Regular Research Articles

A Novel Technique for Automatic Extraction of Roads from High Resolution Satellite Remotely Sensed Images

Imdad A. Rizvi, and B.Krishna Mohan

Abstract

As the information carried in a high spatial resolution image is not represented by single pixels but by meaningful image objects, which include the association of multiple pixels and their mutual relations, the object based method has become one of the most commonly used strategies for the processing of high resolution imagery. This processing comprises of two fundamental and critical steps towards content analysis and image understanding i.e image segmentation and classification. Hence this paper proposes a robust object based segmentation algorithm using multiresolution analysis technique and object based supervised image classification using modified cloud basis functions (CBFs) neural network algorithm to identify road features from high resolution satellite remotely sensed images .

Keywords: *Object based image classification, Supervised Classification, and radial basis functions neural network.*

INTRODUCTION

AUTOMATIC road extraction is a critical feature for an efficient use of remote sensing imagery in most contexts. Roads on medium or low resolution satellite images usually appears with widths of one or two pixels and some details of roads cannot be observed e.g. vehicles, shadows, markings and trees along the roads etc, hence high resolution satellite remotely sensed images are used for extraction of roads. The object-based image approach is employed here for road extraction from high resolution satellite remotely sensed image because it can reduce the spectral variations during image segmentation. Some specific feature vectors are considered for extraction of road objects. A lot of research has been undertaken and is being carried out for developing an accurate classifier for extraction of roads, with varying success rates. Most of the commonly used classifiers are based on support vector machines, which use radial basis functions, for defining the boundaries of the classes. The drawback of such classifiers is that the boundaries of the classes as taken according to radial basis function networks are spherical while the same is not true for majority of the real data. The boundaries of the classes vary in shape, thus leading to poor accuracy in the case of support vector machine based classifiers. The new basis functions, called cloud basis functions (CBFs) use a different feature weighting, derived to emphasize features relevant to class discrimination (De Silva et al., 2008). Further, these basis functions are designed to have multiple boundary segments, rather than a single boundary as for RBFs. This new enhancements to the basis functions along with a suitable training algorithm allow the neural network to better learn the specific properties of the problem domain. Some existing methodologies for road extraction are discussed in section II. The proposed methodologies adopted in this paper are discussed in Section III. Section IV presents experimental result and discussions. Summary and conclusions drawn from this work are discussed in section V.

Review of literature

Automatic road extraction has been an active research area in computer vision and digital photogrammetry for over two decades. During the past 20 years, a number of semi-automatic and automatic methods and algorithms for road extraction have been developed. In literature, there are many methods described to extract line and point objects using both pixels based and object based approaches. The main problems with road extraction in urban areas are the more complex scene content and the different structure of the road network compared to rural areas. Urban or suburban scenes consist of many different objects like houses, trees and vehicles, which lead to a scene that is composed of many small regions (Benediktsson et al., 2003). Supervised classification is one of the most commonly undertaken analyses of remotely sensed data. The output of a supervised classification is effectively a thematic map that provides a snapshot representation of the spatial distribution of a particular theme of interest such as land cover. The goal of a supervised image classification system is to group images into semantic categories giving thus the opportunity of fast and accurate image search. To achieve this goal, these applications should be able to group a wide variety of unlabelled images by using both the information provided by unlabelled query image as well as the learning databases containing different kind of images labeled by human observers. In practice, a supervised image classification solution requires three main steps (Duda, 2001): pre-processing, feature extraction and classification. Based on this architecture, many image classification systems have been proposed, each one distinguished from others by the method used to compute the image signature and/or the decision method used in the classification step.

Artificial Neural Network (ANN) and Support Vector Machine (SVM) are commonly used advanced methods for supervised classification of remotely sensed data (Tso and Mather 2001). The neural network classifier is based on Empirical Risk Minimization (ERM) principle. Empirical risk minimization technique minimizes the misclassification error on the training set. From the range of network types suitable for classification applications (e.g. Ito and Omatu 1997, 1998, Carpenter et al. 1999), multi-layer feedforward networks are the most widely used in remote sensing. In particular, the multi-layer perceptron (MLP) network trained with a backpropagation or related learning algorithm has been frequently used for image classification (Day 1997). Feedforward neural networks are especially attractive for supervised classification applications as, unlike conventional statistical classification algorithms, they do not make a range of assumptions about the datasets used that are often untenable in remote sensing applications. Additionally, on a more pragmatic level, numerous comparative studies have demonstrated that such networks are often, but not always, able to classify remotely sensed imagery more accurately than a variety of other widely used approaches (Peddle et al. 1994, Paola and Schowengerdt 1995, Yang et al. 1999). The active contour models (snakes) technique has received considerable attention (Persaresi et al., 2001). In snakes, a linear feature in the image is modeled by energy which is expressed by geometric and photometric constraints. The extraction of features is performed by optimizing the total energy. Automatic approaches pursue automatic location of a road in the image by recognizing the road and defining its positions accurately. Much effort in the existing methods has been made in automatic determination of the starting point or segment of a road using the knowledge on local

photometric properties of the road point or segment such as the intensity value and contrast (Salembier et al., 1994). Recently, some knowledge-based methods using artificial intelligence techniques have been developed (Shackelford, et al., 2004). These methods use various types of knowledge about roads and the world and inference mechanisms to extract the road network. The road extraction by means of multi-resolution analysis offers a control about the width of roads in the image. Thereby an efficient tool for narrow road detection in high resolution images (big scales) will also serve in the freeways recognition in low resolution images (small scales). In (Heipke et al., 1995) different characteristics of objects such as roads are detected in different scales. In (Baumgartner et al., 1996) it is proposed that the multi-resolution analysis based on the analysis of profiles taken perpendicularly to the axes of roads. Similar study appears by the same author in (Baumgartner et al., 1997). Here, roads are modeled as a network of intersections and links between the intersections. In (Baumgartner et al.,1999) which is based on the topological relations between roads and other cartographic objects. In (Couloigner and Ranchin, 2000) a method to hierarchically extract urban road networks from very high spatial resolution space-borne imagery using the wavelet transform imagery is presented. These methods use various types of knowledge about roads and the world and inference mechanisms to extract the road network. The serious drawback of SVM is that the boundaries of the classes as taken according to radial basis function networks are spherical while the same is not true for majority of the real data. The boundaries of the classes vary in shape, thus leading to poor accuracy. This work is developed on the modified RBFs neural network based classifier for object based classification of high resolution satellite remotely sensed images. The new basis functions, called cloud basis functions (CBFs) use a different feature weighting, derived to emphasize features relevant to class discrimination as discussed in (De Silva et al., 2008). Further, these basis functions are designed to have multiple boundary segments, rather than a single boundary as for RBFs. This new enhancements to the basis functions along with a suitable training algorithm allow the neural network to better learn the specific properties of the problem domain. In this proposed technique, the boundaries of classes considered are not spherical but a set of boundaries is considered for each class, which promises higher accuracy theoretically. Thus it was aimed to propose a suitable classifier for the high resolution satellite remotely sensed images and to test the applicability of cloud basis function neural networks with some modification in the field of remote sensing with especial emphasize on multispectral satellite images. For achieving this objective, an algorithm for object based segmentation is developed for segmentation of high resolution multispectral satellite images, which is an essential step for any classification and then a modified cloud basis function neural network is created, training of the network is done with sample training data and then use the network for classifying high resolution satellite remotely sensed images.

Methodology

Proposed methodology for Object Based Image Segmentation

As the goal is to extract objects, the object based methods are preferred over pixel based methods due to various advantages discussed in (Hay, 2006).Classification of high-resolution satellite images using standard per-pixel approaches is difficult because of the high volume of data, as well as high spatial variability within the objects. One way to deal with this problem

is to reduce the image complexity by dividing it into homogenous segments prior to classification. This has the added advantage that segments can not only be classified on basis of spectral information but on a host of other features such as neighborhood, size, texture and so forth. 90

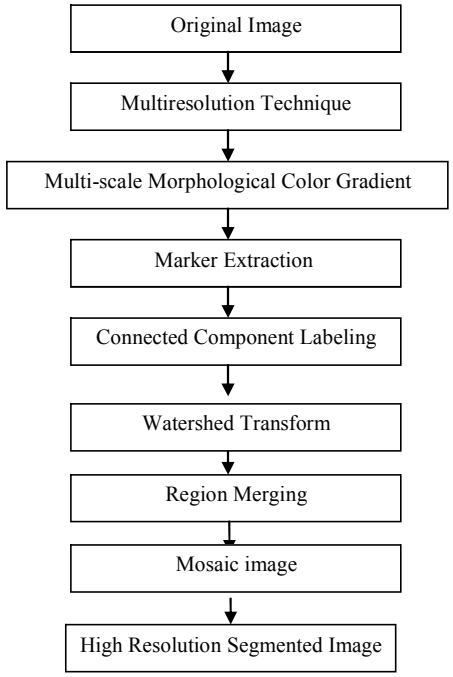

Fig 1. Proposed methodology for image segmentation

Segmentation of the images is carried out using the region based algorithms such as morphological marker based watershed transform by employing the advantages of multi-resolution framework and multi-scale gradient algorithms. The segmentation of the color images is obtained using watershed transform to get its homogenous regions. Classification technique is then applied into these homogenous regions taking the shape, texture and spectral properties of the regions. The proposed methodology for image segmentation is shown in Fig 1.

The Daubechies wavelet is considered because it's an orthogonal transform, compact support and requires small computational complexity. The morphological gradient of each band of the image is calculated using the formula

$$G(f) = (f \oplus B) - (f \Theta B) \tag{1}$$

where $G(f)$: morphological color gradient, f : given image and B be the structuring element. The morphological color gradient of the image is then calculated using:

$$G(f) = \sqrt{\left(G_r(f)^2 + G_g(f)^2 + G_b(f)^2\right)} \tag{2}$$

$G_r(f)$ is gradient of the red band, $G_g(f)$ is gradient of the green band and $G_b(f)$ is gradient of the blue band. The multi-scale morphological color gradient is then calculated using the formula

$$MG(f) = \frac{1}{n}\sum_{i=1}^{n}[G(f)\Theta B_{i-1}] \tag{3}$$

The multi-scale morphological color gradient is dilated with a square structuring element of size 2x2. The constant h is then added to the dilated image; the final gradient is then obtained by reconstruction by erosion of *MG(f)* from .

Hence, the final multi-scale morphological color gradient algorithm is:

$$FG(f) = \Phi^{rec}\left(\left(MG(f) \oplus B\right) + h\right) \quad (4)$$

The markers are the interior of the objects of the interest and the backgrounds are to be extracted from the image to get the segmented image. To many markers results in over-segmentation and too few markers results in under-segmentation. The markers can be extracted from white top-hat or black top-hat transform. But extracted markers from either white or black top-hat will miss some of the objects. So, to utilize the advantage of both top-hat, markers are extracted using morphological laplacian defined as:

$$L(f) = g^{+}(f) - g^{-}(f) \quad (5)$$

where $g+(f)$: White top hat and $g-(f)$: Black top hat transform (Eo et al, 1999)

To utilize the spectral property of the image, markers are extracted from morphological color laplacian of the image; is calculated using the formula:

$$L(f) = \sqrt{\left(L_r(f)^2 + L_g(f)^2 + L_b(f)^2\right)} \quad (6)$$

where $L(f)$ is morphological color gradient, $L_r(f)$ is gradient of the red band, $L_g(f)$ is gradient of the green band and $L_r(f)$ gradient of the blue band.

The marker extracted from image using morphological Laplacian, are to label using connected component labeling. Each connected marker is assigned with a unique label.

Watershed segmentation algorithm applied to the image *C(f)* which is obtained from the marker image *M(f)* and final multi-scale morphological color gradient image *FG(f)*. For any pixel p at position (i,j), *C* is obtained by

$$C(f)_p = FG(f)_p \text{; p is black in } M(f) \quad (7)$$

$$C(f)_p = \frac{1}{2}(FG(f)_p + MG(f)_p) \text{ ; p is white in } M(f) \quad (8)$$

As marker image *M(f)*, provides rough partition of the objects and the final gradient image *FG(f)* avoids over merging. Thus, average of the marker and final gradient multi-scale morphological color images preserves the contours of the object. The algorithm used for the implementation of the watershed transform is ordered queue as it based on maker based watershed transform and easy to implement.

The output of the watershed transform may result in over-segmentation. To merge the adjacent region or the homogenous regions; region merging using criterion is implemented. Each segmented object or region is assigned the average grayscale of each band to generate the mosaic color image. To get the final segmentation at high resolution image; low frequency coefficient of the wavelet is replaced with mosaic image; while detailed coefficients of the wavelet are modified so as to avoid noise introduced back into the finer image. Inverse wavelet transform is then applied on these modified images to get the high resolution segmented image.

Feature Vector Extraction

Physical features in general have certain associations with spectral features, so they can be identified by using multi-spectral information from the remotely sensed images. However land use information cannot be determined by land cover information directly. Properties of objects can be further divided into three categories

- Geometric
- Spectral or thematic
- Textural

A feature vector of all the regions present in the image is calculated. For this work totally 8 features were calculated. The first three values correspond to the values of region's average color in multispectral space. The next three features are related to the shape of the region such as solidity, aspect ratio and eccentricity. The next features correspond to the texture features of each region like contrast ASM etc. Some specific features are consider for extraction of road objects like Density, Width Constancy, Relative border to etc.

Proposed methodology for object based Image Classification

Rather than treating image as set of pixels if we treat it as a set of objects more information can be extracted, as with pixels only intensity values can be used. And with the construction of regions, knowledge is given to the system to classify. This is similar to the way human brain analyzes an image by breaking it down into various objects and uses features such as shape, texture, color and context along with the its cognizance powers to interpret the image. Therefore, dividing the image into regions and then opt for classification is better than per pixel classification.

i. Radial Basis Function Neural Networks:

The general architecture of a radial basis function neural network is as shown in Fig. 1. The network can be seen as a series of two transforms applied on the input data, i.e. a two-tiered mapping. The first mapping is a non-linear mapping from the d-dimensional input space, X, to a (h+1) dimensional basis function space, H, while the second mapping is a linear mapping from the (h+1) dimensional basis space to q-dimensional output space. The first non-linear mapping is usually a Gaussian mapping of the form

$$\varphi(r) = \exp(-\beta r^2) \tag{9}$$

where 'β' is a scale factor, which defines the steepness of the Gaussian as the data points get radially farther from the center of the Gaussian. The second tier of mapping is the set of linear weights that map the post basis function space to the output space. These weights are adjusted during training cycles and after training is completed, these weights are saved and then the network is used for classification. These networks, using Gaussian basis functions try to represent the data distribution spherically, while in real data, it does not hold good. For example, the following figure shows a data distribution similar to the distribution available in remotely sensed images, in 3 classes and the representation of the data using spherical basis functions. Consequently the accuracy of the classifiers using spherical basis functions suffers for such distributions. To overcome this inability to model such data, instead of using a single boundary segment for the basis function, multiple boundary segments are learnt and applied to each basis function based on the neighboring classes.

ii. Cloud Basis Function Neural Network

This form of neural network is a modification of the radial basis function neural network, the algorithm is as follows:

- ***Creating the modified radial basis function neural network***
- Define the input nodes, which take in as input the data from the images
- Define the intermediate nodes for basis function mapping, which map the inputs to the basis space through the Gaussian functions
- Define the output nodes, which form the classes in the image
- ***Programming the training algorithm for the neural network***
- Apply k-means clustering for the initial data to find the possible basis function centers, μ as

$$\mu_j = \frac{1}{N_j} \sum_{x_l \in C_j} x_l \tag{10}$$

- Form the basis function mappings
- Calculate the scale factors, for each of the basis function centers with respect to each of the other basis function centers as

$$\omega_{p,j} = \sqrt{\frac{1}{2} \sum_{i=1}^{d} u_{ii,j} \left(\mu_{ip} - \mu_{ij}\right)^2} \tag{11}$$

And the default scale factor as the mean of all the scale factors as

$$\omega_{0,j} = \frac{1}{k`} \sum_{p=1}^{k`} \omega_{p,j} \tag{12}$$

- Compute the output matrix of the basis function mapping, φ, for the input samples as

$$\phi_j\left(x \mid \mu_j, \{\omega\}_j, U_j\right) = \exp\left(- \frac{\sum_{i=1}^{d} u_{ii,j} \left(x_i - x_j\right)^2}{\left(Sel\left(\{\omega\}_j \mid x\right)\right)^2} \right) \tag{13}$$

- Compute the post basis function weight matrix, W as

$$W = (\varphi)^{\neg} T \tag{14}$$

where $(\varphi)^{\neg}$ represents the pseudo inverse of the output of the basis function matrix

- Compute the output of the network for the input samples and the error in the output with respect to the target vector T as the Euclidean distance from the target vector
- Update the scale factors and the basis function centers based on the error in the output of the network using the supervised iterative gradient descent algorithm as:

$$\{\omega\}_j^m = \{\omega\}_j^{m-1} - \frac{\partial E}{\partial \{\omega\}_j^{m-1}} \tag{15}$$

- After iterative gradient descent is complete for the training iteration, the network output for all the training samples is calculated as

$$Network\_Output = \Phi * W \tag{16}$$

Where, Φ = basis function output matrix and
W = post basis function weight matrix

- According to the network output, classify the pixels and partition the training set into two sets of classified $\{X^C\}$and misclassified samples $\{X^M\}$
- If the number of misclassified samples is less than a set threshold, or if the number of misclassified samples doesn't change in successive cycles, stop training
- For all the classes for which the number of misclassified samples is greater than the set threshold, add a basis function to improve the representation of the class
- Repeat the training algorithm till a maximum number of epochs are completed or till the number of misclassified samples do not change with the increasing basis functions

Classifying the test images using the network

- Input the test images for classification
- Obtain the output matrix for the classification details of the image
- Calculate the classification accuracy of the network

These three phases outline the project and the encompassed objectives within them are to be achieved to move on to the next phase of the project.

Study Area

The study areas considered in the present study is a part Mumbai City image which is multispectral QuickBird image data of 2.44 metre spatial resolution, sharpened by 0.61 metre panchromatic data with an area of 500x500 pixels shown in Fig 2.

Fig 2. Mumbai City Image

Result and Discussions

Object based image segmentation of Mumbai City image is shown in Fig 3. Various features (as stated above) are considered before object based image classification using cloud basis functions neural network is shown in the Fig 4. Five different classes are considered for

supervised classification. Eight samples from each class is taken as training data. Misclassifications because of shadows are

Fig 3 Object based image segmentation of Mumbai Image

common problem; specially road areas. Producer's , user's and overall accuracy was found 92%, 94% and 92% respectively.

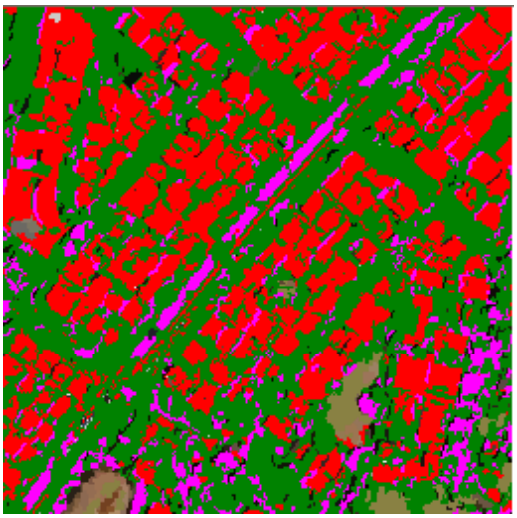

Fig 4 Object based classification of Mumbai image

Vegetation Build up Open Area

Road Shadow

The extracted road from Mumbai City image using proposed technique are shown with pink color in the Fig 5 below. It is seen from the output that proposed technique gives promising

result as compared to other techniques discussed in literatures. There is some misclassification between vegetation and road due to same spectral properties.

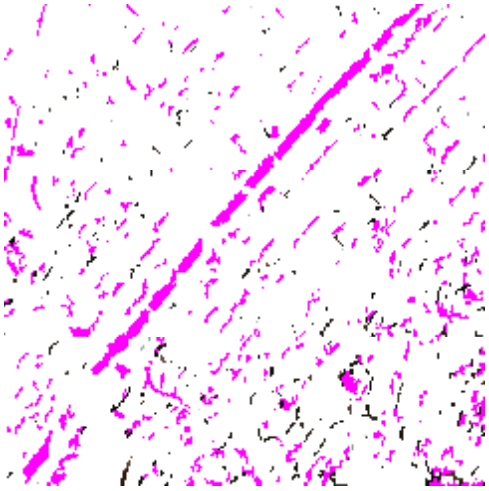

Fig 5. Extracted Roads from Mumbai image

Some unclassified black objects are also seen in the output image, which can be avoided by considering proper classes as well as training sets. Both these proposed algorithms i.e. object based image segmentation and object based image classification are implemented in C++ language, using Code::Blocks IDE on Windows XP platform, and has been successfully tested with various multispectral images.

Summary and Conclusions

In the present work Cloud Basis Functions Neural Network (CBFs NN) classification method was used for extraction of road from high resolution satellite remotely sensed images. Object based image analysis is heavily dependent on the quality and resolution of the image data. From the results presented in this paper it can be concluded that the object based approach enables the usage of various features, making full use of high resolution images information. Beyond purely spectral information, image objects contain a lot of additional attributes which can be used for classification and this method is more suitable and will be the trend for the high resolution remotely sensed data. Object-based approach has the advantage to produce compact objects which correspond to human eye perception of the environment and it reduces the variance problem of very high resolution satellite data. It also provides possibilities to bring in additional knowledge on the image objects of interest, on object inter-relations and relations to external map or GIS information.

REFERENCES

Baumgartner, A., Steger, C., Wiedemann, C., Mayer, H.,Eckstein, W., Ebner, H.,1996. Update of roads in GIS from aerial imagery: Verification and multi-

resolution extraction. Internat. Arch. Photogrammet. Remote Sensing. 1 (Part B3), 53–58.

Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., 1997. Multi-resolution, semantic objects and context for road extraction. In: Forstner, W., Plumer, L. (Eds.), Semantic Modeling for the Acquisition of Topographic Information from Images and Maps. Birkhauser-Verlag, Basel, pp. 150–156.

Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., Ebner,H., 1999. Automatic road extraction in rural areas. Internat. Arch. Photogrammet. Remote Sensing, 107–112.

Benediktsson, J. A., Pesaresi, M., and Arnason, K. 2003. Classification and feature extraction from remote sensing images from urban areas based on morphological transformations. IEEE Transactions on Geoscience and Remote Sensing, 41(9):1940–1949.

Carpenter, G. A., Gopal, S., Macomber, S., Martens, S., Woodcock, C. E., and Franklin, J., 1999, A neural network method for efficient vegetation mapping Remote Sensing of Environment, 70, 326–338.

Couloigner, I., Ranchin, T., 2000. Mapping of urban areas: A multi-resolution modeling approach for semiautomatic extraction of streets. Photogrammet. Eng. Remote Sensing 66 (7), pp. 867–874.

De Silva C.R., Ranganath S., and De Silva L.C., "Cloud Basis Function Neural Network: A modified RBF network architecture for holistic for holistic Facial Expression Recognition", Elsevier Pattern Recognition 41 (2008) 1241-1253, 2008

Duda R.O, Hart P.E., Stork D.G., 2001, Pattern Classification. 2nd edition, John Wiley & Sons.Eo J. and Kim H, 1997, A Detail Extraction Technique For Image Coding Using morphological Laplacian Operator, IEEE TENCON, Dankook Univerity, Hannam-Dong, Yongsan-Ku, 140-714 Seoul, Korea.Heipke, C., Steger, C., Multhammer, R., 1995. A hierarchical approach to automatic road extraction from aerial imagery. In: McKeown, Jr. D.M., Dowman, I.J. (Eds)., Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision II, Proc. SPIE, Vol. 2486, pp. 222–231.

Ito, Y., and Omatu, S., 1997, Category classification method using a self-organising neural network. International Journal of Remote Sensing, 18, 829–845.

Ito, Y., and Omatu, S., 1998, Polarimetric SAR data classification using competitive neural networks. International Journal of Remote Sensing, 19, 2665–2684.

Paola, J. D., and Schowengerdt, R. A., 1995, A detailed comparison of backpropagation neural network and maximum likelihood classification for urban land use classification. IEEE Transactions on Geoscience and Remote Sensing, 33, 981–996.

Peddle, D. R., Foody, G. M., Zhang, A., Frankilin, S. E., And Ledrew, E. F., 1994, Multisource image classification II: an empirical comparison of evidential reasoning and neural network approaches. Canadian Journal of Remote Sensing, 20, 396–407.

Persaresi, M. and Benediktsson, J. A. 2001. A new approach for the morphological segmentation of high resolution satellite imagery. IEEE Transactions on Geoscience and Remote Sensing, 39(2):309–320.

Shackelford, A. K., Davis, C. H., and Wang, X, 2004. Automated 2-D building footprint extraction from high-resolution satellite multispectral imagery, Proceedings of International Geoscience and Remote Sensing Symposium Vol. 3, pp1996–1999.

Salembier, P. and Pardas, M. 1994. Hierarchical morphological segmentation for image sequence coding. IEEE Transactions on Image Processing, 3(5):639- 651.

Tso, B., and Mather, P. M., 2001, Classification methods for remotely rensed data (London: Taylor and Francis).

Yang, H., Van Der Meer, F., Bakker, W., and Tan, Z. J., 1999, A back- propagation neural network for mineralogical mapping from AVIRIS data. International Journal of Remote Sensing, 20, 97–110.

Imdad A. Rizvi received B.E. in 2000 and M.E in 2004 in Electronics Engineering from University of Mumbai, India. Since January 2008 he has been pursuing Ph.D from Centre of Studies in Resource Engineering, Indian Institute of Technology Bombay, Mumbai, India. His area of research is Supervised image classification, Object based Image analysis and multiresolution algorithms.

B.Krishna Mohan received M.Tech in 1983 and Ph.D in 1991 in Electrical Engineering from Indian Institute of Technology Bombay, Mumbai, India. Currently he is an Associate Professor in Centre of Studies in Resource Engineering, Indian Institute of Technology, Bombay, Mumbai, India. His research interests include application of image processing and pattern recognition techniques to various remote sensing problems, geographic information systems and satellite image compression.