ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

Mixed Pixels: A Challenge in Remote Sensing Data Classification for Improving Performance

A. L. Choodarathnakara, Dr. T. Ashok Kumar, Dr. Shivaprakash Koliwad, Dr. C. G. Patil

***Abstract*— Remote Sensing (RS) refers to the science of identification of earth surface features and estimation of their geo-biophysical properties using electromagnetic radiation as a medium of interaction. Spectral, spatial, temporal and polarization signatures are major characteristics of the sensor/target, which facilitate target discrimination. Earth surface data as seen by the sensors in different wavelengths (reflected, scattered and/or emitted), is radio-metrically and geo-metrically corrected before extraction of spectral information. RS data, with its ability for a synoptic view repetitive coverage with calibrated sensors to detect changes, observations at different resolutions, provides a better alternative for natural resources management as compared to traditional methods. The traditional classification methods have been developed using the image statistics, but their applicability to the processing of data is limited due to the presence of mixed pixels which affect the use of remotely sensed data in Land-Use and Land-Cover (LU/LC) classification and change detection. Various techniques are being used to attempt to 'unmix' the information and identify mixing proportions from the pixels and have lead to sub-pixel classification.**

The purpose of this paper is to understand mixed pixel problems, its significance and techniques developed to 'unmix' the pixels, and also to study the various soft classification methods. It is understood that mixed pixel classification could be solved efficiently by Soft Classification methods, and Contextual Classification adopting relaxation labeling would give better results both in Neural Network based and Fuzzy Logic based Classification.

Some of the major operational application themes, in which India has extensively used remote sensing data, are agriculture, forestry, water resources, land use, urban sprawl, geology, environment, coastal zone, marine resources, snow and glacier, disaster monitoring and mitigation, infrastructure development, etc.

***Index Terms*— Remote Sensing, Image Classification, Pure Pixels, Mixed Pixels, Supervised, Unsupervised, Hard Classification, Soft Classification**

Manuscript received Nov 20, 2012.

***A. L. Choodarathnakara**, Department of Electronics & Communication Engineering, Government Engineering College Kushalnagar-571234, South Kodagu INDIA, 9945466988.*

***Dr. T. Ashok Kumar**, Department of Electronics & Communication Engineering, Vivekananda College of Engineering & Technology, Puttur (Dakshina Kannada), INDIA.*

***Dr. Shivaprakash Koliwad**, Department of Electronics & Communication Engineering, Malnad College of Engineering Hassan-573201, INDIA.*

***Dr. C. G. Patil**, Master Control Facility, Hassan-573201, Karnataka State, INDIA.*

I. Introduction

Digital Processing of remotely sensed data has gained momentum in the last ten to fifteen years with the availability of digital data. Remotely sensed data of the Earth may be analyzed to extract useful thematic information. A thematic map shows the spatial distribution of identifiable Earth surface features and provides an informational description over a given area, rather than a data description (Robert A, 1997). Image classification is one of the most often used methods of information extraction and used to produce thematic maps from imagery. The themes can range, for example, form categories such as soil, vegetation and surface water in a general description of a rural area, to buildings, parking lots, roads etc in urban area. A number of factors can cause confusion among spectral signatures, including topography, shadowing, atmospheric variability, sensor calibration changes, and class mixing with in the Instantaneous Field of View (IFOV) (Robert A, 1997).

A pixel is considered as the minimum spatial unit in a digital imagery and it is assumed that a pixel covers a homogeneous land cover data. But very often, the IFOV of a sensor contains multiple cover types and generates confusion when performing any type of classification. Every satellite imaging system has some characteristics which limit the use of the data being produced. The occurrence of mixed pixels depends on the spatial resolution of system and the spatial distribution of the cover types under analysis. For example, Landsat TM satellite images have a spatial resolution of 30 meter. It is usually very difficult to interpret successfully the objects from such coarse resolution images, especially in urban area. Images at such coarse resolutions result in mixed pixels (Thakkar, 2004). Graeme G. W. (2005) has opinioned that there has been no demonstrable improvement in classification performance over the last 15 years period. The mean value of the Kappa coefficient across all experiments was found to be 0. 6557 with a standard deviation value of 0.1980. Hence, mixed pixels provide a substantial challenge in classifying imagery.

A. *Types of Pixels*

1) *Pure Pixels:* Pure Pixels are usually the pixels in the image which represent a single class. These pixels represent areas covered by a single component type. One of the main steps in the mixed pixel analysis is to find such pure pixels. These pure class pixels are the key input to most approaches for un-mixing the mixed pixels.

2) *Mixed Pixels:* Mixed pixels problem is created in digital imagery by those pixels not completely occupied by a single, homogeneous category. When a pixel area is composed of two or more areas that differ greatly with

respect to brightness, then the average is composed of several very different values, and the single digital value that represents the pixel may not accurately represent any of the categories present. The mixing proportion is also called class fractions. Mixed pixels occur often at the edges of large parcels, or along long linear features, such as rivers or highways, where contrasting brightness are immediately adjacent to one another.

The mixed pixel problem cannot be solved simply by increasing the spatial resolution. In general, the proportion of mixed pixels decreases as the spatial resolution becomes higher, and the smaller pixel size allows more pure pixels to fit within the object boundaries. In some cases, however, the proportion of mixed pixels can actually increase because the finer detail resolves the features not recorded before and introduces new spectral classes. Further, at higher resolutions, the within-class variation increases as local differences in humidity, elevation, illumination, etc., become more apparent. Another reason is that increase in spatial resolution usually is achieved at the expense of the spectral of radiometric resolution, because the reduction in received energy due to a smaller IFOV must be compensated for, by broadening the spectral band at which the reflectance measured. A further disadvantage of higher spatial resolution is that the number of pixels can become very large, which adds to the cost of processing. Together with the fact that there is a growing interest in data of global coverage recorded on more and smaller spectral bands, new scanners may be designed to have coarser rather than finer spatial resolution (Gebbinek, 1998).

II. Image Classification Process

Image classification is a complex process that may be affected by many factors. Remote-sensing research focusing on image classification has long attracted the attention of the remote-sensing community because classification results are the basis for many environmental and socioeconomic applications. Scientists and practitioners have made great efforts in developing advanced classification approaches and techniques for improving classification accuracy [Lu and Weng, 2007].

Remote-sensing image classification requires consideration of many factors. The major steps of image classification may include determination of a suitable classification system, selection of training samples, image pre-processing, and feature extraction, selection of suitable classification approaches, post-classification processing, and accuracy assessment. This section focuses on the description of the major steps that may be involved in image classification.

A. Selection of Remotely Sensed data

Remotely sensed data, including both airborne and space borne sensor data vary in spatial, radiometric, spectral, and temporal resolutions. Understanding the strengths and weaknesses of different types of sensor data is essential for the selection of suitable remotely sensed data for image classification. The selection of suitable sensor data is the first important step for a successful classification for a specific purpose. It requires considering factors such as user's need, scale and characteristics of a study area, availability of various image data and their characteristics, cost and time constraints, and the analyst's experience in using the selected image. Scale study area, image resolution and the user's need are the most important factors affecting the selection of suitable remotely sensed data and determine the nature of classification.

Another important factor influencing the selection of sensor data is the atmospheric condition. The frequent cloudy conditions in the moist tropical regions are often an obstacle for capturing high-quality optical sensor data. Therefore, different kinds of radar data serve as an important supplementary data source. Since multiple sources of sensor data are now readily available, image analysts have more choices to select suitable remotely sensed data for a specific study [Lu and Weng, 2007].

B. Selection of a Classification System and Training Samples

Suitable classification system and a sufficient number of training samples are prerequisites for a successful classification. A sufficient number of training samples and their representativeness are critical for image classifications. Training samples are usually collected from fieldwork, or from fine spatial resolution aerial photographs and satellite images. Different collection strategies, such as single pixel, seed, and polygon, may be used, but they would influence classification results, especially for classifications with fine spatial resolution image data .When the landscape of a study area is complex and heterogeneous, selecting sufficient training samples becomes difficult. This problem would be complicated if medium or coarse spatial resolution data are used for classification, because a large volume of mixed pixels may occur. Therefore, selection of training samples must consider the spatial resolution of the remote-sensing data being used, availability of ground reference data, and the complexity of landscapes in the study area.

C. Data Pre-processing

Image pre-processing may include the detection and restoration of bad lines, geometric rectification or image registration, radiometric calibration and atmospheric correction, and topographic correction. If different ancillary data are used, data conversion among different sources or formats and quality evaluation of these data are also necessary before they can be incorporated into a classification procedure. Accurate geometric rectification or image registration of remotely sensed data is a prerequisite for a combination of different source data in a classification process.

When multi-temporal or multi-sensor data are used, atmospheric calibration is mandatory. This is especially true when multi-sensor data, such as Landsat TM and SPOT or Landsat TM and radar data, are integrated for an image classification. A variety of methods, ranging from simple relative calibration and dark-object subtraction to calibration approaches based on complex models (e.g. 6S), have been developed for radiometric and atmospheric normalization and correction. Topographic correction is another important aspect if the study area is located in rugged or mountainous regions.

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

D. Selection of a Suitable Classification Method

Many factors, such as spatial resolution of the remotely sensed data, different sources of data, a classification system, and availability of classification software must be taken into account when selecting a classification method for use. Different classification methods have their own merits and demerits also.

E. Post-classification Processing

Traditional per-pixel classifiers may lead to 'salt and pepper' effects in classification maps. A majority filter is often applied to reduce the noises. Most image classification is based on remotely sensed spectral responses. Due to the complexity of biophysical environments, spectral confusion is common among land-cover classes. Thus, ancillary data are often used to modify the classification image based on established expert rules. For example, forest distribution in mountainous areas is related to elevation, slope, and aspects. Data describing terrain characteristics can therefore be used to modify classification results based on the knowledge of specific vegetation classes and topographic factors.

F. Evaluation of Classification Performance

Evaluation of classification results is an important process in the classification procedure. Different approaches may be employed, ranging from a qualitative evaluation based on expert knowledge to a quantitative accuracy assessment based on sampling strategies. To evaluate the performance of a classification method, researchers have proposed six criteria: accuracy, reproducibility, robustness, ability to fully use the information content of the data, uniform applicability, and objectiveness. In reality, no classification algorithm can satisfy all these requirements nor be applicable to all studies, due to different environmental settings and datasets used.

Few of them have suggested the use of multiple criteria to evaluate the suitability of algorithms. These criteria include classification accuracy, computational resources, stability of the algorithm, and robustness to noise in the Improving classification performance training data. Classification accuracy assessment is, however, the most common approach for an evaluation of classification performance.

III. Digital Image Classifiers

Image classification may be performed using spatial, spectral or temporal information. Spatial pattern recognition determines classes based on spatial relationships (e.g. texture, proximity, size, shape, directionality, repetition and context) between pixels and temporal pattern recognition uses temporal changes to enhance classification (spatial or spectral). But such techniques are typically complex and computationally intensive.

Image classification is a complex process that may be affected by many factors. The effective use of multiple features of remotely sensed data and the selection of a suitable classification method are especially significant for improving classification accuracy. Non-parametric classifiers such as fuzzy logic, neural network, decision tree classifier, and knowledge-based classification have increasingly become important approaches for multisource data classification. Integration of remote sensing, geographical information systems (GIS), and expert system emerges as a new research frontier. More research, however, is needed to identify and reduce uncertainties in the image-processing chain to improve classification accuracy [19].

In general, image classification approaches can be grouped as supervised and unsupervised, or parametric and nonparametric, or hard and soft (fuzzy) classification, or per-pixel, sub-pixel, and per-field. Based on whether training samples used or not used image classification can be of two types:

A. Supervised Classification

In this type of classification, the image analyst "supervises" the pixel categorization process by specifying, to the computer algorithm, numerical descriptors of the various land cover types present in a scene. To do this representative sample sites of known cover type, called training areas, are used to compile a numerical "interpretation key" that describes the spectral attributes for feature type of interest. Each pixel in the data set is then compared numerically to each category in the interpretation key and labelled. There are a number of numerical strategies that can be employed to make this comparison between unknown pixels and training set pixels.

Supervised classification procedures require considerable interaction with the analyst, who must guide the classification by identifying areas on the image that are known prior to the classification, which belongs to specific LU/LC categories. These areas are referred to as training sites. The training sites or samples of known identity are then used to classify pixels of unknown identity. The locations of the training site pixels should be based on ground truth, whenever possible. The computer uses the spectral characteristics of the training pixels to identify other pixels in the image with similar characteristics. Choosing these training pixels is the key to the success of any supervised classification method, which includes Parallelpiped, Maximum Likelihood, Minimum Distance and Mahalanobis Distance [19].

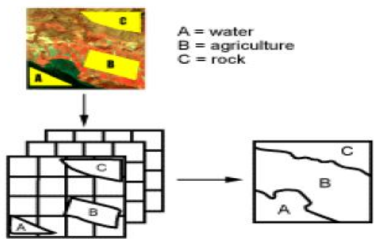

Fig. 1 Supervised Classification Scheme

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

1) *Maximum Likelihood Classifier:* The maximum likelihood classifier quantitatively evaluates both the variance and covariance of the category spectral response patterns when classifying an unknown pixel. An assumption that the distribution of the cloud of points forming the category training data is Gaussian is made. With this assumption the distribution of a category response pattern can be completely described by the mean vector and the covariance matrix. Given these parameters statistical probability of a given pixel value being a member of a particular land cover class can be computed [19].

In essence, the maximum likelihood classifier delineates ellipsoidal "equi-probability contours" in the scatter diagram and these decision regions are shown in Fig. 2. The shape of the equi-probability contours expresses the sensitivity of the likelihood classifier to covariance. For example because of the sensitivity it can be seen that pixel 1 would be appropriately assigned to the "corn" category.

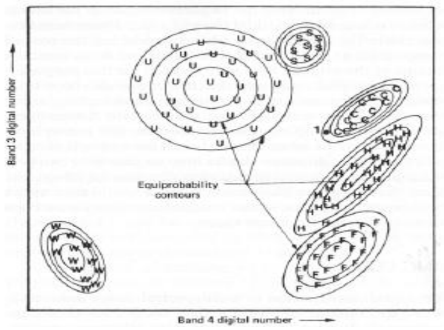

Fig. 2 Equi-probability contours defined by a MLC

2) *Minimum-Distance-To-Means-Classifier:* One of the simpler classification strategies is shown in Fig. 3 called Min-Distance-To-Mean classifier. First, the mean, or average, spectral value in each band for each category is determined. These values comprise the mean vector for each category. The category means is indicated by +'s symbol as shown in Fig. 3. By considering the two channel pixel values as positional coordinates, a pixel of unknown identity may be classified by computing the distance between the value of the unknown pixel and each of the category means. In Fig. 3, an unknown pixel has been plotted at a point 1. The distance between this pixel value and each category is illustrated by dashed lines. In this case the class is "corn". The pixel value plotted at point 2 is identified as "sand".

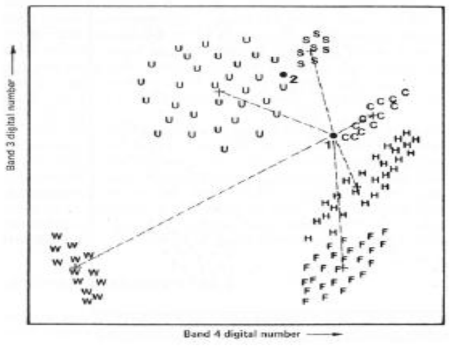

Fig. 3 Minimum distance to means classification strategy

B. Unsupervised Classification

Unsupervised classification involves the process of automatically segmenting an image into spectral classes based on the natural groupings found within the data set. In unsupervised classification, any individual pixel is compared to each discrete cluster to see which one it is closest to, in terms of spectral value.

Mixed pixels cannot be appropriately accommodated by conventional image classification techniques and consequently the spatial representation of land cover classes and estimates of their, extend derived from such classification may be erroneous. Whereas soft classification techniques allow for the partial and multiple class membership within each mixed pixel and, therefore, may be used to refine the standard mapping process and to increase the accuracy of land cover mapping from remote sensing.

Unsupervised classification (commonly referred to as clustering) is an effective method of partitioning remote sensor image data in multispectral feature space and extracting land-cover information. Compared to supervised classification, unsupervised classification normally requires only a minimal amount of initial input from the analyst. This is because clustering does not normally require training data. The unsupervised procedures are applied in two separate steps. In the unsupervised approach the image data are first classified by aggregating them into the natural spectral groupings or clusters as shown in Fig. 4. Then the image analyst determines the land cover identity of these spectral groups by comparing the classified image data to ground reference data [20]. Examples of unsupervised classification techniques are K-means clustering, ISODATA clustering, etc.

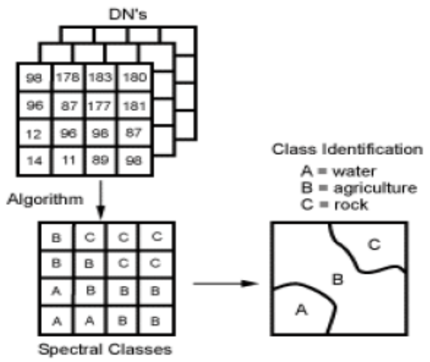

Fig. 4 Unsupervised Classification Scheme

1) *K-means Clustering Algorithm:* An initial mean vector (seed) is arbitrarily specified for each of K-clusters; each pixel is then assigned to the class whose mean vector is closest to the pixel vector; A new set of cluster mean vector is then calculated from this classification, and pixels are reassigned accordingly. In each iteration, K-means will tend to migrate to the true concentrations of data. The iterations are continued until there is no significant change in terms of the net mean migration from one iteration to the next. The final, stable result is not sensitive to the initial specification of seed vector, but more iteration may be required for convergence if the final vectors are not close to the seed vectors. The final cluster mean vectors may be used to classify the entire image with a minimum-distance classifier in one additional pass, or the covariance

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

matrices of the clusters may be calculated and used with the mean vectors in a maximum-likelihood classification.

2) ISODATA Clustering: The Iterative Self-Organizing Data Analysis Technique (ISODATA) represents a comprehensive set of heuristic (rule of thumb) procedures that have been incorporated into an iterative classification algorithm. Many of the steps incorporated into the algorithm are a result of experience gained through experimentation. ISODATA is self-organizing because it requires relatively little human input. The ISODATA algorithm is a modification of the k-means clustering algorithm. It includes merging clusters if their separation distance in multi-spectral feature space is below a user-specified threshold and Rules for splitting a single cluster into two clusters. ISODATA is iterative because it makes a large number of passes through the remote sensing dataset until specified results are obtained, instead of just two passes. Better results will be obtained if all bands have the similar data ranges.

C. Per-pixel classification approaches

Traditional per-pixel classifiers typically develop a signature by combining the spectra of all training-set pixels for a given feature. The resulting signature contains the contributions of all materials present in the training pixels, but ignores the impact of the mixed pixels. Per-pixel classification algorithms can be parametric or non-parametric. The parametric classifiers assume that a normally distributed dataset exists, and that the statistical parameters (e.g. mean vector and covariance matrix) generated from the training samples are representative. However, the assumption of normal spectral distribution is often violated, especially in complex landscapes. In addition, insufficient, non-representative, or multimode distributed training samples can further introduce uncertainty to the image classification procedure. Another major drawback of the parametric classifiers lies in the difficulty of integrating spectral data with ancillary data. The MLC may be the most commonly used parametric classifier in practice, because of its robustness and its easy availability in almost any image-processing software.

D. Subpixel classification approaches

Subpixel classification approaches have been developed to provide a more appropriate representation and accurate area estimation of land covers than per-pixel approaches, especially when coarse spatial resolution data are used. One major drawback of subpixel classification lies in the difficulty in assessing accuracy. One example of sub-pixel classification is the spectral un-mixing. Algorithms for spectral un-mixing uses a variety of different mathematical techniques to estimate end-members and abundances. The decomposing of the complete end to end un-mixing problem comprises a sequence of three consecutive stages: 1) Dimension reduction 2) End-member determination and 3) Inversion (Keshava, 2003).

E. Per-field classification approaches

The heterogeneity in complex landscapes results in high spectral variation within the same land-cover class. With per-pixel classifiers, each pixel is individually grouped into a certain category, and the results may be noisy due to high spatial frequency in the landscape. The per-field classifier is designed to deal with the problem of environmental heterogeneity, and has shown to be effective for improving classification accuracy. The per-field classifier averages out the noise by using land parcels (called 'fields') as individual units Geographical information systems (GIS) provide a means for implementing per-field classification through integration of vector and raster data. The vector data are used to subdivide an image into parcels, and classification is then conducted based on the parcels, thus avoiding intra-class spectral variations.

Spectral un-mixing is the procedure by which the measure spectrum of a mixed pixel is decomposed into a collection of constituent spectra, or end-members, and a set of corresponding fractions or abundances, that indicate the proportion of each end-member present in the pixel. End-members normally correspond to familiar macroscopic objects in the scene, such as water, soil, metal, any natural or man-made materials (Keshava, 2003).

F. Hard Classification

In hard (crisp) classification each pixel is forced or constrained to show membership to a single class as shown in Fig. 5. The characteristics of Hard classification are:

- Making a definitive decision about the land cover class that each pixel is allocated to a single class.
- The area estimation by hard classification may produce large errors, especially from coarse spatial resolution data due to the mixed pixel problem.

Examples of hard classifiers are maximum likelihood, minimum distance, etc.

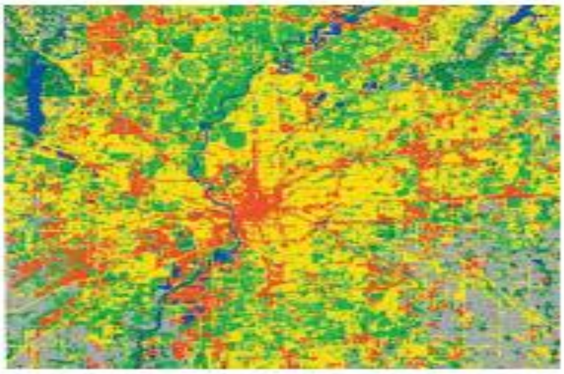

Fig. 5 Example of Hard Classification

G. Soft Classification

In soft classification each pixel may display multiple and partial class membership as shown in Fig. 6. Soft classification has been proposed in the literature as an alternative to hard classification because of its ability to deal with mixed pixels. The characteristics of soft classification are:

- Providing for each pixel a measure of the degree of similarity for every class.
- Soft classification provides more information and potentially a more accurate result, especially for coarse spatial resolution data classification.

Examples of soft classifiers are: Fuzzy-set classifiers, subpixel classifier, Artificial Neural Networks and spectral mixture analysis.

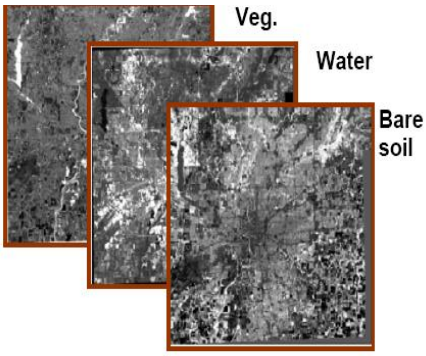

Fig 6. Example of Soft Classification

1) Artificial Neural Networks

Artificial Neural Networks are gaining popularity today as a tool for classification of remotely sensed image. While this algorithm has been successfully employed in various image classification problems, the errors associated with per-pixel classifier still persist, on a smaller scale in other methods. Artificial Neural Networks (ANN) has also been applied from remote sensing image classification to mixed-pixel handling (Buddhiraju, 2001).

Until recently, supervised classification of space-borne remotely sensed data has been achieved traditionally with Maximum Likelihood (ML) approach. The main problem with this statistical method is the assumption that the actual Probability Density Function (pdf) of the class in feature space follows a Gaussian distribution. The earth's feature is too complex to get fit into this simple distribution. Also, it is quite likely that a single class can have multiple representations in the feature space. For example, the class like urban can contain features not only buildings and roads but also features as grass and water bodies which are, in many a normal classification context, to be dealt with as separate classes other than the urban. Moreover, it is essential that the all classes should have non-singular (invertible) covariance matrices for implementing the ML.

To overcome these limitations, several research workers have explored artificial neural networks (ANNs) for low and high dimensional data classification problems. Even though the ANNs show great promise in high dimensional data classification the advantage is not significant when applied to low-dimensional multispectral data. It is commonly accepted now that the ML still yields good result for some classes for which the ANN shows poor result, and vice versa.

The ANN approach is understood to learn the underlying class-wise pdf from the data itself and is, by design, a nonparametric approach. This is true provided that there is no limit posed either on available training sample sizes or by computational resources. With the data complexity mentioned above, the classification performance of the ANN is largely decided by the choice of the net size, given the limited training data size practically possible by ground survey. In fact, the ANN of a fairly reasonable size can learn completely training data set, but it cannot guarantee a good generalization performance over data outside the training set. A way to improve the generalization is to train multiple classifiers and employ appropriate consensus schemes to combine their results.

Artificial Neural Networks are software and hardware models inspired by the structure and behaviour of biological neurons and nervous system, but after this point of inspiration all resemblance of biological systems ceases. There are about 50 different types of neural networks in use today. ANN is a parallel distributed processor that has a natural tendency for storing experimental knowledge. Image classification using neural networks is done by texture feature extraction and then applying the back propagation algorithm.

ANNs can provide suitable solutions for problems, which are generally characterized by non-linearities, high dimensionality noisy, complex, imprecise, and imperfect or error prone sensor data, and lack of a clearly stated mathematical solution or algorithm. A key benefit of neural networks is that a model of the system or subject can be built just from the data. Supervised learning is a process of training a neural network with examples of the task to learn, ie, learning with a teacher. Unsupervised learning is a process when the network is able to discover statistical regularities in its input space and automatically develops different modes of behaviour to represent different classes of inputs.

2) Genetic Algorithm

The techniques of image classification ranging from maximum likelihood to neural networks depend on the feature vectors formed by the intensity values in each spectral channel for each pixel. But the spectral information alone is not sufficient to exactly identify a pixel. The features of its neighbourhood, like texture, or the average value of nearby pixels are necessary to get good spectral information. Hence to choose these features automatically a new evolutionary hybrid genetic algorithm is used.

Genetic algorithm is based on the assumptions that computation or development of scoring function is non-trivial. Genetic algorithm can be used in feature classification and feature selection. It is primarily used in optimization. It can handle large, complex, non differentiable and multimodal spaces. It is good at refining irrelevant and noisy features selected for classification.

3) Decision Tree

Decision tree is one of the inductive learning algorithms that generate a classification tree to classify the data. Decision tree is non parametric classifier. Decision tree is an example of machine learning algorithm. They involve a recursive partitioning of the feature space, based on a set of rules that are learned by an analysis of the training set. A tree structure is developed where at each branching a specific decision rule is implemented, which may involve one or more combinations of the attribute inputs. A new input vector then "travels" from the root node down through successive branches until it is placed in a specific class. . It is based on the "divide and conquer" strategy.

The classification tree is made by recursive partitioning of the feature space, based on a training set. At each branching, a specific decision rule is implemented, which may involve one or more combinations of the attribute inputs or features. The thresholds used for each nodal decision are chosen using minimum entropy or minimum error measures. It is based on

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

using the minimum number of bits to describe each decision at a node in the tree based on the frequency of each class at the node. With minimum entropy, the stopping criterion is based on the amount of information gained by a rule.

A decision tree is composed of a root node, a set of interior nodes, and terminal nodes, called "leaves". The root node and interior nodes, referred to collectively as non-terminal nodes, are linked into decision stages. The terminal nodes represent final classification. The classification process is implemented by a set of rules that determine the path to be followed, starting from the root node and ending at one terminal node, which represents the label for the object being classified. At each non-terminal node, a decision has to be taken about the path to the next node. The advantages of decision tree classifier over traditional statistical classifier include its simplicity, ability to handle missing and noisy data, and non- parametric nature i.e., decision trees are not constrained y any lack of knowledge of the class distributions.

4) Fuzzy Classifiers

Zadeh in 1965 proposed a new theory called "Fuzzy Sets". Fuzzy classification attempts to handle the mixed pixel problem by employing the fuzzy set concept in which a given pixel ma have partial membership in more than one category fuzzy set classification belongs to the soft classification category. In this classification the assumption is that each pixel has more than one class and classification can be made using either supervised or unsupervised techniques. Fuzzy memberships allow a pixel to belong to more than one class with a degree of membership in each class.

Representing a geographical object is very difficult, as in most of the cases they do not have well defined boundaries, meaning that the boundaries between different phenomena are fuzzy, and/or there is heterogeneity within the class. If the class does not have sharp boundary then the assignment of the pixel to a class is uncertain and this uncertainty can be expressed by fuzzy class membership function. Fuzzy set theory provides useful concepts and methods to deal with uncertain information. It is achieved by applying a function called "membership function" on remotely sensed images.

Fuzzy memberships differ from probabilities interpretation. A fuzzy set is defined by a membership function, which defines how each point in the input set is mapped to a membership value (or degree of belongingness) between 0 and 1. If the membership value of an element is 0, it means that, it does not belong to that set and if it is 1, then it belongs fully to the set. In crisp sets, the membership value is either 1 or 0. In probability theory only one class or set is present and it expresses the degree to which its presence is likely as a probability. The class with highest probability is interpreted as actual class. Fuzzy set theory accepts that multiple classes or sets can be present at one place or at one time, and expresses the possibility to which each class or set is present as a membership value or belongingness value (Hedge, 2003).

The set associated with a membership function and each element in this set has its own membership value towards that particular set. The membership values range between 0 and 1. If the membership value of an element is 0, it means that, it does not belong to that set and if it is 1, then it belongs completely. But, in crisp sets, the membership value is 1 or 0. For fuzzy classification, this function takes values between 0 and 1. Therefore every pixel has certain membership values in every class. For example, a vegetation classification might include a pixel with grades of 0.68 for class "forest", 0.29 for class "urban" and 0.03 for "riverbed". We can see that pixel has higher membership value in class forest than other classes, and therefore it will be assigned to forest class. Fuzzy logic makes no assumption about statistical distribution of the data and it provides more complete information for a thorough image analysis, such as fuzzy classification results. It is interpretable and can use expert knowledge and training data at the same time.

The fuzzy knowledge based system should have the following three important stages:

- *Fuzzification of input data:* In this stage the crisp input is transformed into fuzzy set, which can be used in the fuzzy inference. In the satellite images, the DN values for a particular class is a crisp set and is fuzzified by applying the membership functions to the satellite images to classify them into different classes.
- *Fuzzy Inference for Analysis:* The fuzzy implications operators like IF, THEN, ELSE rules are generated in this stage for the purpose of analysis.
- *Defuzzification of Fuzzy Output:* This process converts the fuzzy values back to crisp values once the analysis is over. The map has to be 'hardened' to represent the output classes as one cannot represent two or more classes for a pixel in a single output map. Hardening is done by the defuzzification process using fuzzy operators like MAXIMUM, MINIMUM, etc. (Hegde, 2003).

The popular fuzzy set based approaches are the fuzzy c-means clustering (FCM), the probabilistic c-means clustering (PCM), as well as the fuzzy supervised classification (Thomas Blaschke et. al., 2002).

H. Hybrid Classification

This technique involves the aspects of both supervised and unsupervised classification and are aimed at improving the accuracy or efficiency (or both) of the classification process. Hybrid classifiers are valuable in analyses where there is complex variability in the spectral response patterns for individual cover types present Guided clustering is a hybrid approach which is used in applications like vegetation mapping. In guided clustering the analyst delineates numerous "supervised-like" training sets for each cover type to be classified. The data from all the training sites are then used in an unsupervised clustering routine to generate several spectral signatures. These signatures are examined by the analyst; some may be discarded or merged and the remainder is considered to represent spectral subclasses of the desired information class. Once sufficient subclasses are obtained, a maximum likelihood classification is performed. The spectral classes are then aggregated back into the information classes.

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

IV. Evaluation Of Classifiers

Classification process is not complete until its accuracy is assessed. Accuracy assessment can be performed by comparing two sources of information (Jensen, 1996): Remote-sensing derived classification data and reference test data. The relationship of these two sets is summarized in an error matrix where columns represent the reference data while rows represent the classified data. An error matrix is a square array of numbers laid out in rows and columns that expresses the number of sample units assigns to a particular category relative to the actual category as verified in the field. A classification accuracy assessment generally includes three basic components: sampling design, response design, and estimation and analysis procedures. Selection of a suitable sampling strategy is a critical step. The major components of a sampling strategy include sampling unit (pixels or polygons), sampling design, and sample size. Possible sampling designs include random, stratified random, systematic, double, and cluster sampling.

Before implementing a classification accuracy assessment, one needs to know the sources of errors. In addition to errors from the classification itself, other sources of errors, such as position errors resulting from the registration, interpretation errors, and poor quality of training or test samples, all affect classification accuracy. In the process of accuracy assessment, it is commonly assumed that the difference between an image classification result and the reference data is due to the classification error. However, in order to provide a reliable report on classification accuracy, non-image classification errors should also be examined, especially when reference data are not obtained from a field survey.

The accuracy of a classification has traditionally been measured by the overall accuracy. The overall accuracy of classification is obtained by dividing the sum of the correctly classified pixels (i.e. summed up values on the major diagonal of the error matrix) by the total number of pixels classified of the reference points. The kappa statistic, also called *KHAT* value, is a measure of how well the classification agrees with the reference data. It is also a measure of overall accuracy [R. G. Congalton et. al., 1986] and most commonly employed to evaluate the performance of a classifier.

But the overall accuracy alone gives no insight into how well the classifier is performing for each of the different classes. In particular, a classifier might perform well for a class which accounts for a large proportion of the test data and this will bias the overall accuracy, despite low class accuracies for other classes. To avoid such a bias, it is important to consider the individual class accuracy under producer's accuracy and user's accuracy. Producer's accuracy is a measure of the probability of a reference pixel being correctly classified and also called a measure of omission error. It is obtained by dividing the total number of correct pixels in a category by the total number of pixels of that category as derived from the reference data [R. G. Congalton, 1989]. User's accuracy can be obtained by dividing the total number of correct pixels in a category by the total number of pixels that were classified in that category and also called a measure of commission error. It is an indicative of the probability that a pixel classified on the image actually represents that category on the ground [R. G. Congalton, 1991].

There are many kinds of accuracy assessment techniques like spatial accuracy, thematic accuracy, temporal accuracy and topological accuracy. Spatial accuracy assessment is the determination of positional accuracy of objects (points, lines, polygons, or pixels) relative to known locations. Thematic accuracy concerns the measure of errors in the attributes associated with the objects. Thematic accuracy is assessed by comparing the reported values with that of the standard values. Topological accuracy also called the logical consistency is measuring the errors associated more with the processed data than interpretation. Temporal accuracy assessment has not much importance as in large scale map preparation; very negligible change may occur in between the field observation and map preparation. When performing accuracy assessment for the whole classified image, the known reference data should be another set of data, different from the set used for training the classifier. If training samples are used as reference data then the result of the accuracy assessment only indicates how the training sample are classified, but does not indicate how the classifier performs elsewhere in a scene. The following are the most commonly used methods to do the accuracy assessment.

A. *The Error Matrix*

The error matrix approach is the one most widely used in accuracy assessment. In order to properly generate an error matrix, one must consider the following factors: (1) reference data collection, (2) classification scheme, (3) sampling scheme, (4) spatial autocorrelation, and (5) sample size and sample unit. After generation of an error matrix, other important accuracy assessment elements, such as overall accuracy, omission error, commission error, and kappa coefficient, can be derived. Previous literature has defined the meanings and provided computation methods for these elements. The Error matrix as shown in Table 1 is a square, with the same number of information classes that will be assessed as the row and column. Numbers in rows are the classification result and numbers in columns are reference data (ground truth). In this square, elements along the main diagonal are pixels that are correctly classified. Overall accuracy, user's accuracy, and producer's accuracy is calculated using error matrix.

TABLE I
ERROR MATRIX

	Reference Data					
Classification Data		Class1	Class2	…..	Class N	Row Total
	Class 1	a_{11}	a_{12}	…..	a_{1n}	$\sum_{K=1}^{N} a_{1k}$
	Class 2	a_{21}	a_{22}	…..	a_{2n}	$\sum_{K=1}^{N} a_{2K}$
	…….					………
	Class N	$_{n1}$	a_{n2}	…..	a_{nn}	$\sum_{K=1}^{N} a_{NK}$
	Column Total	$\sum_{K=1}^{N} a_{K1}$	$\sum_{K=1}^{N} a_{K2}$	……	$\sum_{K=1}^{N-} a_{KN}$	N = $\sum_{1,K=1}^{N} a_{iK}$

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

1) Overall Accuracy: Overall accuracy is the proportion of all reference pixels, which are classified correctly. It is computed by dividing the total number of correctly classified pixels (the sum of elements along the main diagonal) by the total number of reference pixels. According to the error matrix above, the overall accuracy can be calculated as:

$$OA= \frac{\sum_{K=1}^{N} a_{kk}}{\sum_{i,k=1}^{N} a_{ik}} = \frac{1}{n}\sum_{K=1}^{N} a_{kk} \quad \ldots\ldots\ldots \quad (1)$$

2) Producer's Accuracy: Producer's accuracy estimates the probability that a pixel, which is of class I in the reference classification, is correctly classified. It is estimated with the reference pixels of class I divided by the pixels where classification and reference classification agree in class I. Producer's accuracy tells how well the classification agrees with reference classification. Given the error matrix above, the producer's accuracy can be calculated as:

$$PA\ (class\ I) = \frac{a_{ii}}{\sum_{i=1}^{N} a_{ki}} \quad \ldots\ldots\ldots \quad (2)$$

3) User's Accuracy: User's accuracy is estimated by dividing the number of pixels of the classification result for class I with the number of pixels that agree with the reference data in class I. User's accuracy predicts the probability that a pixel classified as class I is actually belonging to class I. It can be calculated as:

$$UA\ (class\ I) = \frac{a_{ii}}{\sum_{i=1}^{N} a_{ik}} \quad \ldots\ldots\ldots \quad (3)$$

B. Kappa Statistics

The Kappa analysis is a discrete multivariate technique used in accuracy assessment for statistically determining if one error matrix is significantly different than another. Kappa Statistic is based on the difference between the actual agreement in the error matrix (i.e., the agreement between the remotely sensed classification and the reference data is indicated by the major diagonal) and the chance agreement, which is indicated by the row and column totals (i.e., marginals).

The Kappa coefficient is a measure of overall statistical agreement of an error matrix, which takes non-diagonal elements into account. Kappa analysis is recognized as a powerful method for analyzing a single error matrix and for comparing the differences between various error matrices. Modified kappa coefficient and tau coefficient have been developed as improved measures of classification accuracy. Moreover, accuracy assessment based on a normalized error matrix has been conducted, which is regarded as a better presentation than the conventional error matrix.

The error matrix approach is only suitable for 'hard' classification, assuming that the map categories are mutually exclusive and exhaustive and that each location belongs to a single category. This assumption is often violated, especially for classifications with coarse spatial resolution imagery. 'Soft' classifications have been performed to minimize the mixed pixel problem using a fuzzy logic. The traditional error matrix approach is not appropriate for evaluating these soft classification results. Accordingly, many new measures, such as conditional entropy & mutual information and parametric generalization of Morisita's index have been developed. However, one critical issue in assessing fuzzy classifications is the difficulty of collecting reference data. More research is thus needed to find a suitable approach for evaluating fuzzy classification results.

In summary, the error matrix approach is the most common accuracy assessment approach for categorical classes. Uncertainty & confidence analysis of classification results has gained some attention recently & spatially explicit data on mapping confidence are regarded as an important aspect in effectively employing classification results for decision making.

V. Conclusion

Image classification has made great progress over the past decades in the following three areas: (1) development and use of advanced classification algorithms, such as sub pixel, per-field, and knowledge-based classification algorithms; (2) use of multiple remote-sensing features, including spectral, spatial, multi-temporal and multi-sensor information; and (3) incorporation of ancillary data into classification procedures, including such data as topography, soil, road, and census data.

The mixed pixels continue to exist regardless of the size of the pixel or sensor resolution. The higher the resolution, the longer is the processing time and also the cost of the image data.

One drawback of the linear mixture model is that it does not account for certain factors such as multiple reflections, which can result in complex nonlinearities in the spectral mixing process. In this situation, a more sophisticated nonlinear spectral mixture model may be required and hence artificial neural networks (ANN) may be well suited as ANNs do not require the input data to have a Gaussian distribution and they do not assume that spectra mix linearly.

The knowledge extracted using decision tree approach gives better results than traditional statistical classifier such as maximum likelihood classifier. Due to its nonparametric nature it is easy to add ancillary layers to it. It also doesn't require any statistical assumption about the distribution of the training sets such as normal distribution as required by MLC. Decision trees are easy to train and they learn quickly from examples. The main advantage of the decision tree is that we can convert decision tree into classification rules.

Introduction of Neural Network for training increases the iterations and hence increases the overall time taken for classification. Neural networks are superior to statistical methods used, in terms of classification accuracy of the training data. It has the advantage that it is distribution-free and we therefore we don't have to know anything about the statistical distribution of the data. It also avoids the problem of determining how much influence a source should have in the classification, which remains a problem for statistical methods.

Contextual information is inherently used by the human visual system while inferring the content of the scene presented to it. It is an important fact that it is not taken into consideration in pixel based classifiers. Hence Contextual Classification adopting relaxation labeling would give better results both in Neural Network based and Fuzzy Logic based classification.

The mixed pixel classification problem could be solved efficiently by adopting soft classification methods based on Fuzzy Logic and Artificial Neural Networks. Time taken for Fuzzy Classification is less than the time taken for conventional Maximum Likelihood Classification.

However, there has been no demonstrable improvement in classification performance over a period of last 15 years. The mean value of the Kappa coefficient across all experiments was found to be 0.6557 with a standard deviation of 0.1980.

ACKNOWLEDGMENT

It is a pleasure to recognize the many individual who have helped me in completing this paper. I sincerely express heartiest thanks to my guide Dr. Shivaprakash Koliwad for all the technical guidance, encouragement and support throughout this process. Also, I thank to Dr. C. G. Patil (Director of MCF) for his external support and guidance.

REFERENCES

[1] Bosdogianni Panagiota, Maria Petrou, Josef Kittler, "Mixed Pisel Classification with Robust Statistics", *IEEE Trans. On Geoscience and Remote Sensing*, Vol. 35, No. 3, May 1997.

[2] Buddhiraju Krishna Mohan, "Contextual Refinement of Neural Network Classification using Relaxation Labelling Algorithms", *Paper presented at the 22nd Asian Conference on Remote Sensing*, Singapore, November 2001.

[3] Daniel Newland, "Evaluation of stepwise spectral unmixing with HYDICE date", *Centre for Imaging Science*, Rochester Institute of Technology, Rochester, l998.

[4] Gebbinck Maurice Stefen Klein, "Decomposition of mixed pixels in remote sensing images to improve the area estimation of agricultural fields", *A Thesis submitted to university of Nijmegen*, November 1998.

[5] Graeme G. Wilkinson, Member, IEEE, "Results and implications of s study of fifteen years of satellite image classification experiments", *IEEE Transaction on Geoscience and Remote Sensing*, Vol. 43, No. 3, March 2005.

[6] Hedge Santosh, "Modelling land cover change: A Fuzzy Approach", M.Sc. Thesis submitted to the *International Institute for Geo-information Science and Earth Observation*, December 2003.

[7] Joseph George, "Fundamentals of remote sensing", *Universities Press (India) Pvt. Ltd.*, 2nd Edition, 2005.

[8] Keshava Nirmal, "Survey of spectral unmixing algorithms", *Lincon Laboratory Journal*, Vol. 14, No.1, 2003.

[9] Lillesand, Thomas M, Ralph Kiefer, Chipman, "Remote sensing and interpretation", *John wile and sons Inc.*, 5th edition, pp 579-583, 2004.

[10] Melgani Farid, Bakir A R, Al Hashemy, Saleem M. R., Taha, "An explicit fuzzy supervised classification method for multispectral remote sensing images", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, Bo. 1, January 2000.

[11] Nedelijkovic, "Image classification based on fuzzy logic", *Map soft Ltd.*, Zahumska, 2005.

[12] Raymond Bonnet, Campbell B, Campbell, "Introduction to remote sensing", *Taylor and Francis*, 3rd edition, pp. 277-280, 2002.

[13] Robert A., Schowrnger D T., "Remote sensing", *Academic Press*, 2nd edition, pp. 389-459, 1997.

[14] Tateishi R., Shimazaki Y., Gunin P. D., "Spectral and temporal linear mixing model for vegetation classification", *Int. Journal of Remote Sensing*, Vol. 25, No. 20, 4203-4218, October 2004.

[15] Thakkar Pathik, "Mixed pixel analysis", GIS Master Project, School of Social Sciences, *University of Texas at Dallas*, 2004.

[16] Thomas Blaschke, Stefan Lang, Eric Lorup, Josef Stroble, Peter Zeill, "Object oriented image processing in an integrated GIS/Remote Sensing Environment and perspectives for environmental applications", *Department of Geography and Geo-information*, University of Salzburg, 2002.

[17] Wang, Fangiu, "Fuzzy supervised classification of remote sensing images", *IEEE Transaction of Geoscience and Remote Sensing*, Vol. 28, No. 2, March 1990.

[18] Wang Y Z., Zhang S., "Remote sensing modeling in natural resources mapping: understanding the distribution and condition of biodiversity", *Department of Natural Resources Science*, University of Rhode Island Kingston, Rhode Island.

[19] Barandela R. and M. Juraez, "Supervised Classification of Remotely Sensed Data with Ongoing Learning Capability", IJRS, , Vol. 23, No. 22, 2002, pp. 4965-4970.

[20] Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques", 2nd ed., Elsevier, 2006.

[21] R. G. Congalton and A. Roy Mead, "A review of three discrete multivariate analysis techniques used in assessing the accuracy of remotely sensed data from error matrix", IEEE Trans. Geosci. Remote Sensing, Vol. GE 24, No. 1, pp. 169-174, Jan. 1986.

[22] R. G. Congalton, "Considerations and techniques for assessing the accuracy of remotely sensed data", Geosci. Remote Sens. Symp., Vol. 3, Jul. 10-14, 1989, pp. 1847-1850.

[23] Congalton R. G., "A review of assessing the accuracy of classification of remotely sensed data', Remote Sensing of Environment", 37, pp. 35–46, 1991.

[24] L. A. Zadeh, "Calculus of fuzzy restrictions, Fuzzy Sets and their Application to Cognitive and Decision Processes", IEEE Conference on Fuzzy Systems, 683-697, 1992.

[25] F. Aminzadeh and M. Jamshidi, "Soft Computing, Fuzzy Logic, Neural Networks and Distributed Artificial Intelligence", Prentice Hall, New Jersey, 1994.

[26] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance", IJRS, Vo. 28, No. 5, 823-870, March 2007.

A. L. Choodarathnakara was born in Hassan District of Karnataka State in INDIA. He received the B.E. degree in Electronics & Communication Engineering in the year 2002 and M.Tech. degree in Digital Electronics & Communication Systems in the year 2008, both from the Malnad College of Engineering, Hassan affiliated to Visvesvaraya Technological University, Belgaum, INDIA.

He is an Associate Professor of Electronics & Communication Engineering Department at Government Engineering College, Kushalnagar of South Kodagu District in Karnataka. His subjects of interest include Natural Language Processing, Language Learning Technologies, Text, Speech, Image, Signal & Remote Sensed Data Processing.

Asso. Prof. Choodarathnakara A. L. is a Life Member of the Indian Society of Technical Education (ISTE-43168), Indian Society of Remote Sensing (ISRS-3533) and Indian Society of Geomatics (ISG-1258).

Dr. Shiva Prakash Koliwad received the B.E. degree in 1974 from SJCE, Mysore (Mysore University), M. Tech. from IIT, Delhi, and the Ph. D. degree in 2000 from Anna University, Chennai, India. His subjects of interest are Image Processing, Signal & Remote Sensed Data Processing & Bio-medical Engineering. His specialization is in Split Band Bio-feedback Technique for bio-medical signal analysis.

At present, he is a Professor in the Department of Electronics and Communication Engineering at Malnad College of Engineering, Hassan. He is in teaching since 1975, and is a life member of the Indian Society for Technical Education (ISTE).

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

Dr. C. G. Patil received the M.Sc. degree in Physics from IIT, Kanpur and the Ph. D. degree from Karnatak University, Dharwad, India. His areas of interest are Satellite Image Processing, Remote Sensed Data Processing, Spacecraft Altitude Control, Communication, Mission Management and Expert Systems. His specialization is in Transport Properties in Ferromagnetic Semiconductors.

He is serving Master Control Facility since 1980 and presently is a Director of MCF, Hassan. Dr. C. G. Patil is involved in all the 17 mission operations of geostationary satellites controlled from MCF & operation director for INSAT-2A & INSAT-2B satellites. He is a recipient of Astronautical Society of India Award for the year 2001-02.

Dr. T. Ashok Kumar received the B.E. degree in Electronics and Communication Engineering from Malnad College of Engineering, Hassan (Mysore University) in the year 1990, and M.E. in Digital Electronics form SDMCET, Dharwad (Karnataka University), India, in 1995.

He is a Professor of Electronics and Communication Engineering, and currently he is serving as the Principal of Vivekananda College of Engineering and Technology, Puttur, Karnataka. He is in teaching for 20 years. His subjects of interest include Image Processing, Communication Engineering, Data Mining and Remote Sensing.

Prof. T. Ashok Kumar is a life member of Indian Society of Geomatics (ISG), Indian Society of Remote Sensing (ISRS), ISTE, and member of IEEE.