COMPARISON OF OBJECT BASED AND PIXEL BASED CLASSIFICATION OF HIGH RESOLUTION SATELLITE IMAGES USING ARTIFICIAL NEURAL NETWORKS

B.K. Mohan and S. N. Ladha

Centre for Studies in Resources Engineering
IIT Bombay, Powai, Mumbai – 400 076

ABSTRACT

Segmentation and classification of high resolution imagery is a challenging problem due to the fact that it is no longer meaningful to carry out this task on a pixel-by-pixel basis. The fine spatial resolution implies that each object is now an aggregation of a number of pixels in close spatial proximity, and accurate classification requires that this aspect be considered. In this study we focused on classification of high resolution satellite images using Artificial Neural Network (NN) and compare two different classification approaches, Object and Pixel based classifications. Object based classification involves segmentation of input image. We used the morphological watershed transform to extract regions. A number of properties of the regions were computed – spectral mean vector, average texture, departure from circularity, length-to-breadth ratio, area, perimeter, compactness and others. Image was then classified on the basis of the regions instead of the pixels (as in Pixel based classification). Where the pixels were of the order of a few millions, the regions were of the order of a few thousands. The results are encouraging and the scheme developed in this study is being evaluated with a range of images and a number of other classifiers.

Index Terms— Image classification, Neural network application, Object oriented analysis, Segmentation, Watershed

1. INTRODUCTION

High resolution remotely sensed imagery offers an exciting possibility for feature extraction and spatial modeling. In the present study, a region based approach for classification of high resolution multispectral images has been presented.

The authors wish to thank the ISRO-IIT(B) Space Technology Cell for supporting this research through a grant and Head, CSRE for his encouragement during the course of the work.

Region based approaches do not operate directly on individual pixels but on regions consisting of many spatially adjacent pixels that have been grouped together in a meaningful way by image segmentation. In addition to the textural measures employed in the pixel-based classification methods, image objects also allow shape characteristics and neighborhood relationships. Their attributes are applicable to all the pixels inside the objects. This method basically includes three steps. 1) Image segmentation to extract the regions from the pixel information based on homogeneity criteria. 2) Calculation of spectral parameters like mean vector, texture, NDVI and spatial/shape parameters like aspect ratio, convexity, solidity, roundness and orientation for each region. 3) Classification of image using the region feature vectors using suitable classifiers such as NN.

2. IMAGE SEGMENATION AND FEATURE EXTRACTION

Segmentation of the images is generally done using three techniques, namely, edge based methods, region based methods and hybrid methods. Here our emphasis is on region based segmentation.

2.1. Region Growing Using Morphological Watershed Transformation

Region growing algorithms take one or more pixels, called seeds, and grow the regions around them based upon certain homogeneity criteria. If the adjoining pixels are similar to the seed, they are merged within a single region. The process continues until all the pixels in the image are assigned to one or more regions.

Watershed transformation is a powerful tool for image segmentation. Regions of terrain that drain to the same point are defined to be part of the same watershed. The same analysis can be applied to images by viewing intensity as height. In this case, the image gradient is used to predict the direction of drainage in an image. By following the image gradient downhill from each point in the image, the set of points (dark regions), which drain to each local intensity minimum, can be identified. These disjoint regions are called the watersheds of the image. Similarly, the gradients can be followed uphill to local intensity maximum in the image, defining the inverse watersheds (bright regions) of the image. See [1], [2] for detailed introduction to this area.

2.1.1. Watershed Transform Implementation

Watershed segmentation for high-resolution satellite images starts with preprocessing of image by suitable operators, prominent being alternating close-open filters defined in mathematical morphology. The Gaussian smoothing operator is also applied to `blur' images and remove excessive detail and noise. The second stage in segmentation is to locate regional maxima in the images. The regional maxima is a flat zone (a maximal connected component of grey scale image with sample pixel values) surrounded by flat zone of strictly lower grey values. One of the crucial steps in the watershed transform is marker extraction. A marker must be placed in representative sample of the region of the object to be extracted. One of the methods to find the marker is using the regional maxima of the filtered image. One advantage of these methods is its independence of grey-scale threshold value. Labels are assigned to the binary markers thus extracted. The gradient of the image is generated using Robert's gradient operator. Using the markers as seeds, the regions are grown by simulating a flooding process of the terrain, i.e., adding adjacent pixels to grow regions. When two growing floods (regions) meet, the region boundaries occur at those points.

2.2. Connected Component Labeling

Connected component labeling is a process where pixels in each non-overlapping region are given independent identity (label) so that the region parameters can be computed. Usually it takes at least two passes through the segmented image, once from top to bottom and then once from bottom to top to identify pixels falling within regions and label them. An algorithm for the same can be found in [3]. Once each region is uniquely labeled, then the region shape, size, average grey level / spectral / textural properties can be computed for each region. This gives an opportunity to deal with the image in an object oriented manner rather than on a per-pixel basis.

2.3. Region Parameters

While it is possible to compute a large number of parameters for each region, some of the prominent ones may be listed as follows
1) Size; 2) Spatial Moments; 3) Roundness, Convexity, and Solidity 4) Mean vector; 5) NDVI (for multispectral images); 6) Texture statistics (entropy, contrast, angular second moment, etc.).

3. IMAGE CLASSIFICATION USING ARTIFICIAL NEURAL NETWORKS

Many classifiers are available for classification of multispectral satellite images. These include discriminate analysis, maximum likelihood classification scheme, etc. A major disadvantage of these classifiers is that they are not distribution free. This has prompted in significant increase in use of NN for classification of remotely sensed images [4] and to a certain extent in fuzzy logic [5]. Several other reasons can be sited in favor of NN based classifiers as listed below [6].

1) Each of the (region) parameters will be in a different numerical range, some in [0,1], some in [0, 255], etc. Rescaling all parameters to a single range can affect the inter-class and intra-class separation; 2) NN classifiers can detect and use to their advantage non-linearity in data patterns; 3)
Ancillary data can be included in NN classifiers; 4) NN architectures are flexible which can be easily optimized for performance; 5) NN can handle multiple subcategories per class.

Much of the NN classification work in remote sensing has used Multi Layer Feed Forward (MLFF) NN. In this study we consider one more type of NN classifier called Radial Basis Function (RBF) NN. Following paragraphs provide a brief overview of both the classifiers [7].

3.1. Multi Layer Feed Forward Neural Networks

Typically an MLFF NN consists of a set of sensory units (source nodes) that constitute the input layer, one or more hidden layers of computation nodes, and an output layer of computation nodes. The input signal propagates through the

Fig. 1. High resolution image used for classification study

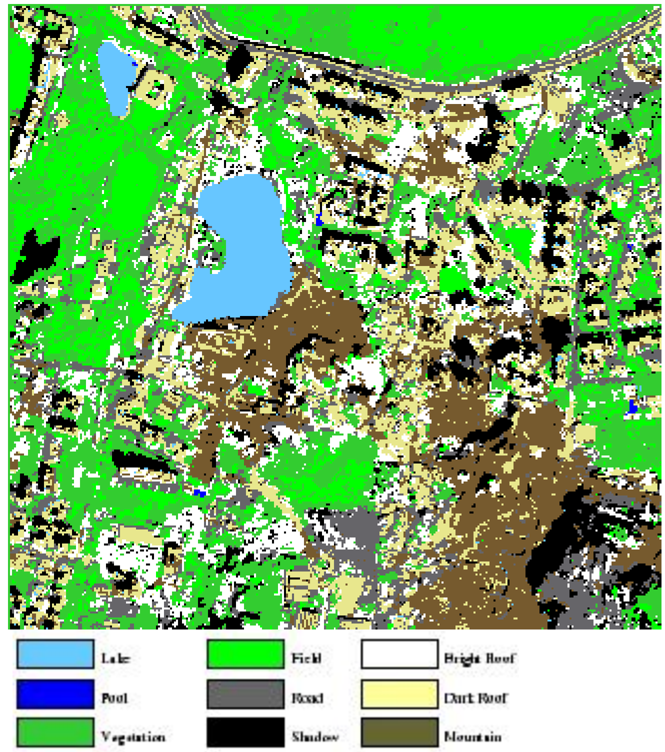

Fig. 2. Object oriented segmentation and classification by backpropagation neural network classifier

network in a forward direction on a layer-by-layer basis. Learning in MLFF NN consists of two passes through the different layers of the network: a forward pass and a backward pass. In the forward pass, an input pattern is applied to the sensory nodes of the network and its effect propagates through the network layer by layer. Forward pass is followed by a backward pass. During backward pass the error signal (difference between actual output of the network and the desired output) is propagated backward through the network, against the direction of synaptic connections. Hence the name back-propagation. The synaptic weights are adjusted to make the actual response of the network move closer to the desired response.

3.2 Radial Basis Function Neural Networks

Design of RBF NN can be viewed as a curve fitting approximation in a high dimensional space. Learning is equivalent to finding a surface in a multidimensional space that provides a best fit to the training data. Correspondingly generalization is equivalent to the use of this multidimensional surface to interpolate test data. The construction of a RBF NN, in its most basic form, involves three layers with entirely different roles. The input layer is made up of source nodes (sensory units) that connect the network to its environment. The second layer, the only hidden layer in the network, applies a nonlinear transformation from the input space to the hidden space; where hidden space is in general of high dimensionality. The output layer is linear, supplying the response of the network to the activation pattern applied to the input layer. A mathematical justification for the rationale of a nonlinear transformation followed by a linear transformation may be traced back to an early paper by [8]. According to this paper, a pattern-classification problem cast in a high dimensional space is more likely to be linearly separable than in a low-dimensional space.

4. EXPERIMENT

For the purpose of this study, we have implemented a MLFF NN and a RBF NN, in order to compare the performance of different types of NN classifiers.

A high resolution (2000 x 2000) multispectral remotely sensed image illustrating various types of land use and land cover was used as the test image for classification as shown in figure 1.

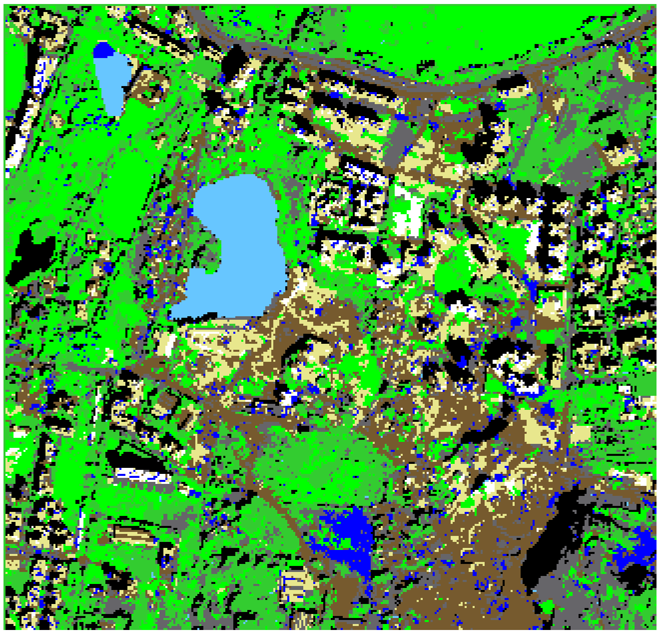

Fig. 3. Object oriented segmentation and classification by radial basis function network classifier

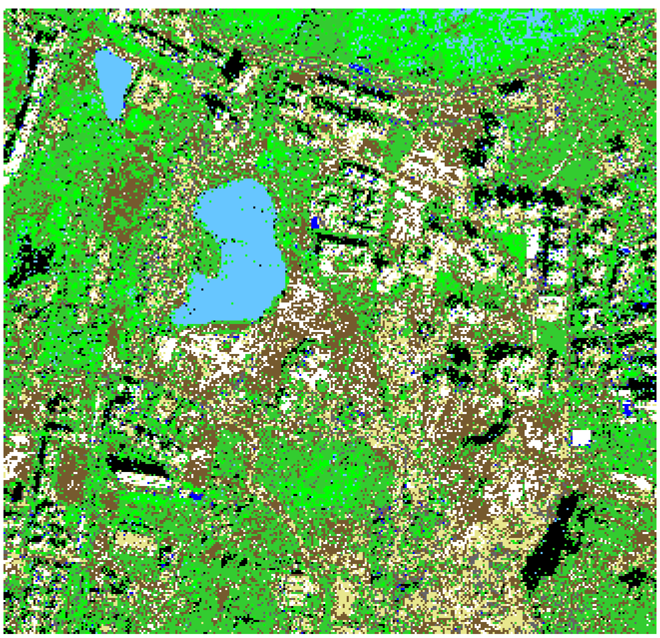

Fig. 4. Per-pixel classification using backpropagation neural network classifier

The image was classified into 9 prominent classes covering a majority of the land cover features, as shown in the legend of figure 2.

Both Object based and Pixel based (using spectral features for classification) approaches were used to classify the image using MLFF and RBF networks. Accuracy and error statistics were computed for each approach/NN combination. Figures 2. and 3. depict output of Object based classification using MLFF and RBF networks respectively. Figure 4. and 5. depict output of Pixel based classification using MLFF and RBF networks.

TABLE 1. Classification statistics of Object and Pixel based classification approaches

	Object Based		***Pixel Based***	
	MLFF	**RBF**	**MLFF**	**RBF**
Accuracy	0.8750	0.8472	0.7972	0.8083
Kappa Coeff	0.8794	0.8519	0.7685	0.8065
Producer's Accuracy				
Lake	1.0000	0.8750	0.8974	0.9127
Pool	0.8105	0.8045	1.0000	1.0000
Vegetation	1.0000	0.8750	0.8473	0.8494
Field	0.9000	0.8750	0.9308	0.8242
Road	0.8954	1.0000	0.5064	0.6873
Shadow	0.9000	1.0000	1.0000	0.9711
Bright Roof	0.7988	0.8873	1.0000	1.0000
Dark Roof	0.9059	0.7660	0.5760	0.6388
Mountain	0.8514	0.8000	0.5395	0.6445
Consumer's Accuracy				
Lake	1.0000	1.0000	0.9722	0.9730
Pool	1.0000	0.8571	0.9756	1.0000
Vegetation	0.8889	0.8750	0.6604	0.7143
Field	1.0000	0.8750	0.8810	0.9091
Road	0.7778	0.6667	0.5833	0.5357
Shadow	1.0000	1.0000	0.8889	0.9070
Bright Roof	0.8571	0.8750	0.9302	0.8670
Dark Roof	0.875	1.0000	0.7143	0.8947
Mountain	0.6364	0.6667	0.5417	0.6487

5. RESULTS AND CONCLUSION

Table 1. listed above gives a summary of accuracy and error statistics of the two classification approaches.

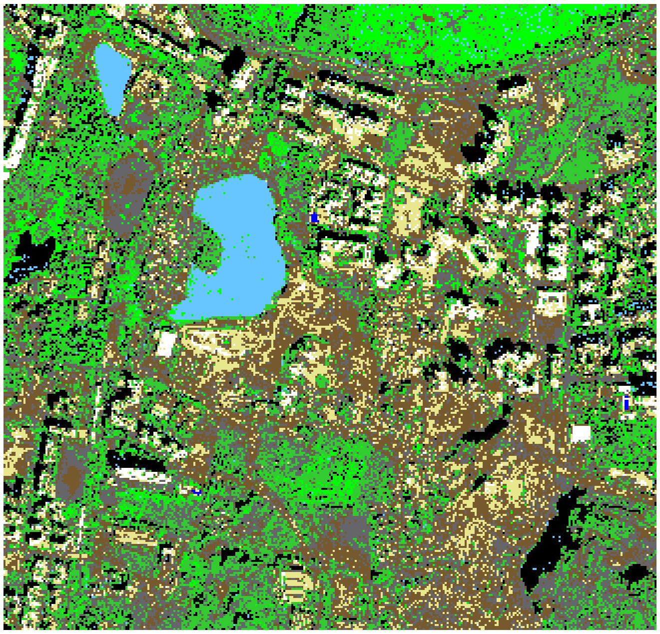

Fig. 5. Per-pixel classification using radial basis function neural network classifier

As seen in the above images the output of Object based classification is more smooth and coherent (more homogeneous region wise) as compared to that of Pixel based approach. With Pixel based approach differences in spectral density of pixels lying in close spatial proximity (belonging to the same region) are reflected in the output. As a result of which the output is very granular (pixelized).

Fig 3. clearly indicates that Object based classification is not a panacea, it is evident that regions are misclassified. For example, a part of the lake is being classified as a pool, an entire lake is classified as shadow, etc. This happens due to spectral closeness of these regions. In order to reduce misclassification we need to take into consideration ancillary data (contextual data) available about the image. For example, if a region is classified as a shadow then there has to be tall structure in the vicinity of the shadow.

In conclusion, Object based image classification approach yields better results than Pixel based approach. A combined approach to classification using Object based methods and contextual information available about the image, seems promising and needs further exploration.

6. REFERENCES

[1] F. Meyer, S. Beucher, "Morphological Segmentation," *JOURNAL. Visual Communication and Image Representation.*, 1990, vol. 1, pp. 21-46.

[2] P. Soille, "Morphological Image Analysis," Springer-Verlag, Berlin 2003.

[3] L. G. Shapiro, G. C. Stockman, "Computer Vision," Prentice Hall, NJ, 2001.

[4] P. M. Mather, "Computer Processing of Remotely Sensed Images," John Wiley, NY, 2004.

[5] B. M. Foody, "Image Classification with a Neural Network: From Completely Crisp to Fully Fuzzy Situations," *Advances in Remote Sensing and GIS Analysis.*,John Wiley, NY, 1999.

[6] J. Han, S. Lee, K. Chi, K. Ryu, "Comparison of Neuro-Fuzzy, Neural Network, and Maximum Likelihood Classifiers for Land Cover Classification using IKONOS Multispectral Data" *Geoscience and Remote Sensing Symposium 2002. IGARSS 02. 2002 IEEE International*, 2002, vol 6, pp. 3471-3473.

[7] S. Haykin, "Neural Networks, A Comprehensive Foundation," Second Edition, 1999, pp. 156, 256

[8] T. Cover, "Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition," *IEEE Transactions on Electronic Computers*, vol. EC-14, pp.326-334