Fusion of ETM+ Multispectral and Panchromatic Texture for Remote Sensing Classification

Mahesh Pal

***Abstract*—**This paper proposes to use ETM+ multispectral data and panchromatic band as well as texture features derived from the panchromatic band for land cover classification. Four texture features including one 'internal texture' and three GLCM based textures namely correlation, entropy, and inverse different moment were used in combination with ETM+ multispectral data. Two data sets involving combination of multispectral, panchromatic band and its texture were used and results were compared with those obtained by using multispectral data alone. A decision tree classifier with and without boosting were used to classify different datasets. Results from this study suggest that the dataset consisting of panchromatic band, four of its texture features and multispectral data was able to increase the classification accuracy by about 2%. In comparison, a boosted decision tree was able to increase the classification accuracy by about 3% with the same dataset.

***Keywords*—**Internal texture; GLCM; decision tree; boosting; classification accuracy.

I. Introduction

ACCURATE classification of terrain from remotely sensed data is essential, especially for agricultural and forest monitoring, ecological monitoring of vegetation communities, land cover mapping and monitoring, and many other similar applications. Much work related to the classification of land use/land cover categories using satellite data is reported in the literature. To achieve an accurate classification of terrain, an image at a suitable resolution for the terrain needs to be acquired first, and then the characteristics of each small segment of the image must be classified accurately.

Several types of remote sensing data are being used in land cover classification. Data used can be obtained by using optical or microwave regions of the spectrum, and can be hyperspectral or multispectral in nature, depending on their availability and quality for a particular region. Usually medium spatial resolution satellite sensor data, such as Landsat ETM+ have been used in several researches. Data acquired by this sensor operate in the visible and near-infrared part of the electromagnetic spectrum. The availability of a panchromatic band of 15 m spatial resolution not only reduces mixed pixels, but also provides a rich texture and contextual information than multispectral bands with 30 m spatial resolution [1].

Mahesh Pal is with National Institute of Technology Kurukshetra, Kurukshetra, Haryana, 136119, India (e-mail:mpce_pal@yahoo.co.uk).

This paper will examine the potential of high spatial resolution panchromatic band and some of its texture features in combination with multispectral data for land cover classification. A univariate decision tree with and without boosting will be used to classify different datasets.

II. Dataset

The study area used in this paper is an agricultural area located near the town of Littleport in Cambridgeshire, in the eastern part of England. ETM+ data acquired on 19 June 2000 are used. The classification problem involved in identification of seven land cover types, namely, wheat, potato, sugar beet, onion, peas, lettuce and beans that cover the bulk of the area of interest. Ground resolution for ETM+ data is 30 m, except for the thermal band in which the resolution is 60 m. A panchromatic band with 15 meter resolution is also added for rectification and image sharpening. Landsat 7 provides data with a swath width of 185 km and a repeat coverage interval of 16 days.

A sub-image consisting of 307-pixel (columns) by 330-pixel (rows) covering the area of interest was used from both multispectral and panchromatic bands for subsequent analysis and classification. As the resolution of the ETM+ panchromatic data is 15m, bilinear resampling was used to reduce the resolution to 30m (i.e. the resolution of ETM+ multispectral data). Three GLCM texture features used for this study were extracted from the 30m resolution resampled panchromatic image. For this study, field data printouts for the relevant crop season were collected from farmers and their representative agencies, and other areas were surveyed on the ground to prepare the ground reference images.

III. Decision Tree Classifier

In the usual approach to classification, a common set of features is used jointly in a single decision step. An alternative approach is to use a multistage or sequential hierarchical decision scheme. The basic idea involved in any multistage approach is to break up a complex decision into a union of several simpler decisions, hoping the final solution obtained in this way would resemble the intended desired solution. Hierarchical classifiers are a special type of multistage classifier that allows rejection of class labels at intermediate stages. Classification trees offer an effective implementation of such hierarchical classifiers. Indeed, classification trees have become increasingly important due to their conceptual simplicity and computational efficiency. A decision tree classifier has a simple form which can be compactly stored and that efficiently classifies new data. Decision tree classifiers can perform automatic feature selection and complexity reduction, and their tree structure provides easily understandable and interpretable information regarding the predictive or generalisation ability of the classification. To construct a classification tree by heuristic approach, it is assumed that a data set consisting of feature vectors and their corresponding class labels are available. The decision tree is then constructed by recursively partitioning a data set into purer, more homogenous subsets on the basis of a set of tests applied to one or more attribute values at each branch or node in the tree. A number of approaches have been developed to split the training data at each internal node of a decision tree into regions that contain examples from just one class, and this is the most important element of a decision tree classifier. These algorithms either minimise the impurity of the training data or maximise the goodness of split. There are many approaches to the selection of attributes used for decision tree induction, and these approaches have been studied in detail by researchers in machine learning [2,3,4,5,6]. The procedure of creating a tree classifier involves three steps: splitting nodes, determining which nodes are terminal nodes, and assigning class label to terminal nodes. The assignment of class labels to terminal nodes is straightforward: labels are assigned based on a majority vote or a weighted vote when it is assumed that certain classes are more likely than others. A tree is composed of a root node (containing all the data), a set of internal nodes (splits), and a set of terminal nodes (leaves). Each node in a decision tree has only one parent node and two or more descendent nodes. A data set is classified by moving down the tree and sequentially subdividing it according to the decision framework defined by the tree until a leaf is reached. Decision tree classifiers divide the training data into subsets, which contain only a single class. The result of this procedure is often a very large and complex tree. In most cases, fitting a decision tree until all leaves contain data for a single class may overfit to the noise in the training data, as the training samples may not be representative of the population they are intended to represent. If the training data contain errors, then overfitting the tree to the data in this manner can lead to poor performance on unseen cases. To reduce this problem, the original tree can be pruned to reduce classification errors when data outside of the training set are to be classified.

A. Boosting

Boosting is a method used to improve the accuracy of any classifier by producing a series of classifiers. The training set chosen for a classifier depends on the performance of it earlier classifier. Sample, which is incorrectly classified by an earlier classifier, is selected more often than a correctly classified. Thus, boosting produce a new classifier, which is able to perform well on the new data set. In this study, a boosting algorithm called AdaBoost M1 [7]. Boosting assigns a weight to each observation - the higher the weight; the more that observation influences the classifier. At each trial, the vector of weights is adjusted to reflect the performance of the corresponding classifier, with the result that the weight of misclassified observations is increased. The final classifier aggregates the classifiers generated after each iteration by voting and each classifier's vote is a function of its accuracy. Studies carried out using boosting with a univariate decision tree classifier suggest that the resulting classifier perform quite well in comparison with individual classifier [8,9,10]. In order to evaluate the effects of boosting a total of 50 iterations were carried out with all three datasets with univariate decision tree as base classifier.

IV. Texture Feature Extraction

Methods to generate features based on combinations or transformations of primary features are called feature extraction. Image derived features, such as measures of spatial and spectral features may provide useful information for classification. Some features obtained by transforming primary features tend to suppress undesirable variability in remote sensing signatures, such as noise, so it is wise to use such features in classification because they allow the classifier to better distinguish spectral classes. A number of methods have been developed to deal with spectral and spatial information, in order to achieve improved classification performance. In comparison with tonal measures, the definition of texture features appears more difficult. The main difficulty faced by the researcher is to define a set of meaningful features to characterize the texture properties.

The extraction of texture features from high resolution remote sensing imagery provides a complementary source of data for some applications and found to be working well in improving the classification accuracy [1]. Based on the texture descriptors available in the literature, this study uses *correlation, entropy, and inverse different moment* texture features derived using grey level co-occurrence matrix [11] as well as a texture measure called *internal texture*. Internal texture was extracted from the panchromatic band in a way to reduce the image to 30m resolution. A program to calculate the difference between the maximum and minimum value in a 2×2 window was used for this purpose. The texture feature derived this way was able to reduce the image size from 15m to 30m resolution while simultaneously generating the internal texture image. The image generated by this procedure is georeferenced to the multispectral image and an area of 307 column and 330 rows was extracted for further study in combination with multispectral data.

V. Results

The aim of the present study is to evaluate the effect of the texture features obtained from the panchromatic data on the level of overall classification. The results obtained from two data sets were compared with ETM+ multispectral data. Two datasets used for the classification consists of the following combinations:

1. Combination of multispectral, panchromatic band and internal texture of panchromatic band referred to as data set 1.
2. Combination of data set 1 and three GLCM based texture features (correlation, entropy, and inverse different moment) of panchromatic band, referred to as data set 2.

Equalized random sampling plan was used to collect the training and test pixels from the ETM+ multispectral data as well as from datasets 1 and 2. The program used for this purpose uses ground reference image to collect the required pixels from different datasets. Selected pixels were divided in two parts in order to remove any bias in using same pixels for training and testing. Table I provides number of training and test pixels used in classification with different datasets.

Classifications were performed in order to evaluate the effects of various texture features on the level of classification accuracy achieved by the boosted and unboosted decision tree classifier. While using a decision tree classifier gain ratio as attribute selection measures and error-based pruning approach was used to prune overgrown decision tree [9].

Tables II and III provide the results obtained using different datasets with unboosted and boosted decision tree classifier respectively. Results suggests that dataset 2 perform well with or without boosting the decision tree classifier and an increase of about 2 to 3% in classification accuracy is achieved in comparison to the accuracy achieved with ETM+ multispectral data. Results from Table II suggest that the inclusion of the panchromatic band and internal texture feature with multispectral data does not increase classification accuracy by a large amount. Further, results with the data set 2 (Tables II and III) suggests that classification accuracies with this dataset improve by about 2 to 3%, thus suggesting the utility of panchromatic data and its texture for land cover classification.

TABLE I
TRAINING AND TEST DATA SET USED WITH DIFFERENT DATASETS

Dataset	Training pixels	Test pixels
ETM+	2700	2037
Dataset 1	2700	2034
Dataset 2	2700	2043

TABLE II
CLASSIFICATION ACCURACY AND KAPPA VALUES WITH DIFFERENT DATASETS USING A UNIVARIATE DECISION TREE CLASSIFIER

Dataset	Classification accuracy (%)	Kappa value
ETM+	83.55	0.810
Dataset 1	84.76	0.820
Dataset 2	85.76	0.834

TABLE III
CLASSIFICATION ACCURACY AND KAPPA VALUES WITH DIFFERENT DATASETS USING BOOSTING WITH A UNIVARIATE DECISION TREE CLASSIFIER

Dataset	Classification accuracy (%)	Kappa value
ETM+	87.78	0.857
Dataset 1	88.89	0.870
Dataset 2	90.75	0.892

VI. CONCLUSION

This paper was aimed at assessing the usefulness of fused multispectral, panchromatic data and its texture features for land cover classification. A major conclusion of this study is that the dataset consisting of multispectral, panchromatic and four tsture features derived from panchromatic data works well with or without boosting the decision tree classifier. Although 2% to 3% increase in classification accuracy may appear to be a small increase, it should be borne in mind that even small percentage increases are difficult to generate when the overall classification accuracy level exceeds 80%. Thus it can be concluded that the texture features derived from ETM+ panchromatic data can be used to improve the classification accuracy with decision tree classifier.

REFERENCES

[1] Lu, D and Weng, Q, Urban Classification Using Full Spectral Information of Landsat ETM+ Imagery in Marion County, Indiana, Photogrammetric Engineering & Remote Sensing, 71(11), 2005, 1275–1284.
[2] L.,Breiman, J.H. Friedman, R.A.,Olshen and C.J.,Stone, Classification and Regression Trees, Wadsworth, Monterey, CA, 1984.
[3] S. K.Murthy, S. Kasif and S. Salzberg, A system for induction of oblique decision trees. Journal of Artificial Intelligence Research, 2, 1994, 1-32.
[4] I. Kononenko and J. S. Hong, Attribute selection for modelling. Future Generation Computer Systems, 13, 1997, 181-195.
[5] J. Mingers, An empirical comparison of selection measures for decision tree induction. Machine Learning, 3, 1989, 319-342.
[6] J. R. Quinlan, C4.5: Programs for Machine Learning. San Mateo: Morgan Kaufmann, San Francisco, 1993.
[7] Y.Freund and R. Schapire, Experiments with a new boosting algorithm. Machine Learning: Proceedings of the Thirteenth International conference, 148-156, 1996.
[8] M. A. Friedl, C. E. Brodley, and A. H. Strahler, Maximizing land cover classification accuracies produced by decision tree at continental to global scales. IEEE Transactions on Geoscience and Remote Sensing. 37, 1999, 969-977.
[9] M. Pal and P. M. Mather, An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of Environment. 86, 2003, 554-565.
[10] G. J. Briem, J. A. Benediktsson, and J. R. Sveinsson, Multiple Classifiers Applied to Multisource Remote Sensing Data IEEE Transactions on Geoscience and Remote Sensing, 40, 2002, 2291-2299.
[11] R.M. Harlick, K. Shanmugam and I. Dinstein, Texture features for image classification. IEEE Transactions on System, Man, and Cybernetics, 3(6), 1973, 610-621.