ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

Classification of crops using FCM segmentation and texture, color feature

Anandkumar Patil[1], Lalitha Y S[2]

Department of CSE, Appa Institute of Engineering and Technology, Gulbarga, Karnataka, India[1]

Department of ECE, Godutai College of Engineering for women, Gulbarga. Karnataka, India[2]

ABSTRACT— ***The objective of this study is to develop a FCM Segmentation that could distinguish crops as plant, soil and residue parts. This classification will help agronomist to decide crop pattern and cultivation practises. In this paper, collected the 10 different types of crops JPEG images from the fields. And stored as a database. After segmentation, color and texture features are applies to get the features of color and texture of the each crops. And the Euclidian distance algorithm used to identify the crops. Results show that classification accuracy is significantly improved. Hence, finally project has been demonstrated by using the plotted graphs. One for the accuracy of the images and another error rate in the classification of images.***

Keywords—Image classification, FCM, Euclidian distance, Feature extraction etc.

I. INTRODUCTION

Plants exist everywhere we live, as well as places without us. Many of them carry significant information for the development of human society. A common, well-defined system of crop classification is important in crop science and agriculture. Having standardized botanical names will also facilitate efficient communication, dissemination and retrieval of scientific information. Additionally, the grouping of crops will indicate that these crops may have similar uses, adaptation, growth habits and methods of culture. The confusion due to the absence of a universal standard should be a major concern for those who are engaged in the dissemination of the basics of crop production. Hence crop classification is the main purpose of this project. We have used texture and color feature for crop classification. We have used Sony cyber shot DSC W310 camera to take images of crops. Many kinds of machine vision technology have been employed, mainly including spectral devices and digital cameras. Bruno D.V. Marino et al given an example of total cross-sectional area of vegetation within a scene estimated from digital visible and near-infrared images obtained from simple and inexpensive camera systems. Rob E. Schouten et al used a 3CCD digital camera to get the color data of cucumber, built a generic model to predict batch keeping quality.

To segment crop JPEG image, digital image processing method Fuzzy C-Means were used. The main tools were feature extraction i.e. color, texture features operations in MATLAB. And gray co-occurrence matrix was also used to convert into gray color. And Euclidian distance used to classify the crops all these things are discussed in below chapters.

A. *Objective of the project*

To classify the crops use Fuzzy C-Means algorithm for segmenting JPEG crop image into plant, residue and soil parts. Then color and texture features are used to extract the crops color, texture features. In this project, the Euclidian distance algorithm used to classify the crop images.

II. RELATED WORK

In paper [1] the author **su-xia wang** discussed, to improve the precision of image recognition and estimation result, the grid algorithm of the crop image feature extraction was proposed. In paper [2] the author **Chong- Yaw Wee** used a scaled invariant Zernike moment based feature extractor to extract the relevant information from rice grain images for the purpose of classification. An incremental supervised learning and multidimensional maps neural network called fizzy ARTMAP (FA) has been proposed to reduce the learning time while maintaining high accuracy. In paper[3] the author **D. M. Hobson** specified, the ability to recognize defining characteristics for identification is desirable as fraudulent mislabeling of rice grain varieties is a growing problem Eight different common rice varieties were used in tests for defining features. In paper [4] the author **F. Cointault** presented firstly the use of color and texture image processing together to detect the ears, before to propose and compare different

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

texture image segmentation techniques based on feature extraction by first and higher order statistical methods. In paper [5] author **Dawid Olesiuk** given a method and results for artificial neural networks crops classification based on HyMap hyper spectral data. The method that uses an ANNs does not only depend on statistical parameters of particular class and hence makes it possible to include texture information.

In paper [6] the author **O.Kryvobok** discussed, the technique of winter crops classification using satellite data is suggested, which to separate winter crops in depending from their conditions. The technique based on textural measures computed using Gray Level Difference Vector approach (GLDV). In paper [7] the author **Sajad Kiani** discussed about in order to reduce the quantities of herbicides applied to fields, proposed to exploit the advantages of image processing to automatically detect and localize the crops and to remove all other undesired plants growing within rows and between two crops. In paper [8] the author **M. Ameer Ali** discussed by considering object similar surface variations (SSV) as well as the arbitrariness of the fuzzy c-means (FCM) algorithm for pixel location, a *fuzzy image segmentation considering object surface similarity* (FSOS) algorithm was developed, but it was unable to segment objects having SSV satisfactorily. In paper [9] the author **Juraj Horváth** describes fuzzy c-means clustering method in image segmentation. Segmentation method is based on a basic region growing method and uses membership grades' of pixels to classify pixels into appropriate segments. In paper [10] the author **M. Gomathi** presents medical image segmentation approach using Modified Fuzzy C-Means (FCM) algorithm and Fuzzy Possibilistic c-means algorithm (FPCM). This approach is a generalized version of standard Fuzzy C-Means Clustering (FCM) algorithm. In paper [11] the author **Dengsheng Zhang** proposed a method combining both color and texture features of image to improve the retrieval performance. Given a query, images in the database are firstly ranked using color features.

In paper [12] the author **Noureddine Abbadeni** considered textured images and proposes to model their textural content by a set of features having a perceptual meaning and their application to content-based image retrieval. In paper [13] the author **Xin Wang** presented inexpensive computer vision techniques allowing measuring the texture characteristics of woven fabric, such as weave repeat and yarn counts, and the surface roughness. In paper [14] the author **Jianhong Liu** extracting cropland parcels from high-resolution remote sensing images is an important issue for dynamic land-use monitoring, precision agriculture and other fields.

III. PROPOSED WORK

This section describes about proposed work of the project and system design. In the following system design figure 1, first we read every image stored in the feature database. Then one test crop image is selected, then that image converted into grayscale image using gray co-matrix function. Then segmentation is done to grayscale image, in this image segmented to soil, plant and residue part is found out using FCM segmentation. Then texture and color features are found out for each image. Then one image is taken out of all images, and then by using minimum distance algorithm crop is classified. In this project we are going to find out some texture and color feature of each crop image. So we collected a large database of crops for our project. The different types of crops are collected from the agriculture fields using digital camera. The classification project deals to find out texture and color features. The test image is to be segmented by using FCM segmentation. Finally test image is classified by using minimum distance .algorithm and plotted the graphs for accuracy an error rate.

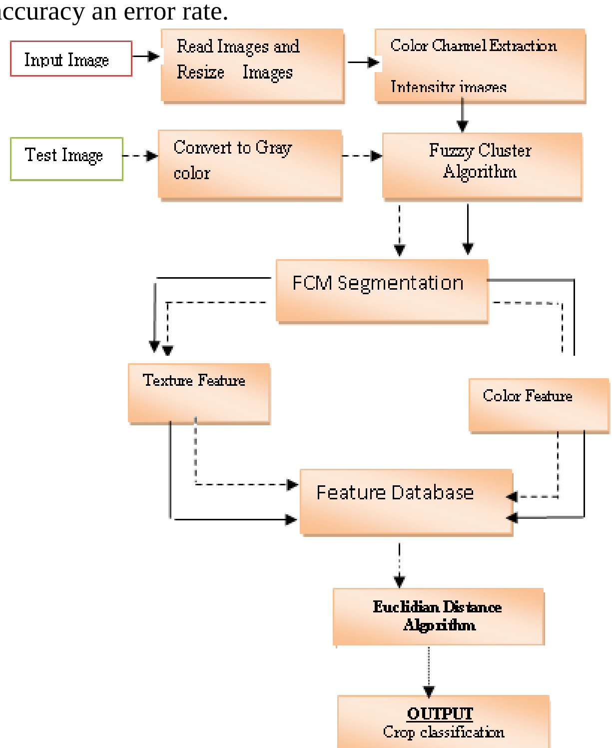

Figure 1 System Design

IV. SEGMENTATION

Modeling the statistical relations in images is an important framework for image processing and synthesis algorithms. In many applications a fixed representations such as the fourier transformation is assumed to model a large number of different images. Image processing techniques that use a more flexible model that is adapted to the structure of the underlying data can achieve better results. Adaptive techniques such as principal component analysis approximate the intrinsic structure of image data up to its second order

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

statistics. In segmentation we are using Fuzzy C Means algorithm for the extracting plant, soil and residue part to get accurate image plant. First we converted into grayscale image then applying the FCM algorithm to segment.

Image segmentation is a necessary task for image understanding and analysis. A large variety of methods have been proposed in the literature. Image segmentation can be defined as a classification problem where each pixel is assigned to a precise class. Image segmentation is a significant process for successive image analysis tasks. In general, a segmentation problem involves the division a given image into a number of homogeneous segments, such that the union of any two neighboring segments yields a heterogeneous segment.

Segmentation algorithms fall into two general classes, based on whether they searching for discontinuities or similarities. Algorithms focusing on locating discontinuities in the data are primarily edge-based, while algorithms concerned with locating adjacent pixels based on similarities are primarily region-based. Threshold techniques, a major category of algorithms, can fall into either class. In addition to these two major classes, there are also a number of general subcategories. For instance, algorithms either process color or gray-scale data, operate on either an individual pixel basis (global) or a neighborhood of pixels (local), and may use different window sizes or different color representations. For examples of surveys of segmentation algorithms. Cheng discussed the major segmentation approaches for segmenting monochrome images: histogram threshold, characteristic feature clustering, edge detection, region-based methods, fuzzy techniques, neural networks, etc.

Fuzzy c-means (FCM) is a data clustering technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade. The Fuzzy Logic Toolbox command line function `FCM` starts with an initial guess for the cluster centers, which are intended to mark the mean location of each cluster. The initial guess for these cluster centers is most likely incorrect. Additionally, `fcm` assigns every data point a membership grade for each cluster. By iteratively updating the cluster centers and the membership grades for each data point, `fcm` iteratively moves the cluster centers to the right location within a data set. This iteration is based on minimizing an objective function that represents the distance from any given data point to a cluster center weighted by that data point's membership grade. The fuzzy c-means algorithm generalizes a hard clustering algorithm called the c-means algorithm, which was introduced in the ISODATA clustering method. The (hard) c-means algorithm aims to identify compact, well-separated cluster. Basic segmentation methods based on fuzzy c-means clustering are working as follows:

1 Convert image into feature space of clustering method (usually is used RGB color space, but IHS, HLS, L*u*v* or L*a*b* color spaces are used too).
2 Run fuzzy c-means method on converted image.
3 Use some defuzzification rule or rules to classify each pixel to segment.

Simple defuzzification rule is based on maximal membership grade of pixel to cluster As shown below figure 2, the input image which is covered full area of the field. The input image is converted to grey color because gray having mono color so it is easy to function on the image. Then gray color image is segmented as soil, residue and plant part using FCM segmentation as specified in cluster 1, cluster 2, cluster 3. Here as we see the image having full area of crop so it won't segment properly as soil and residue segment parts.

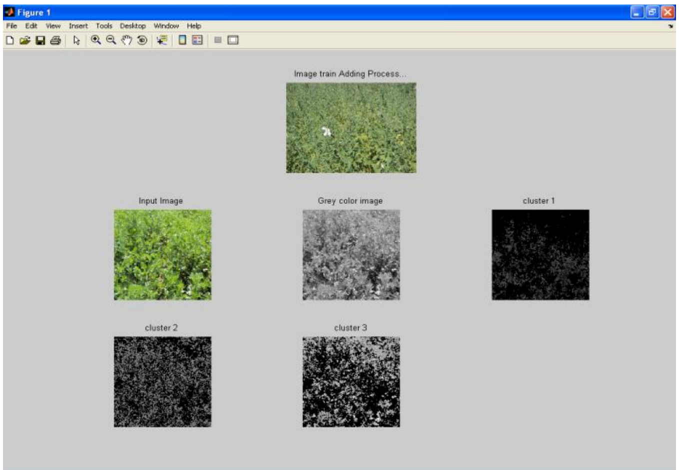

Figure 2 Segmented test image

V. FEATURE EXTRACTION

Texture feature of the image can be described by some parameters that were calculated through a couple of pixels gray co-occurrence matrix. The gray co-occurrence matrix describes the emergence probability of a couple of pixels which have *i* and j gray value respective and space d pixels in the direction of θ. The member of co-occurrence matrix is regard as P(I,j / d, θ) and also regard as when the Pi,j when the ,θ and d are constant. The gray co-occurrence matrix is a function of the distance and direction, which counts the number of the pixels couple according with the condition in the image area or the accounting window, so it reflects the integrated information of the direction, the border upon space and the change scope, which can be considered as image pixels and array structure information, that is a symmetrical matrix whose orders are decided by the amount of the grey

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

level storey in the image, therefore the gray co-occurrence matrix is used for describing texture feature quantificational.

A. Texture feature Extraction

Texture is the one of the feature of the crop. In this project we are interested to find out 9 different parameters for texture of crop by horizontally and vertically. **1. Autocorrelation 2. Contrast 3. Correlation 4. Cluster Prominence and Cluster Shade 5. Dissimilarity 6. Energy 7.Entropy 8. Homogeneity 9. Maximum probability**

Notation

C(I,j) is the (I,j)th entry in a co-occurrence matrix C

$\sum_{i}$ means $\sum_{i=1}^{i=M}$ where M is the number of rows.

$\sum_{J}$ means $\sum_{j=1}^{ij=N}$ where N is the number of rows.

$\sum_{i,i}$ means $\sum_{i}\sum_{J}$

μ_i is defined as $\mu_i = \sum i \sum C(i,j)$

μ_j is defined as $\mu_j = \sum j \sum C(i,j)$

σ_i is defined as $\sigma_i = \sum (i-\mu_i)^2 \sum C(i,j)$

σ_j is defined as $\sigma_j = \sum_j (j-\mu_j)^2 \sum_i C(i,j)$

μ_x μ_y Are the means of row and column sums respectively

σ_x σ_y Are the standard deviations of row and column sums respectively

1) Autocorrelation

An important property of many textures is the repetitive nature of the placement of texture elements in the image. The autocorrelation function of an image can be used to assess the amount of regularity as well as the fineness/coarseness of the texture present in the image. Formally, the autocorrelation function of an image is I(x,y) defined as follows:

$$\rho(x,y) = \frac{\sum_{u=0}^{N}\sum_{v=0}^{N} I(u,v)I(u+x,v+y)}{\sum_{u=0}^{N}\sum_{v=0}^{N} I^2(u,v)}$$

2) Contrast

Contrast is a measure of the local variations present in an image.If there is a large amount of variation in an image the **P[i,j]**'s will be concentrated away from the main diagonal and contrast will be high (***typically k=2, n=1***).

$$C(k,n) = \sum_i \sum_j (i-j)^k P_d[i,j]^n$$

3) Correlation

Correlation is a measure of image linearity

$$C_c = \frac{\sum_i \sum_j [ijP_d[i,j]] - \mu_i \mu_j}{\sigma_i \sigma_j}$$

$$\mu_i = \sum i P_d[i,j], \qquad \sigma_i^2 = \sum i^2 P_d[i,j] - \mu_i^2$$

Correlation will be high if an image contains a considerable amount of linear structure. The correlation feature measures the correlation between the elements of the matrix. When correlation is high the image will be more complex than when correlation is low.

4) Cluster Prominence

Cluster Shade and Cluster Prominence had to be normalized because the co-occurrence matrix is normalized. The elements $(i-\mu_i)$ and $(j-\mu_j)$ are large because μ_i and μ_i are calculated with the normalized co-occurrence matrix. In the worst case $((i-\mu_i)+(j-\mu_j))^3$ is (i+j)3. Therefore, Cluster Shade and Cluster Prominence are normalized by dividing it by (i + j) in every iteration.

$$\text{Cluster Prominence} = \sum_{i,j} \frac{((i-\mu_i)+(j-\mu_j))^4}{(i+j)^4 C(i,j)}$$

$$\text{Cluster Shade} = \sum_{i,j} \frac{((i-\mu_i)+(j-\mu_j))^3}{(i+j)^3 C(i,j)}$$

Due to this normalization Cluster Shade will vary between -1 and 1 and Cluster Prominence will vary between 0 and 1. This normalization is permitted because it has no effect on the outcome of the feature, it only rescales it.

5) . Dissimilarity

To measure the dissimilarity between unknown crop and a certain training class k, a normalized frequency $V_{L,k}$, where $\hat{g}_L$ is the frequency of run length *L* for the unknown crop? And the degree of dissimilarity R_k between the unknown crop and the class k can be obtained by summing the Over all run lengths. Finally, the degree of dissimilarity issorted from smallest to largest and given sorting grade Wk in order, {**Wk**}= {**1, 2,** □**, 10**} . As the sorting grade decreases, the similarity increases.

$$V_{L,K} = \frac{|\hat{g} - \mu_{L,K}|}{\sigma_{L,K}} \qquad R_K = \sum_{L=1}^{I} V_{L,K}$$

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

$$\{ W_K \}=\text{grade}\{\text{sort}\{ R_K \}\}$$

6) Energy

One approach to generating texture features is to use local kernels to detect various types of texture. After the convolution with the specified kernel, the ***texture energy measure (TEM)*** is computed by summing the absolute values in a local neighborhood:

$$L_e = \sum_{i=1}^{m} \sum_{j=1}^{n} |C(i, j)|$$

If ***n*** kernels are applied, the result is an ***n***-dimensional feature vector at each pixel of the image being analyzed.

7) Entropy

Entropy is a measure of information content. It measures the randomness of intensity distribution. Such a matrix corresponds to an image in which there are no preferred gray level pairs for the distance vector **d**. Entropy is highest when all entries in **P[i,j]** are of similar magnitude, and small when the entries in **P[i,j]** are unequal.

$$C_e = -\sum_i \sum_j P_d[i, j] \ln P_d[i, j]$$

8) Homogeneity

A homogeneous image will result in a ***co-occurrence matrix*** with a combination of high and low **P[i,j]**'s. Where the ***range of gray levels*** is small the **P[i,j]** will tend to be clustered around the main diagonal. A heterogeneous image will result in an even spread of **P[i,j]**'s.

$$C_h = \sum_i \sum_j \frac{P_d[i, j]}{1+|i-j|}$$

9) *Maximum probability*

This is simply the largest entry in the matrix, and corresponds to the strongest response. This could be the maximum in any of the matrices or the maximum overall.

$$C_m = \max_{i,j} P_d[i, j]$$

B. *Color Feature Extraction*

1) *Mean*

The *mean* is the average value, so it tells us something about the general brightness of the image. A bright image will have a high mean, and a dark image will have a low mean. We will use L as the total number of intensity levels available, so the gray levels range from 0 to L − 1. For example, for typical 8-bit image data, L is 256 and ranges from 0 to 255. We can define the mean as follows:

$$\bar{g} = \sum_{g=0}^{L-1} gP(g) = \sum_r \sum_c \frac{I(r,c)}{M}$$

2) *Variance*

A key problem in practical image processing in the detection of specific features in a noisy image. It tells us something about the contrast. Analysis of variance techniques can be very effective in some situations in image processing.

3) Standard deviation

The standard deviation, which is also known as the square root of the variance, tells us something about the contrast. It describes the spread in the data, so a high contrast image will have a high variance, and a low contrast image will have allowed variance. It is defined as follows:

$$\sigma_g = \sqrt{\sum_{g=0}^{L-1} (g-\bar{g})^2 P(g)}$$

VI. RESULTS

The above graph shows accuracy of the output. It is specified as total number of groups Vs percentage of accuracy. The accuracy graph is plotted depending on the performance i.e number of accurate output classified

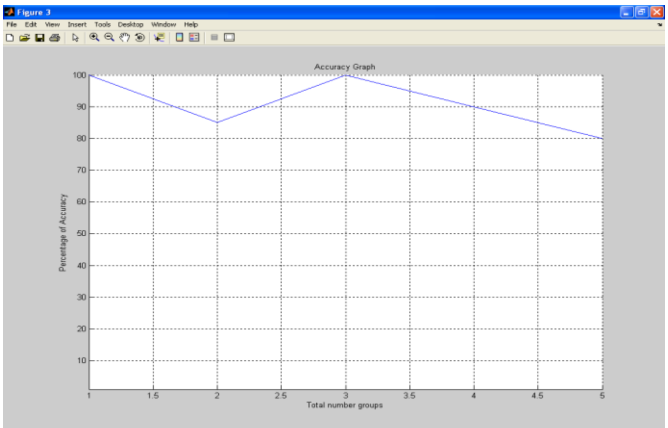

Fig 3.Accuracy graph

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

The following graph show the error rate graph plotted depending on total number of groups Vs error rate. Error rate calculated depending on the feature factor, if the input image is having same to the database image class. The classified crop class is having minimum error rate and other classes ar having hishest error rate.

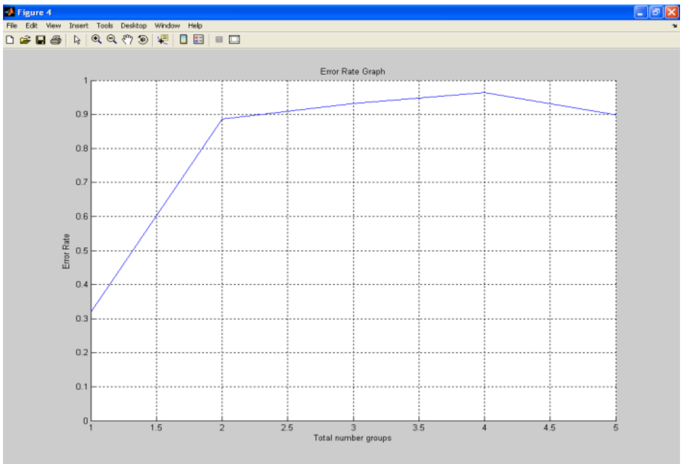

Fig 4.Error rate graph

Table 1.

Classification percentage

Sl. No.	Crop class	Classification rate
1	Banana Class	100%
2	Javar Class	85%
3	Coriander Class	100%
4	Groundnut Class	90%
5	Wheat Class	100%
6	Sugarcane Class	100%
7	Palak Class	95%
8	Onion Class	100%

The above classification percentage table shows the classification rate of each crop class. The Banana, Wheat, Onion classes gives 100% classification rate because these crops texture, color features. The Togari, Palak crops gives 95% because some images classes' texture, color features won't match with database crop classes.

Information about classification percentage

The above table shows the classification rate for different types of the image classes. We found class 1 and 3 are having good accuracy. Classification rate is low for image depending on the feature factors.

VII. CONCLUSION

In this project, we used color and texture features are extracted for each crops and then Euclidian distance algorithm for the classification of crops. The color and texture feature extraction which we used in this project can be analyzed in different planting situations.

The essential foundation was established for the crop recognition, which was reduced the time of operation and improved real time character and dependability of crop recognition greatly

REFERENCES

[1] O.Kryvobok,"Winter crops classification using satellite data" IEEE 1994.

[2] Dengsheng Zhang,"Improving Image Retrieval Performance by Using Both Color and Texture Features" ICIG'04 IEEE 2004.

[3] D. M. Hobson, R. M. Carter, Y. Yan "Characterisation and Identification of Rice Grains through Digital Image Analysis", IEEE 2007.

[4] SU-XIA WANG, ZHENG-HE SONG, ZHONG-XIANG ZHU, BANG-JIE YANG, EN-RONG MAO, RUI ZHANG **"**Study on crop image feature extraction of vehicle-based estimation system on large scale crop acreage" IEEE 2007.

[5] Jiazhi pan, Min Huang, Yong He "Crop and Weed Image Recognition by Morphological Operations and ANN model", IEEE 2007

[6] Sajad Kiani, Zohreh Azimifar, Saadat Kamgar "Wavelet-Based Crop Detection and Classification" ICEE 2010, May 11-13, 2010

[7] Dawid Olesiuk, Martin Bachmann, Martin Habermeyer, Wieke Heldens, Bogdan Zagajewski, "SNNS application for crop classification using hymap data" IEEE 2010

[8] Dr.G.Padmavathi, Mr.Muthukumar, "Image segmentation using fuzzy c means clustering method with thresholding for underwater images", Vol: 02, Issue: 02, Pages: 514-518 (2010)

Biography

Anandkumar G Patil born on June 05, 1983 in India. He received B.E. In Computer Science and Engineering and M.Tech in Computer Science and Engineering from Visweswaraya Technological University Belgaum. He is working as Assistant professor in APPA Institute of Engineering anf Technology, Gulbarga. His research interest includes image processing, fuzzy logic, ANN. He presented paper one international journal paper and one National Conference.

Lalitha Y.S was born on December 7, 1969 in India. She **received** B.E degree in Electronics and Communication Engineering and M.E. degree in Power Electronics from Gulbarga University Gulbarga, India, in 1991 and 2002 respectively. She is working as Professor in Godutai Engineering College

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

for Women Gulbarga, India. Her research interests include image Processing, Wavelet Transform coding. She presented paper in Four National Conferences and three International Conferences. She has published five international journal papers.