ABSTRACT - In these days an image retrieval system has become a challenging task. The expectation maximization (EM) algorithm has been of considerable train multilayer perception (MLP). The EM algorithm to train MLP networks may be of limited value and discuss based on the text based retrieval but the need of image based retrieval system that takes an image as the input query and retrieves images based on image content is more complicated task. For ME networks, it is reported in the literature that networks trained by the EM algorithm using iteratively Content Based Image Retrieval is an approach for retrieving semantically-relevant images from an image database based on automatically-derived image features. Image reweighted least squares (IRLS) algorithm in the inner loop of the M-step, often performed poorly in multiclass classification. The aim and Objective of the Paper is classifying the soils using multiclass classification. The expectation maximization algorithm, a Neural Network concepts for more efficient and effective results. Forthcoming problems or disasters are easily studied and predicted with help of EM algorithm. So that priory we can rescue the human kind and mother earth.

Keywords - CBIR, EM Algorithm, Clustering, multilayer perceptron (MLP).

I. INTRODUCTION

This section gives an introduction to content based image retrieval system (CBIRS) [15] and the technologies used in them. The interest in CBIR has grown because of the limitations inherent in metadata-based systems, as well as the large range of possible uses for efficient image retrieval. Textual information about images can be easily searched using existing technology, but this requires humans to manually describe each image in the database. This can be impractical for very large databases or for images that are generated automatically, e.g. those room surveillance cameras. It is also possible to miss images that use different synonyms in their descriptions. Systems based on categorizing images in semantic classes like "cat" as a subclass of "animal" can avoid the miscategorization problem, but will require more effort by a user to find images that might be "cats", but are only classified as an "animal". Many standards have been developed to categorize images, but all still face scaling and miscategorization issues. [2]

In the past, present and future of image retrieval is highlighted. There are several reasons why there is a need for additional, alternative image retrieval method like mage reweighted least square (IRLS) apart from the steadily growing rate of image production. Image retrieval is the process of browsing, searching and retrieving images from a large database of digital images. Advances in data storage and image acquisition technologies have enabled the creation of large image datasets. In order to deal with these data, it is necessary to develop appropriate information systems to efficiently manage these collections. It is simple to identify a desired image from a small collection simply by browsing, but we need more effective techniques with collections containing thousands of items. Image searching is one of the most important services that need to be supported by such systems. In general, two different approaches have been applied to allow searching on image collections: one based on image textual metadata and another based on image content information. The first retrieval approach is based on attaching textual metadata to each image and uses traditional database query techniques to retrieve them by keywords [5, 6]. In fact, different users tend to use different words to describe a same image characteristic. The lack of systematization in the annotation process decreases the performance of the keyword-based image search. Image retrieval systems have not kept pace with the collections they are searching. The shortcomings of these systems are due both to the image representations they use and to their methods of accessing those representations to find images. The problems of image retrieval are becoming widely recognized, and the search for solutions an increasingly active area for research and development.

Geographical Information Systems (GIS) and remote sensing:

Although not strictly a case of image retrieval, managers responsible for planning marketing and distribution in large corporations need to be able to search by spatial attribute (e.g. to find the 10 retail outlets closest to a given warehouse). And the military are not the only group interested in analysing satellite images. Agriculturalists and physical geographers use...
such images extensively, both in research and for more practical purposes, such as identifying areas where crops are diseased or lacking in nutrients – or alerting governments to farmers growing crops on land they have been paid to leave fallow. There by we can overcome the disadvantages of the text based retrieval systems. The main advantages of this approach is the possibility of an automatic retrieval process, contrasting to the effort needed to annotate images. In this paper it was focused on soil classification of various fields on the earth map/remote sensed image. Generally classification can be done with aid of various filter techniques but in order to classify the soils we are using an advanced platform called Neural Networks.

II RELATED WORKS

Image Processing and Analysis
Many image processing and analysis techniques have been developed to aid the interpretation of remote sensing images and to extract as much information as possible from the images. The choice of specific techniques or algorithms to use depends on the goals of each individual project. In this section, we will examine some procedures commonly used in analyzing interpreting remote sensing images.

Pre-Processing
Prior to data analysis, initial processing on the raw data is usually carried out to correct for any distortion due to the characteristics of the imaging system and imaging conditions. Depending on the user's requirement, some standard correction procedures[3] may be carried out by the ground station operators before the data is delivered to the end-user. These procedures include radiometric correction to correct for uneven sensor response over the whole image and geometric correction to correct for geometric distortion due to Earth's rotation and other imaging conditions (such as oblique viewing). The image may also be transformed to conform to a specific map projection system. Furthermore, if accurate geographical location of an area on the image needs to be known, ground control points (GCP's) are used to register the image to a precise map (geo-referencing).

Image Enhancement
In order to aid visual interpretation, visual appearance of the objects in the image can be improved by image enhancement techniques such as grey level stretching to improve the contrast and spatial filtering for enhancing the edges. An example of an enhancement procedure is shown here.

Multispectral SPOT image of the same area shown in a previous section, but acquired at a later date. Radiometric and geometric corrections have been done. The image has also been transformed to conform to a certain map projection (UTM projection). This image is displayed without any further enhancement. In the above unenhanced image, a bluish tint can be seen all-over the image, producing a hazy appearance. This hazy appearance is due to scattering of sunlight by atmosphere into the field of view of the sensor. This effect also degrades the contrast between different land covers. It is useful to examine the image Histograms before performing any image enhancement. The x-axis of the histogram is the range of the available digital numbers, i.e. 0 to 255. The y-axis is the number of pixels in the image having a given digital number. The histograms of the three bands of this image are shown in the following figures.

Note that the minimum digital number for each band is not zero. Each histogram is shifted to the right by a certain amount. This shift is due to the atmospheric scattering component adding to the actual radiation reflected from the ground. The shift is particular large for the XS1 band compared to the other two bands due to the higher contribution from Rayleigh scattering for the shorter wavelength. The maximum digital number of each band is also not 255. The sensor's gain factor
has been adjusted to anticipate any possibility of encountering a very bright object.

Hence, most of the pixels in the image have digital numbers well below the maximum value of 255. The image can be enhanced by a simple linear grey-level stretching. In this method, a level threshold value is chosen so that all pixel values below this threshold are mapped to zero. An upper threshold value is also chosen so that all pixel values above this threshold are mapped to 255. All other pixel values are linearly interpolated to lie between 0 and 255. The lower and upper thresholds are usually chosen to be values close to the minimum and maximum pixel values of the image. The Grey-Level Transformation Table is shown in the following graph.

In supervised classification, the spectral features of some areas of known land cover types are extracted from the image. These areas are known as the "training areas". Every pixel in the whole image is then classified as belonging to one of the classes depending on how close its spectral features are to the spectral features of the training areas.

In unsupervised classification, the computer program automatically groups the pixels in the image into separate clusters, depending on their spectral features. Each cluster will then be assigned a land cover type by the analyst. Each class of land cover is referred to as a "theme" and the product of classification is known as a "thematic map".

The following image shows an example of a thematic map. This map was derived from the multispectral SPOT image of the test area shown in a previous section using an unsupervised classification algorithm.

Classification algorithm.
A plausible assignment of land cover types to the thematic classes is shown in the following table. The accuracy of the thematic map derived from remote sensing images should be verified by field observation.
The spectral features of these Land cover classes can be exhibited in two graphs shown below.

The first graph is a plot of the mean pixel values of the XS3 (near infrared) band versus the XS2 (red) band for each class. The second graph is a plot of the mean pixel values of the XS2 (red) versus XS1 bands. The standard deviations of the pixel values for each class are also shown.

In the scatter plot of the class means in the XS3 and XS2 bands, the data points for the non-vegetated land cover classes generally lie on a straight line passing through the origin. This line is called the "soil line". The vegetated land cover classes lie above the soil line due to the higher reflectance in the near infrared region (XS3 band) relative to the visible region.

In the XS2 (visible red) versus XS1 (visible green) scatter plot, all the data points generally lie on a straight line. This plot shows that the two visible bands are very highly correlated. The vegetated areas and clear water are generally dark while the other non-vegetated land cover classes have varying brightness in the visible bands.

**Spatial Feature Extraction**

In high spatial resolution imagery, details such as buildings and roads can be seen. The amount of details depend on the image resolution. In very high resolution imagery, even road markings, vehicles, individual tree crowns, and aggregates of people can be seen clearly. Pixel-based methods of image analysis will not work successfully in such imagery.

In order to fully exploit the spatial information contained in the imagery, image processing and analysis algorithms utilizing the textural, contextual and geometrical properties are required. Such algorithms make use of the relationship between neighboring pixels for information extraction. Incorporation of a-priori information is sometimes required. A multi-resolution approach (i.e. analysis at different spatial scales and combining the resolute) is also a useful strategy when dealing with very high resolution imagery. In this case, pixel-based method can be used in the lower resolution mode and merged with the contextual and textural method at higher resolutions.

**Geographical Information System (GIS)**

Different forms of imagery such as optical and radar images provide complementary information about the land cover. More detailed information can be derived by combining several different types of images. For example, radar image can form one of the layers in combination with the visible and near infrared layers when performing classification.

The thematic information derived from the remote sensing images are often combined with other auxiliary data to form the basis for a Geographic Information System (GIS). AGIS is a database of different layers, where each layer contains information about a specific aspect of the same area which issued for analysis by the resource scientists.

The image shown in Figure 1 has been divided into N = 16 rows and M = 16 columns. The value assigned to every pixel is the average brightness in the pixel rounded to the nearest integer value. The process of representing the amplitude of the 2D signal at a given coordinate as an integer value with L different gray levels is usually referred to as amplitude quantization or simply quantization.

- Common Values
- Characteristics of Image Operations
- Video Parameters
- Common Values
There are standard values for the various parameters encountered in digital image processing. These values can be caused by video standards, by algorithmic requirements, or by the desire to keep digital circuitry simple. Table 1.1 gives some commonly encountered values.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Typical values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rows</td>
<td>N</td>
<td>256, 512, 525, 625, 1024, 1035</td>
</tr>
<tr>
<td>Columns</td>
<td>M</td>
<td>256, 512, 768, 1024, 1320</td>
</tr>
<tr>
<td>Gray Levels</td>
<td>L</td>
<td>2, 64, 256, 1024, 4096, 16384</td>
</tr>
</tbody>
</table>

Table 1.1: Common values of digital image parameters [15]

Quite frequently we see cases of M=N=2K where {K = 8, 9, 10}. This can be motivated by digital circuitry or by the use of certain algorithms such as the (fast) Fourier transform. The number of distinct gray levels is usually a power of 2, that is, L=2B where B is the number of bits in the binary representation of the brightness levels. When B=1 we speak of a gray-level image; when B=1 we speak of a binary image. In a binary image there are just two gray levels which can be referred to, for example, as "black" and "white" or "0" and "1".

III. PROPOSED WORKS

This systems based on features like colour, shape, texture, spatial layout, object motion, etc., are cited in [11], [12]. Colour is one of the most widely used features for image similarity retrieval. Colour retrieval yields the best results, in that the computer results of colour similarity are similar to those derived by a human visual system that is capable of differentiating between infinitely large numbers of colours. One of the main aspects of colour feature extraction is the choice of a colour space. A colour space is a multidimensional space in which the different dimensions represent the different components of colour [13]. Most colour spaces are three dimensional. Example of a colour space is RGB, which assigns to each pixel a three element vector giving the colour intensities of the three primary colours, red, green and blue. The space spanned by the R, G, and B values completely describes visible colours, which are represented as vectors in the 3D RGB colour space. As a result, the RGB colour space provides a useful starting point for representing colour features of images.

The EM Algorithm

The EM algorithm [11] was explained and given its name in a classic 1977 paper by Arthur Dempster, Nan Laird, and Donald Rubin. They pointed out that the method had been "proposed many times in special circumstances" by earlier authors. EM is typically used to compute maximum likelihood estimates given incomplete samples. The EM algorithm estimates the parameters of a model iteratively. Starting from some initial guess, each iteration consists of an E step (Expectation step) and M step (Maximization step).

In the sequel, we shall assume that the neural network is being used in a multiclass classification context. In the classification g context, there are populations or groups $G_1, G_2, ..., G_n$ and the problem is to infer the unknown membership of an unclassified entity with feature vector of $g$-dimensions. This membership can be defined by a $g$-dimensional output vector of zero-one indicator variables, where the $i$th element of the output vector is one or zero, according as the entity does or does not belong to the $i$th group $G_i$.

We let $(x_1^T, y_1^T, ..., x_n^T)^T = (x_1^T, y_1^T, ..., x_n^T)^T$ for $i = 1, 2, ..., n$. In the training process, the unknown parameters in the neural network, denoted by a vector $Ψ$, are inferred from the observed training data given by (1). We let and $x=(x_1^T, ..., x_n^T)^T$ and $y=(y_1^T, ..., y_n^T)^T$. In order to estimate $Ψ$ by the statistical technique of maximum likelihood, we have to impose a statistical distribution for the observed data (1), which will allow us to form a log likelihood function $log L(Ψ, y, x)$, for $Ψ$. In general, we proceed conditionally on the values for the input variable $x$; that is, we shall consider the specification of the conditional distribution of the random variable $Y$ corresponding to the observed output $y$ given the input $x$ as $log L(Ψ, y, x)α log pr(Y | x, Ψ)$.

The EM algorithm is a popular tool for the iterative computation of ML estimates, useful in a variety of circumstances.
incomplete-data problems, where algorithms such as gradient ascent methods may turn out to be more complicated. Further details on the EM algorithm in a general context can be found in the monograph of McLachlan and Krishnan [14]. Within the EM framework, the unknown vector \( \Psi \) is estimated by consideration of the complete-data log likelihood (that is, the log likelihood function for \( \Psi \) based on both the observed and the missing data \( z \)). The complete-data log likelihood formed on the basis of both the observed and the missing data is given by

\[
\log L_c(\Psi; y, z, x) = \log \Pr(Y | x; \Psi) + \log \Pr(Z | x; \Psi),
\]

That is, we need to specify the distribution of the random variable \( Z \), conditional on \( x \), and the conditional distribution of \( Y \) given \( x \) and \( z \). On the one hand, \( \log L_c(\Psi; y, z, x) \) is computed by the approximation

\[
Q(\Psi; \Psi^{(k)}) = E_{\Psi^{(k)}} \{ \log L_c(\Psi; y, z, x) | y, x \},
\]

where \( E_{\Psi^{(k)}} \) denotes the expectation operator using the current value \( \Psi^{(k)} \) for \( \Psi \). On the other hand, \( \Psi^{(k+1)} \), is updated by taking \( \Psi^{(k+1)} = \operatorname{argmax} Q(\Psi; \Psi^{(k)}) \) over all admissible values of \( \Psi \). In some instances, a modified form of the EM algorithm is being used unwittingly by the author(s) in that on the E-step, the conditional expectation of the complete-data log likelihood, the Q-function, is effected simply by replacing the random vector by its conditional expectation. For example, in (3), \[12\], \[16\] is computed by the approximation

\[
Q(\Psi; \Psi^{(k)}) \approx \log L_c(\Psi; y, z, x)
\]

Where

\[ z = E_{\Psi^{(k)}} \{ \Psi; \Psi^{(k)} \}. \]

However, the approximation (4) will be valid only in special cases. It is valid if the complete-data log likelihood were linear in as in the ME neural networks.

### TRAINING MULTILAYER PERCEPTRON NETWORKS

An MLP neural network constructs a decision surface in the data space for discriminating instances with similar features by forming a boundary between them. For a MLP neural network with one hidden layer \( m \) of units (Fig. 1), we can specify a stochastic model of MLP neural network as follows. Let \( z_{hj} (h=1,2,...m; j=1,2,...n) \) be the realization of the zero-one random variable \( z_{hj} \) for which its conditional distribution given \( x_j \) is specified by

\[
\Pr(Z_{hj}=1 | x_j) = \frac{\exp(w_h^T x_j)}{1+\exp(w_h^T x_j)}
\]

Where \( w_h=(w_{h0}, w_{h1}, ..., w_{hp})^T \) is the synaptic weight vector of the \( h_{th} \) hidden unit. The bias term is included in \( w_h \) by adding a constant input \( x_{0j}=1 \) for all \( j=1,2,...,n \) so that the input is now \( x_j=(x_{0j}, x_{1j}, ..., x_{pj})^T \) that is

\[
w_h^T x_j = \sum_{i=1}^{p} w_{hi} x_{ij} + w_{ho} = \sum_{i=0}^{p} w_{hi} x_{ij}
\]

The probability \( \Pr(Z_{hj}=0|x_j) \) is equal to \( 1 - \Pr(Z_{hj}=1|x_j) \)

That is, \( z_{hj} \) has a Bernoulli distribution. The output of \( g \)-dimensional zero-one indicator variables is \( Y_{j} \) distributed according to a multinomial distribution consisting of one draw on \( g \) cells with probabilities

\[
\Pr(Y_{ij}=l | x_j, z_{j}) = \frac{\exp(v_i^T z_j)}{\sum_{r=1}^{g} \exp(v_r^T z_j)}
\]

for \( l=1,...,g \) where \( v_i=(v_i0, ..., v_im)^T \) is the synaptic weight vector of the \( l_{th} \) output unit. The bias term \( v_{i0} \) is included in \( v_i \) by adding a constant hidden unit \( z_{0j}=1 \) for all \( j=1,2,...,n \) so that the hidden layer is now \( z_j=(z_{0j}, z_{1j}, ..., z_{mj})^T \) that is

\[
v_i^T z_j = \sum_{h=1}^{m} v_{ih} z_{hj} \quad v_{i0} = \sum_{h=0}^{m} v_{ih} z_{hj}
\]

In the EM framework, the missing data are then given by \( z=(z_1^T, ..., z_n^T)^T \).

**Table 2.1: Sample training data for soil classification**

<table>
<thead>
<tr>
<th>Class No. (Colour in Map)</th>
<th>Landcover Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (black)</td>
<td>Clear water</td>
</tr>
<tr>
<td>2 (green)</td>
<td>Dense Forest with closed canopy</td>
</tr>
<tr>
<td>3 (yellow)</td>
<td>Shrubs, Less dense forest</td>
</tr>
<tr>
<td>4 (orange)</td>
<td>Grass</td>
</tr>
<tr>
<td>5 (cyan)</td>
<td>Bare soil, built-up areas</td>
</tr>
<tr>
<td>6 (blue)</td>
<td>Turbid water, bare soil, built-up areas</td>
</tr>
<tr>
<td>7 (red)</td>
<td>bare soil, built-up areas</td>
</tr>
<tr>
<td>8 (white)</td>
<td>bare soil, built-up areas</td>
</tr>
</tbody>
</table>
### Table 2.2: Inference results for the soil classification (untrained data)

<table>
<thead>
<tr>
<th>Color of the soil</th>
<th>(Gravel %)</th>
<th>(Sand %)</th>
<th>(Fine grain ed particles %)</th>
<th>(Liquid limit %)</th>
<th>(Plastic limit %)</th>
<th>I.S Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18</td>
<td>82</td>
<td></td>
<td>59</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>0</td>
<td>0.329</td>
<td>0.869</td>
<td>0.71</td>
<td>0.7</td>
<td>0.203(0.2)</td>
</tr>
<tr>
<td>0.1</td>
<td>0</td>
<td>0.341</td>
<td>0.857</td>
<td>0.69</td>
<td>0.7</td>
<td>0.193(0.2)</td>
</tr>
<tr>
<td>0.2</td>
<td>0.11</td>
<td>0.682</td>
<td>0.5</td>
<td>0.50</td>
<td>0.5</td>
<td>0.1(0.1)</td>
</tr>
<tr>
<td>0.2</td>
<td>0</td>
<td>0.548</td>
<td>0.654</td>
<td>0.57</td>
<td>0.6</td>
<td>0.289(0.3)</td>
</tr>
<tr>
<td>0.3</td>
<td>0</td>
<td>0.756</td>
<td>0.452</td>
<td>0.49</td>
<td>0.5</td>
<td>0.129(0.1)</td>
</tr>
<tr>
<td>0.3</td>
<td>0</td>
<td>0.585</td>
<td>0.619</td>
<td>0.61</td>
<td>0.8</td>
<td>0.594(0.6)</td>
</tr>
</tbody>
</table>

### IV. CONCLUSION

This paper contains with Three different domains of sciences i.e Basics of Digital Image Processing, Soil Fundamentals & Neural Networks. We need to study Supervised and Unsupervised learning techniques and we are using the concepts of EM algorithm. The algorithm which are used in the project are predefined functions, which cannot be altered according to our task. The most important algorithms that are used in this paper is EM, Simplified FUZZY ARTMAP for soil classification as well as image recognition. All three algorithms are completely mathematical based tools i.e the functions that are used in these algorithms are predefined one. Presenting these algorithms in Java is really a challenging task, working with dynamic image and collecting the relevant data such as position of the pixel, RGB values and converting it into the intensity values and then giving these values as the inputs to the these algorithms through Java and verifying the output values with the trained data ends the project.

Even though a large number of commercial and open source applications exist to process remote sensing data. According to an NOAA Sponsored Research by Global Marketing Insights, Inc. the most used applications among Asian academic groups involved in remote sensing are as follows: ESRI 30%; ERDAS IMAGINE 25%; ITT Visual Information Solutions ENVI 17%; MapInfo 17%; ERMapper 11%. Among Western Academic respondents as follows: ESRI 39%, ERDAS IMAGINE 27%, MapInfo 9%, AutoDesk 7%, ITT Visual Information Solutions ENVI 17%. It is an attempt to study two important learning techniques which plays a vital role in the image processing through these algorithms makes the project more interesting and challenging using Neural Network and Java Flat form.

**Experimental Results:**

- Presenting the updated weights matrix to the user.
- Updated weights matrix to the user.
- Click to continue...
REFERENCES


K.Eswara Rao received the M.Tech in CSE & NN from JNTU, Kakinada and B.Tech in CSE Degree from the department of Computer Science and Engineering at the Andhra University, Vishakhapatnam respectively. He is currently working as Assistant. Professor in Department of Computer Science &Engineering, in Aditya Institute of Technology and Management, Tekkali, Andhra Pradesh, India. He has 7 years of experience in teaching Computer Science and Engineering subjects. His Research interests include Neural Networks, Data Mining.

A.Nagabhushana Rao received The M.Tech in CSE from JNTU Kakinada from the department of Computer Science & Engineering. He is currently working as Assistant. Professor in Department of Computer Science & Engineering, in Aditya Institute of Technology and Management, Tekkali, Andhra Pradesh, India. He has 4 years of experience in teaching Computer Science and Engineering subjects. His Research interests include Data Mining, Networking and Security.