CHAPTER - 9

COMPARATIVE STUDY AND CONCLUSION

9.1. Comparative Analysis

The vagueness and uncertainty are inherent in the nature of remotely sensed spectral data. Now a days high resolution remote sensing images, the issue of mixed pixels continues to exist as a unique challenge to the remote sensing research community. The feature classification is an important application of remote sensing technology. It is a technology ascertaining the changes of specific features within a certain time interval. It provides the spatial distribution of features and qualitative and quantitative information of features changes in different dimensions. It involves the type, distribution and quantity of changes that is the ground surface types, boundary changes and trends before and after the changes in classification. The comparative study is an important part of any research area. The reasonable analysis of various remote sensing classification approaches are discussed below.

9.2. Comparison of Local, Dynamic Global and Interactive Adaptive Threshold Method in RSI Classification

The remote sensing image threshold method is a simple and straightforward concept. Its advantage is easy to understand and easy to use. Comparing with the method of threshold classification such as LTM, GTM and ATM method reduces the probability of errors. The background gray is restrained and foreground gray is enhanced in image selection processing. It is benefit for picking-up the information such as forest zone, estuaries and underwater channels ditch. It also has a good practical value so it is widely used in detecting coasting environmental changes, tropical forests changes, temperate forests changes, and desertification and crop analysis. A major problem is to find an optimized value for threshold. A small change in optimum threshold value destroys some important image details that may cause blur and artifacts. So, optimum threshold value should be found out which is adaptive to different sub band characteristics.

The histogram thresholding is one of the cheapest and fastest techniques for unsupervised image classification. In the present research, several thresholding techniques have been used and compared for distinguishing changes of multi-temporal remote sensing images. All the described techniques are used to detect thresholds from the histogram of the difference image that represents the modulus of the spectral change vectors associated with each pixel in the study area. Considering the present investigation, among the various thresholding techniques produced the best results compared to other threshold techniques. For all of the thresholding methods, the classification accuracies are higher than 80 % regardless of remote sensing data which validate the effectiveness of investigated methods. By analyzing Figure 3.3, Figure 3.5, Figure 3.7, Table 3.2, Table 3.5 and Table 3.7, we can get the following opinions. The features extracted with the LTM methods are suitable to threshold selection in RSI classification and the classification accuracy is 83.13 %. The Dynamic GTM method is the better, but it is not suitable for RSI classification and the classification accuracy is 77.86 %. The Interactive ATM method is the same in feature extraction in RSI classification and the classification accuracy can reach to 80.59 %. Finally compare to classification accuracy is obtained based on LTM, GTM and ATM feature extraction method, the LTM method is the best and acceptable accuracy of RSI Classification.

The main disadvantage of this method is that it does not reflect changes in categories. So the radiation value of deduction image does not always show the change of the objects because of a variety of factors such as atmospheric conditions, the sun light, sensor calibration, ground water conditions and so on. This method is too simplistic to detect the features of remote sensing images. It is difficult to take into account all factors and likely to cause the loss of information capacity. There may be a lot of noise (attracted by the image correlation). The threshold would be selected separately in change pixels area and unchanged pixels area on the histogram. In actual applications, the choice of threshold is quite difficult. Therefore it does not suitable for change detection of remote sensing areas. Some information would be loss in this simple method. So it is difficult to research the nature of changes and further analysis is needed to recover the next method k-means clustering method. The different thresholding methods are often compensated in the another K-Means Clustering analysis.

9.3. Comparison of K-Means Clustering Method in RSI Classification

Image classification of K-Means Clustering method is useful for the extraction of vegetation in remote sensing images. However, there is still some changes information not being detected such as the lane and the circular features, etc. The influence of the slope and aspect, the shadow or sun angle, radiation changing caused by strong seasonal transformation and multiplication noises can be eliminated or inhibited. It is the advantage of K-Means clustering method. It highlighted different slope features between the groups (Clusters). It has more exact accuracy and applies to research change detection of different image clusters and image nuclei. The choice of threshold value is difficult. Different features of the same gradient are easy to confuse and the clustering images are often compensated in the another PCA analysis. This is the main shortcomings of image clustering method.

From the precision of view, the feature selection of the remote sensing image using K-Means Clustering methods are more accurate, followed by the image thresholding methods. The K-Means clustering method and image Thresholding method have relatively simple operation and cost less time. The method of feature selection classification is more complicated and time consuming. From the application perspective, image threhsolding method of feature selection of remote sensing image classification is not suitable for coastal and forest based regions. The K-Means Clustering method can be used for remote sensing image, particularly analysis of vegetation, water, land and soil. By analyzing Figure 4.4 and Table 4.2, we can get following ideas. The K-Means Clustering method is the better than the threhsolding methods, but it is not suitable for all RSI classification and the classification accuracy is 85.49 percent.

9.4. Comparison of PCA Method in RSI Classification

Principal Component Analysis (PCA) is a mathematical technique for reducing the dimensionality of a remote sensing data set. Because digital remote sensing images are numeric, their dimensionality can be reduced by using this technique. In multi-band remote sensing images, the bands are the original variables. Some of the original bands may be highly correlated and such bands could be combined into new, less correlated Eigen value images by PCA. In this regard, we reduce the dimensions for the purpose that the feature information may not be redundant and the convergent speed of estimating the parameters of classifiers may be accelerated. As indicated, the information content of PCA bands decreases with an increasing number of PCA bands, and most of the information may only be contained in the first few PCA bands. This fact was confirmed by the following analysis performed for remote sensing imagery. In Figure 5.5, which contains 280*260 pixel samples of the bands, clearly shows that the majority of the variability is accounted for in the first few PCA bands and that the remaining bands quickly become noise. Similar studies collected in in the quantitative descriptions of the contents of PCA bands by using the ratios of the eigenvalue of each PCA band to the sum of all eigenvalues. Surprisingly, both the theoretical and experimental studies indicate that PCA bands with smaller eigenvalues may contain apparently visible information that is useful and can contribute to the image classification. As a matter of fact, this method suggested that inference cannot be made solely based on the magnitudes of the Eigen values, and a visual check of the obtained PCA bands is necessary and important.

This principal component is to detect the healthy vegetation component because healthy vegetation reflects highly in the near infrared regions. However, interpretation of the changes associated with the colour composite is very difficult without referring back to each of the input images to interpret the land cover characteristics at the suspected changed locations on each of the remote sensing image. The classification of feature selection result, which are illustrated in Figure 5.5 and the performance matrices are illustrated in Table 5.3. The best significant number of principal components used in visual display for just PC1, PC3, PC4, PC5 and PC7, which contained a cumulative total of 95 % of the data variance. The average number of principle component used in visual display for just PC2, PC6 and PC8, which contained a cumulative total of 86 % of the data variance.

Thus, the correct classification rate increases slowly in a nearly linear manner as the number of the PCA bands used for the classification increases. For remote sensing data sets, the use of the first few finest PCA bands can obtain a correct classification rate of about 95 percent, the second few average PCA bands leads to about 86 percent correct classification rate. The third worst level of PCA band images can get a foulest classification rate of about 75 percent. These PC bands are slightly decrease the correct classification rate. The PCA method is the better than the threhsolding methods and K-Means clustering, but it is suitable for specific RSI classification and the classification accuracy is 86.13 %. Therefore, the PCA approach can effectively ensure a practically acceptable and accurate classification result by handling only a small data set . The Table 5.3 also lists the computational time for the PCA analysis. The total time is the sum of the classification time and the PCA transformation time. For all of the dimension reduction methods, the classification accuracies are higher than 80 % regardless of image dataset which validate the effectiveness of investigated methods. Finally the best classification accuracy is obtained based on PCA then compared to LTM, GTM, ATM and K-Means Clustering method.

9.5. Comparison of FCM Clustering Method in RSI Classification

Fuzzy analysis, which is supported by whole fuzzy mathematic theory and fuzzy logic, presents great potential in dealing with spectral mixture problems in remote sensing image. Here we examined and compared three commonly used traditional classification methods with a high-resolution remote sensing images. We applied the modified FCM algorithm with $\alpha = 0.8$ on a noisy remote sensing image. The results are shown in Figure 6.5. To get the accurate size of a classified region, we calculated the partial volume using the membership value. Figure 6.5 shows the points that have a membership value less than the threshold T. The figure shows that those points are either near the regions boundaries or noise. In this case, we calculated the membership value of the surrounding pixels without the neighbor effect so that the resulting float would be a faithful value of the partial volume.

The volumetric measurements for structures can be accurately determined using the membership value as a guideline to get the partial volume at the boundaries. Using the neighbors to enhance the clustering with the FCM algorithm corrects for noisy images without affecting the edges. The effect of the neighboring pixels at the boundaries in a narrow real region could affect the region size after clustering. This is a fast and efficient method for color remote sensing image classification. The two most common factors are here discussed: First, a new initialization method of cluster centers, allowing faster convergence of the iterative algorithm is presented. Second, we generalize the enhanced fuzzy c-means algorithm to color by using a quantization of the color space. Thus, we keep the ability of the algorithm to be computed quickly. Moreover, using a quantization of the color space allows to obtain a more robust clustering: noisy pixels are allocated to bins shared by noise-free pixels. According to an extensive comparison with state of the art segmentation methods, our approach gives satisfactory results. Moreover, the computation time has been drastically reduced, enabling to process very large images in a reasonable time.

The proposed algorithm has been applied to well-known satellite images such coastal and forest regions. All the performance results have been reported in Table 6.4. The Vitality Factor (VF), Absolute Entropy (AE) and Distinct Entropy (DE) of Table 6.4 represent the optimal range for the number of clusters for the remote sensing images and sprinkles has also been copied from which is based on visual analysis by a group of ten experts. These results have been compared to those of Threshold Methods, KMC Method and PCA method. The results tabulated here for each image cluster, index and nuclei is the mean of 11 simulations. Images of Figure 6.5 are showthe classified output images using our proposed FCM method. The result of Table 6.5 is shown better than Threshold Methods, KMC Method and PCA method which always find a solution within the optimal range. The performance of the proposed algorithm with energetic and enhanced to produce comparable classification results. The result shows that the vitality factor of the proposed method is lowest than K-Means and normal FCM. This technique gives the better results to compare with others methods. The overall accuracy of proposed FCM method in the rate correct classifier is 87.07 %. The selection of the dataset used and different environmental settings also adds up to the performance of the classification accuracy.

9.6. Comparison of Active SVM Method in RSI Classification

The Support Vector Machine (SVM) is a classification method based on the statistical information of remote sensing images. Recently, particular attention has been devoted to support vector machines for the classification of hyper spectral remote sensing images . SVM is a non-parametric binary classifier that locates the optimal hyper plane between the two classes to separate them in a new high-dimensional feature space by taking into account only the training samples that lie on the edge of the class distributions known as support vectors. Moreover, it does not require the assumption of normality and is insensitive to the curse of dimensionality. SVMs have often been found to provide higher classification accuracies than other widely used pattern recognition techniques, such as threshold based methods, KMC Method, PCA Approaches and FCM Method. Furthermore, SVMs appear to be especially advantageous in the presence of heterogeneous classes for which only few training samples are available. The implementation of SVMs in multiclass classification problem is possible by formulating and optimizing the number of the classes needed to be classified and the number of parameters to be estimated to affects the SVMs classification performance in terms of accuracy.

The proposed active SVM learning algorithm provides the best segmentation quality as measured by the b index. This algorithm provides the lowest b-value which is expected use a very small number of training samples. The visual quality of the classified images is shown in the Figure 7.9. Among the supervised classification algorithms, namely, active SVM uses the least number of labeled samples and has minimum training time.

The results in terms of classification and performance matrices are summarized in Table 7.3. The active learning SVM exhibited the best overall accuracy of 94.50 %, i.e., the percentage of correctly classified pixels among all the test pixels considered. Moreover, for the comparative analysis of Threshold methods, KMC Method, PCA approaches and FCM method, the Active learning SVM is better to achieve the classification accuracy. Thus, for the comparative analysis, the performance matrix for the active SVM classifier is calculated as two levels such as finest and average levels. These two levels are nearly equal, but other traditional classification method has three distinct levels. The active SVM classifier except the spectral classes which shows the accuracy of finest level of 97 % and average level of 92 % respectively.

9.7. Comparison of Longitudinal Time Based CCL in RSI Classification

The remote sensing image consists of many oblique foreground pixel lines and background pixel lines. By algorithms, for each such line many provisional labels will be assigned to the pixels of the line, which will finally be combined into the same equivalent label set. Initializing so many equivalent label sets and combining so many equivalent label sets to require a time consuming work. Therefore, the work for initializing equivalent label sets is reduced greatly and we do not need to combine any equivalent label set. The multi pass 8-directing scan labeling approach also requires a multi pass through the image. This approach can also capture essential characteristics of the connected components like size, position and bounding rectangle. The labeling algorithms also resolve label equivalences on a region when merging with the neighbors of interest.

A majority of the algorithms give great results in terms of accuracy and faster execution time. Thus, the algorithms are very efficient for this image. The resultant matrices are shown in Table 8.7. The performance matrices are derived finest level and average level, but not in the case of worst level. Moreover, the execution time for calculating the number of connected components is propositional to the number of provisional labels assigned to connected components. The experimental results of our proposed algorithm is more efficient than the other traditional remote sensing classification algorithm. Different types of metrics are used detect the different levels of performance evaluation of image classification algorithm.

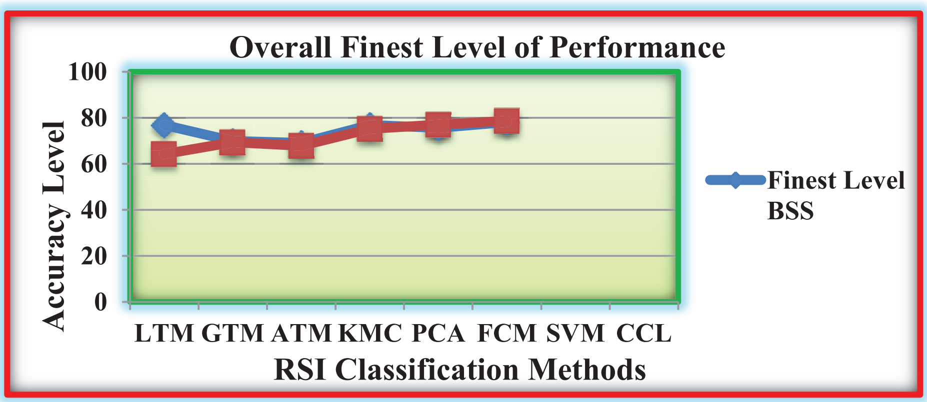
The success of image classification approaches is very much dependent on the quality of the image classification. The overall accuracy is calculated by dividing the number of correctly classified pixels by the total number of reference pixels. The basic advantage associated with performing the checks for a new equivalence in the class domain. Whatever the actual method used to manage equivalences, it is clear that handling a new equivalence results in a computational cost for the labeling process.

Results of the analysis revealed that there is a linear relationship between total number of pixels in ROI (or pixels in connected components) and total number of regions for given images are shown in the figure 8.7. Initially we performed a time based multi pass through a remote sensing image and calculated the total number of pixels in ROI based on user defined threshold value to distinguish background and ROI regions. This shows highly significant improvement compare to SVM results, the proposed CCL method and active SVM method are more or less in equal efficiency. This algorithm completes the connected component labeling with one scan on image and one scan on label connection table and finally, one scan on label matrix to assign final labels in label connection table to label matrix. But according to the results obtained from our experiment it shows that majority of images needed four scans to complete the connected component labeling Hence, it labeled the connected components efficiently and it requires acceptable computational time than others. Results obtained by the proposed technique are very accurate, especially when detecting the boundary pixels, and compared with the results obtained when applying the other techniques. It shows that the new proposed algorithm significantly performs excellent output and with minimum CPU time used for labeling , high accuracy and minimum storage, especially for large images.

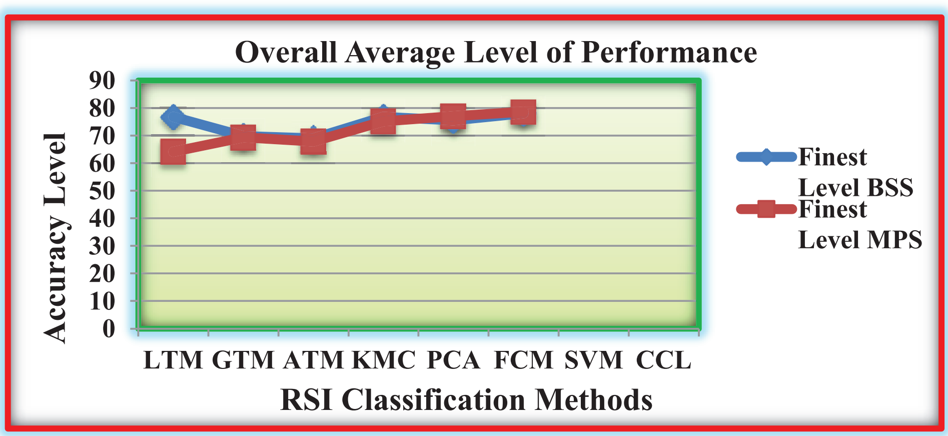
9.8. Performance Comparison of BSS and MPS

All eight algorithms comparison of various results and the performances are shown in Figures 9.1.

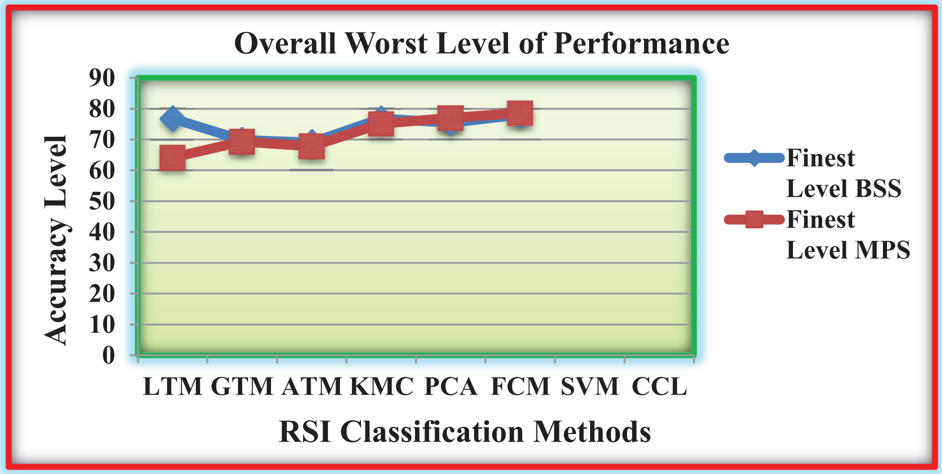
A



B



C



D

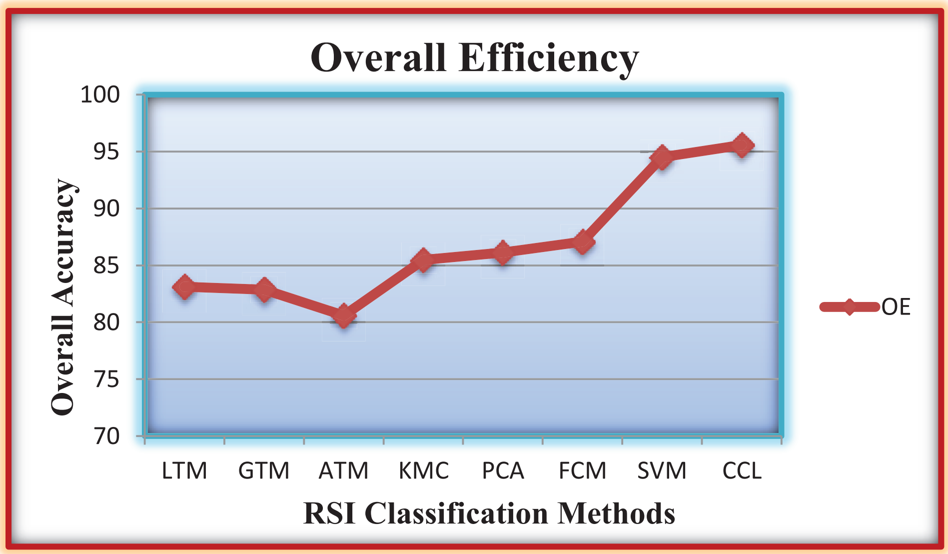


Figure 9.1. Comparison Graphs for Proposed RSI Classification Method, A. Overall Finest Level of Performance, B. Overall Average Level of Performance, C Overall Worst Level of Performance, D. Overall Efficiency

9.9. Conclusion and Recommendations

In the research, the problem of effective feature selection and extraction from remote sensing images have been addressed by developing several image processing and analysis algorithms based on the pattern recognition approach. The proposed techniques aim at solving different remote sensing data analysis problems taking into account the specific issues involved by the typology of employed data. In fact, the growing availability of remote sensing imagery of the

Earth surface, granted by recent and future space missions for Earth observation, provides huge potentialities for environmental management and monitoring, but also claims for accurate image analysis procedures. The proposed methods aim at providing feasible solutions to several open problems in remote-sensing image analysis, in order to exploit effectively such potentialities. In the following paragraphs, the main conclusions about all the developed methods are drawn.

Specifically, the availability of satellite constellations providing repetitive coverage with a very short revisit time represents an important support to environmental monitoring applications, but also requires the definition of effective unsupervised or partially supervised multi-temporal analysis techniques providing accurate classification results for all the acquisition dates without requiring a training set at each date. In this context, a supervised and unsupervised feature selection classification algorithms have been proposed in this thesis. The former deals with couples of images acquired over the same area at different periods and exploits training data and experimental validation over remote sensing image data has highlighted the capability of the methods to generate classification maps with high values are discussed. This research forms the principles of different methods achieving with MATLAB software, analyzing their results and comparing their advantages and disadvantages. At present, people study the change detection automatic technology on the computers. In addition, effective feature classification evaluated and expected the overall accuracy and also allowing an identification of spectral and spatial classes are presented. Globally, the set of presented processing techniques has been developed by integrating concepts drawn from the remote-sensing and the pattern-recognition disciplines in order to exploit the methodological data analysis framework provided by pattern recognition to extract effectively the information conveyed by remote-sensing imagery for Earth observation.

From a classification viewpoint, hyper spectral remote sensing data are known to provide strong class discrimination capabilities, although the resulting high number of features yields different dimensionality issues, usually requiring preliminary feature selection and reduction stages. In this research, an innovative feature extraction algorithm has been proposed which combines the flexibility of feature transformation methods with the availability of a physical meaning for the generated features. The method processes a set of hyper spectral bands in order to generate a set of synthetic multi spectral bands, specifically enhanced for supervised classification algorithms. Three thresholding techniques have been developed by extending to the present feature extraction context five corresponding distinct methods developed in the feature selection context. The method has been tested on real hyper spectral data, providing classification accuracies higher than the ones obtained by a reference feature-selection method and very similar to the ones of a benchmark feature extraction method.

More generally, the application of the pattern-recognition methodologies to the analysis of remote-sensing EO data still presents many open issues, being of interest both from an application-oriented and from a scientific viewpoint. In particular, a fully operational exploitation of the information-extraction potentialities provided by remote sensing data processing still requires to fill several technological gaps. Focusing specifically on the data-analysis issues, for instance, it is worth noting that most remote-sensing image processing methodologies are developed and tested in a laboratory operational situation, but their applicability in integrated environmental planning/monitoring systems to large-scale imagery in real-world situations would require an extensive phase of adaptation and optimization of the related techniques and of further experimental validation. In addition, the classification and estimation accuracies granted by the currently available methods are often still not enough to support a reliable exploitation of the resulting processed data in applicative contexts.

On the other hand, the huge number of distinct processing methods addressing a classification or estimation task in remote sensing would also require the definition of univocal and effective benchmarking criteria which is still an open problem in the remote-sensing community. Current and common remote sensing classification methods are discussed, particularly analysis and comparison of Threhsolding methods (LTM, GTM,ATM), KMC, PCA, FCM, SVM and CCL methods establishing a unified accuracy evaluation system are also explored. In this regards, eight feature extraction methods are comparatively studied. The results may be different with different data sets, but as a whole the CCL and SVM methods are better while extracting the key features of the multispectral remote sensing images. Occasionally, FCM and PCA are also a better method to extract the key features and the Thresholding Methods (LTM, GTM and ATM) and KMC methods are the vilest ones among all of the eight feature extraction methods. Therefore, when classifying the remote sensing images, we suggest that the CCL or SVM method should be used to extract the features in order to obtain higher classification accuracy.

From the perspective of the present research, interesting developments would concern addressing such open issues in order to optimize further the developed techniques aiming at their practical application in environmental management and monitoring schemes and allowing a further complete assessment of their processing capabilities in such operational contexts. It is also recommended to add classification and as well as association rules for the classification of satellite images. With the addition of association rules we can expect higher accuracy of image classification.