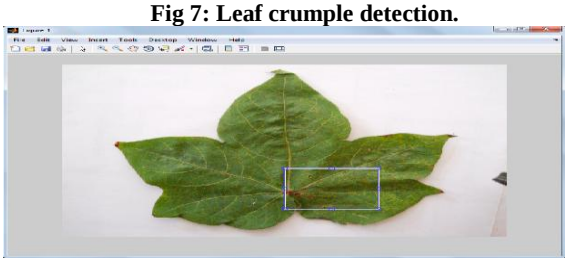


Fig 8: Area Of Interest.

7. CONCLUSION

This paper addresses how the disease analysis is possible for the cotton leaf diseases detection, the analysis of the various diseases present on the cotton leaves can be effectively detected in the early stage before it will damage the whole plant, initially we can be able to detect 2 diseases on the cotton leaves by the methodology of eigen feature regularization and extraction technique.

The result obtained as shown in figure 8 and 9. Motivates us to detect more possible diseases on the leaf of cotton plant. The idea utilized here is having more success rate, than that of the other feature detection methods. Also from this method about 90% of detection of Red spot i.e. fungal disease is detected, it is most dangerous disease, it can highly affect the productivity of the cotton plant in more extent. And if it detects in early stage we can say that, we able to make better productivity.

Here the model presented can able to detect the disease more accurately, The Viderbha Region of Maharashtra state is main producer of cotton, where if this model is applied, we can say that, we can archive good productivity by preventing the various diseases present on the leaves of cotton plant.

8. ACKNOWLEDGMENTS

We are thankful for the cotton farm owner Dashrath Gujar for their supportive nature, also We are very thankful to Pratap Rankhamb who help us to provide the detailed guidance related to various disease on cotton plant. We are thankful to Prof. A. P. Thakare HOD (Department Of Electronics and Telecommunication), Dr. S. A. Ladhake(Principal) Sipna C.O.E.T. Amravati for their encouragement and support.

9. REFERENCES

[1] http://www.pdkv.ac.in/CottonUnit,php(cotten area in viderbha)

[2] Yan Cheng Zhang, Han Ping Mao, Bo Hu, Ming Xili "features selection of Cotton disease leaves image based on fuzzy feature selection techniques" IEEE Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China, 2-4 Nov, 2007

[3] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," ACM Computing Surveys, vol. 35, no. 4, pp. 399-458, Dec. 2003

[4] B. Moghaddam and A. Pentland, "Probabilistic Visual Learning for Object Representation," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 696-710, July 1997

[5] B. Moghaddam, T. Jebara, and A. Pentland, "Bayesian Face Recognition," Pattern Recognition, vol. 33, no. 11, pp. 1771-1782, Nov. 2000.

[6] B. Moghaddam, "Principal Manifolds and Probabilistic Subspace for Visual Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 6, pp. 780-788, June 2002.

[7] D.L. Swets and J. Weng, "Using Discriminant Eigenfeatures for Image Retrieval," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 831-836, Aug. 1996.

[8] C. Liu and H. Wechsler, "Enhanced Fisher Linear Discriminant Models for Face Recognition," Proc. Int'l Conf. Pattern Recognition, vol. 2, pp. 1368-1372, Aug. 1998.

[9] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, July 1997.

[10] D.Q. Dai and P.C. Yuen, "Regularized Discriminant Analysis and Its Application to Face Recognition," Pattern Recognition, vol. 36, no. 3, pp. 845-847, Mar. 2003.

[11] X. Wang and X. Tang, "Dual-Space Linear Discriminant Analysis for Face Recognition," Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 2, pp. 564-569, June 2004.

[12] H. Cevikalp, M. Neamtu, M. Wilkes, and A. Barkana, "Discriminative Common Vectors for Face Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no. 1, pp. 4-13, Jan. 2005.

[13] H. Yu and J. Yang, "A Direct LDA Algorithm for High-Dimensional Data with Application to Face Recognition," Pattern Recognition, vol. 34, no. 10, pp. 2067-2070, Oct. 2001.

[14] J. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos, "Face Recognition Using LDA-Based Algorithms," IEEE Trans. Neural Networks, vol. 14, no. 1, pp. 195-200, Jan. 2003

[15] K. Liu, Y.Q. Cheng, J.Y. Yang, and X. Liu, "An Efficient Algorithm for Foley-Sammon Optical Set of Discriminant Vectors by Algebraic Method," Int'l J. Pattern Recognition Artificial Intelligence.

[16] J. Yang and J.Y. Yang, "Why Can LDA Be Performed in PCA Transformed Space?" Pattern Recognition, vol. 36, pp. 563-566, 2003.