Islamic Azad University
Mashhad Branch



# Object-Oriented Method for Automatic Extraction of Road from High Resolution Satellite Images

**I. Ali Rizvi*[1] and B. Krishna Mohan[2]**

*1. Satellite Image Processing Research Laboratory, Centre of Studies in Resource Engineering,*
*2. Indian Institute of Technology Bombay, Powai, Mumbai 400076, India*
Received 15 January 2010; accepted 25 February 2010

## 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 two fundamental and critical steps towards content analysis and image understanding i.e. image segmentation and classification. This paper proposes a robust object based segmentation algorithm using multi-resolution 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-Oriented Method, Image segmentation, Satellite Images*

## 1. 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. But the presence of high resolution satellite images and their potential to be used in wide variety of applications such as preparing and updating maps have made the automatic extraction of object, especially roads and buildings, a new challenge in remote sensing. Traditionally, road extraction from satellite images has been performed manually by the operator, hence costly, time consuming and the efficiency was by no means very high.
Automatic road extraction provides means for creation, maintaining, and updating transportation network. It also provides data bases for traffic management, automated vehicle navigation and guidance. So the object-based image approach is employed for road extraction from high resolution satellite remotely sensed image. It reduces the spectral variations during image segmentation. Some specific feature vectors are considered for extraction of road objects.

---------------------

*Corresponding author.
*E-mail address (es): imdadrizvi@iitb.ac.in*

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. In most of the commonly used classifiers, radial basis functions are used 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. This new enhancement 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 highlighted in section 2. The proposed methodologies adopted in this paper are discussed in section 3. Section 4 presents experimental results and discussions. Conclusions drawn from this work are discussed in section 5.

## 2. 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 [1]. For an overview a few of recent works are focused here.

An approach designed for a wide variety of imagery based on an object-oriented database which allows the modeling and utilization of relations between roads as well as other objects [2]. Road extraction using statistical modeling in the form of point processes and Reversible Jump Markov Chain Monte Carlo are discussed in [3]. A system for road extraction from multi-spectral imagery based on fuzzy logic is proposed by Amini [4]. Segmentation of color images using the Dempster-Shafer theory of evidence for the fusion of texture, to extract linear features are presented in [5]. A multi-resolution analysis approach based on wavelets, road junction detection, and grouping is proposed in [6]. An approach based on fuzzy logic and mathematical morphology can be seen in [7]. The effectiveness of angular texture signature to discriminate among parking lots and roads using high resolution satellite images is discussed in [8]. In this research, spectral and textural information were used separately for detection of roads and for eliminating non-road pixels respectively. Knowledge-based methods using artificial intelligence techniques have been developed in [9]. 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). A method to hierarchically extract urban road networks from very high spatial resolution space-borne imagery using the wavelet transform imagery is presented in [10]. These methods use various types of knowledge about roads and the world and inference mechanisms to extract the road network.

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: pre-processing, feature extraction and classification [11]. 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 [12]. The multi-layer perceptron (MLP) network trained with a back-propagation or related learning algorithm has been frequently used for image classification. 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 [13]. 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. This technique is emphasized specifically for multispectral satellite images. Thus it was aimed to propose a suitable classifier for the high resolution satellite remotely sensed images and to test the applicability of modified cloud basis functions for the field of remote sensing. For achieving this objective, an algorithm for object based segmentation is developed for segmentation of high resolution multispectral satellite images and then a modified cloud basis function neural network is created. This network is trained using the sample data and finally the network is used for the classification.

## 3. Methodology

In this paper we propose an approach for automatic road extraction from high resolution remotely sensed multi-spectral images, such as IKONOS or QuickBird. While aerial imagery usually consists of three spectral bands, high resolution satellite data comprises four spectral bands with a better radiometric quality compared to film, but a worse geometric resolution. Therefore, strongly making use of the spectral properties of satellite imagery is a way to mitigate the geometric disadvantages and achieve results comparable to those from aerial imagery.

### 3.1 Proposed methodology for Object Based Image Segmentation

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. The proposed methodology for image segmentation is shown in Fig. 1.

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 algorithm is given below:

- Apply multi-resolution framework (here Daubech6 family of wavelet transform is used) to input image.
- Use multi-scale gradient algorithms to calculate color gradient.

The morphological gradient of each band of the image is calculated using equation (1)

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

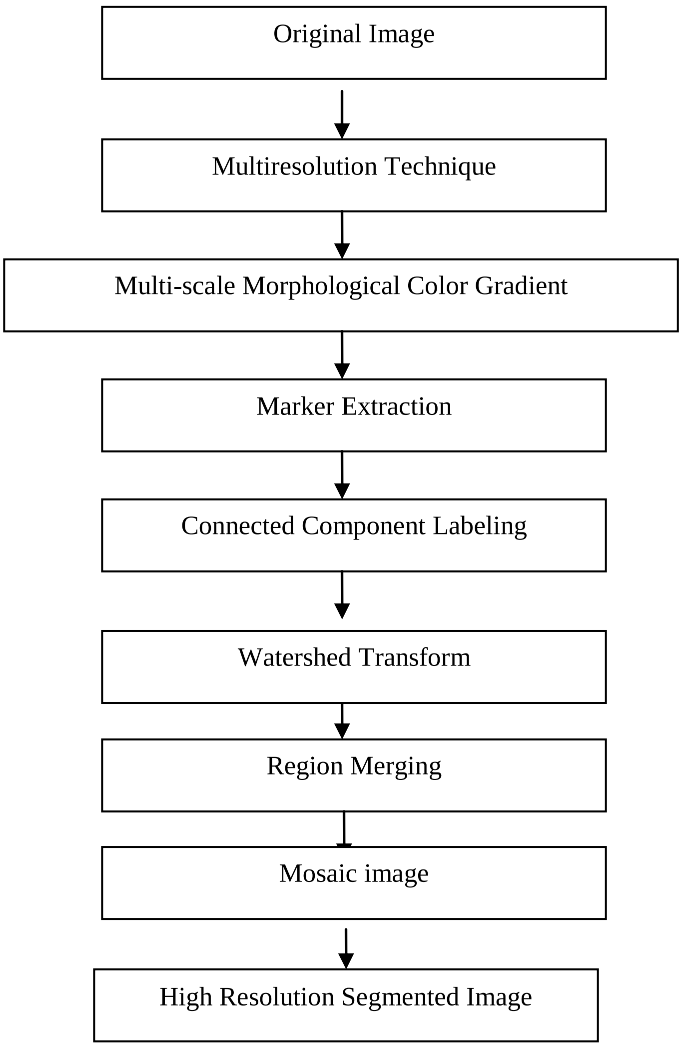


Fig 1. Proposed methodologys for image segmentation

where $G(f)$ = Morphological color gradient,
$f$ = Given image
$B$ = Structuring element.

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

$G_r(f)$ = Gradient of the red band,
$G_g(f)$ = Gradient of the green band and
$G_b(f)$ = Gradient of the blue band.

The multi-scale morphological color gradient is dilated with a square structuring element of size 2x2.

- 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 [14], which can be defined as:

$$L(f) = g^{+}(f) - g^{-}(f) \qquad (3)$$

where $g+(f)$ = White top hat transform and
$g-(f)$ = Black top hat transform

For utilizing the spectral property of the image, markers are extracted from morphological color Laplacian of the image; and is calculated using equation (4)

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

(4) where $L(f)$ = Morphological color gradient,
$L_r(f)$ = Gradient of the red band,
$L_g(f)$ = Gradient of the green band and
$L_r(f)$ = Gradient of the blue band.

- Apply connected component labeling to connect various labels.
- Morphological marker based watershed transform algorithm is used for region segmentation [15].
- Region merging is done to avoid over-segmentation.
- Mosaic image is generated.
- Inverse wavelet transform is used to generate high resolution image.

The output of the watershed transform may result in over-segmentation. To merge the adjacent region or the homogenous regions; region merging using criterion (What 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.

### 3.2 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 multi-spectral 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 considered for extraction of road objects like Density, Width Constancy and Relative border to etc. [16].

### 3.3 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. Hence cloud basis function neural network is used which is essentially a form of neural network with modification in 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 initial data to find the possible basis function centers, μ.
- 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.

$$\omega_{p,j} = \sqrt{\frac{1}{2}\sum_{i=1}^{d}\left(\mu_{ip} - \mu_{ij}\right)^2} \qquad (5)$$

$\omega_{p,j}$ = Scale factor of the boundary segment between $p^{th}$ mean and $j^{th}$ mean

$\left(\mu_{ip} - \mu_{ij}\right)$ = Euclidian distance between $p^{th}$ mean and $j^{th}$ mean

*d* = Number of features in each object

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

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

(6)

$\omega_{0,j}$ = Default scale factor for $j^{th}$ mean

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

$$\phi_j\left(x \mid \mu_j, \{\omega\}_j, U_j\right) = \exp\left(-\frac{\sum_{i=1}^{d} u_{ii,j}\left(x_i - x_j\right)^2}{\left(Sel\left(\{\omega\}_j \mid x\right)\right)^2}\right) \qquad (7)$$

where ,

$\varphi_j\left(x \mid \mu_j, \{\omega\}_j\right)$ = Basis function output for sample *x* belonging $j^{th}$ to cluster

$Sel\left(\{\omega\}_j \mid x\right)$ = Boundary segment selected for sample *x*

$\left(x_i - u_{ij}\right)^2$ = Euclidian distance between sample *x* and $j^{th}$ mean

- Compute the post basis function weight matrix, W.

$$W = (\varphi)^{\neg} T \qquad (8)$$

where

$(\varphi)^{\neg}$ = Pseudo inverse of the output of the basis function matrix

*T* = Target Vector

- 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.

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

(9)

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

$$Network\_Output = \Phi * W$$

(10)

Where, $\Phi$ = 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.

### 3.4 Study Area

The study areas considered in the present study are subsets of Mumbai and Rome City images which are multi-spectral QuickBird image data of 2.44 meter spatial resolution, sharpened by 0.61 meter panchromatic data with an area of 500x500 and 228x434 pixels as shown in Fig. 2 and Fig. 3 respectively.

## 4. Result and Discussions

Object based image segmentation of Mumbai and Rome City images are shown in Fig.4 and Fig.5 respectively. Various features vectors are considered before object based image classification using cloud basis functions neural network, as shown in the Fig.6 and Fig.7. Five different classes are considered for supervised classification for each image. Eight samples from each class are taken as training data. Misclassification in case of shadow regions are observed. Producer's, user's and overall accuracy were found as 92%, 94% and 92% respectively for Mumbai city image and 88%, 90% and 93% respectively for Rome city image. The extracted road from Mumbai City image using proposed technique are shown with pink color in the Fig.8 and in grey color in Fig.9 shows the extracted roads from Rome image. There is some misclassification between vegetation and road due to same spectral properties. Some unclassified low reflectance 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 multi-spectral images.

Fig. 2. Mumbai City Image

Fig. 3. Rome City Image

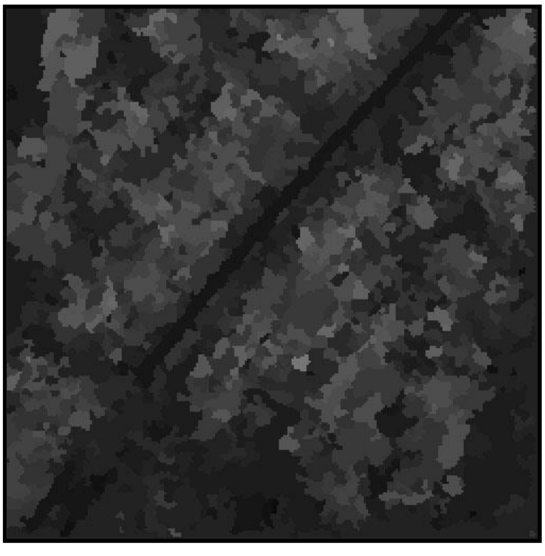

Fig. 4. Object based image segmentation *of Mumbai image*

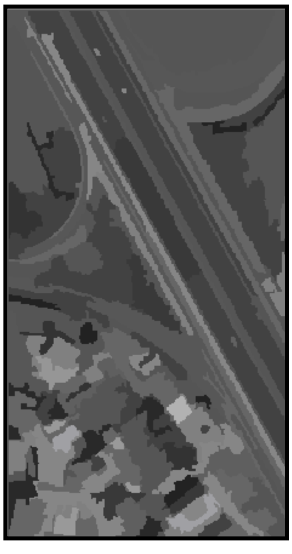

Fig. 5. Object based image segmentation *of Rome image*

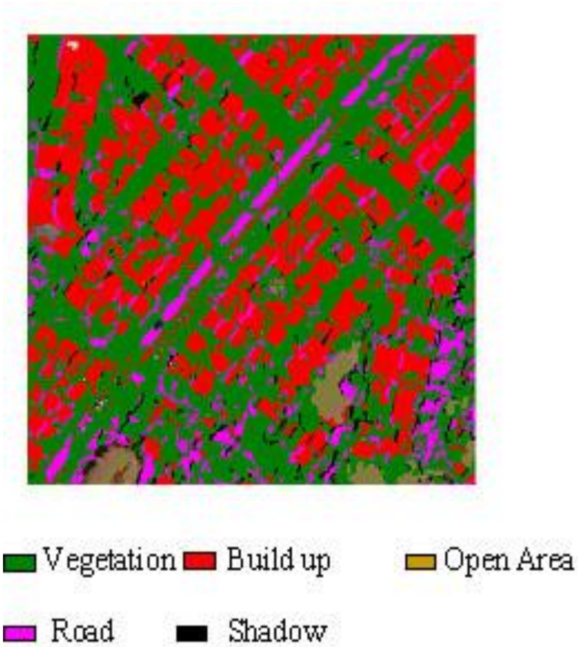


Fig. 6. Object based classification of Mumbai image

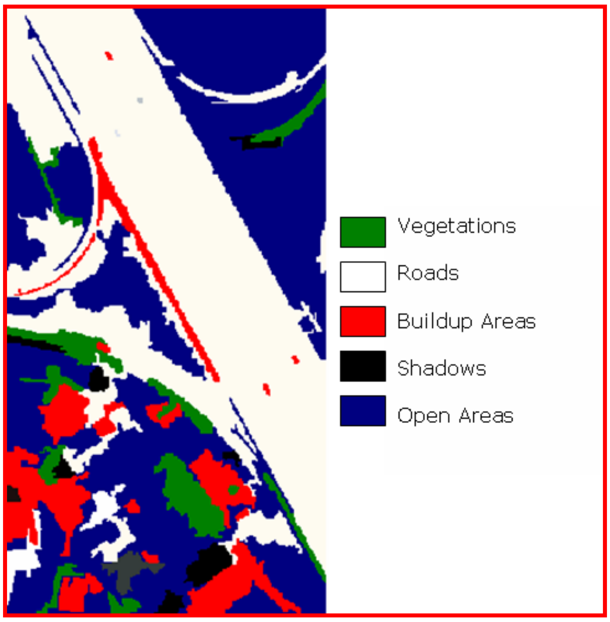


Fig. 7. Object based classification of Rome image

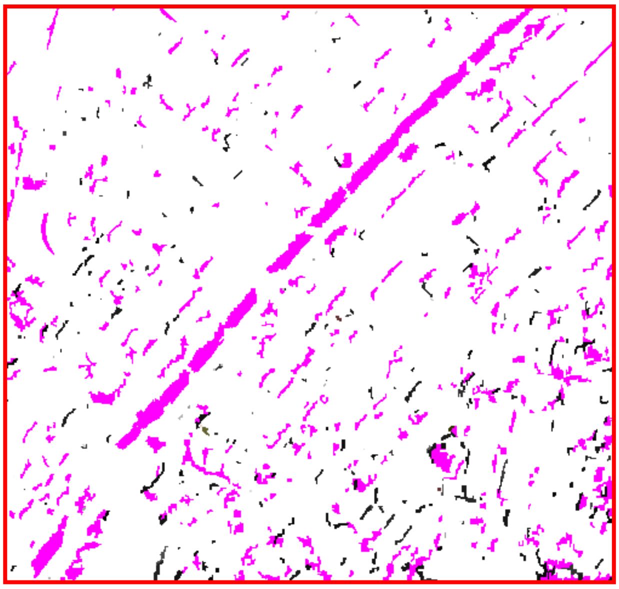

Fig. 8. Extracted Roads from Mumbai image

Fig. 9. Extracted Roads from Rome image

## 5. 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.

### Acknowledgements

Authors are grateful to Digital Globe Inc., USA for granting permission to use the browse images on their website for this research work.

## References

[1] 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, v. 41, no.9, p.1940–1949.

[2] Wallace, S., Hatcher, M., Priestnall, G. and Morton, R., 2001. Research into a Framework for Automatic Linear Feature Identification and Extraction. In: Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), Balkema Publishers, Lisse, Niederlande, p. 381–390.

[3] Stoica, R., Descombes, X. and Zerubia, J., 2004. A Gibbs Point Process for Road Extraction from Remotely Sensed Images. International Journal of Computer Vision v. 57, no.2, p. 121–136.

[4] Amini, J., Lucas, C., Saradjian, M., Azizi, A. and Sadeghian, S., 2002. Fuzzy Logic System for Road Identification Using IKONOS Images. Photogrammetric Record v. 17, no.99, p. 493–503.

[5] Mena, J. and Malpica, J., 2003. Color Image Segmentation Using the Dempster-Shafer Theory of Evidence for the Fusion of Texture, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, v.34, 3/W8, p. 139–144.

[6] Zhang, Q. and Couloigner, I., 2004. A Wavelet Approach to Road Extraction from High Spatial Resolution Remotely Sensed Imagery, Geomatica v.58, no. 1, p. 33–39.

[7] Mohammadzadeh, A., Tavakoli, A. and Zoej, M., 2004. Automatic Linear Feature Extraction of Iranian Roads from High Resolution Multi-Spectral Satellite Imagery, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 35, B3, p. 764–767.

[8] Zhang, Q., and Couloigner, I., 2006. Benefit of the angular texture signature for the separation of parking lots and roads on high resolution multi-spectral imagery. Pattern Recognition Letters, v.27, no.9, p. 937-946.

[9] Shackelford, A. K., Davis, C. H., and Wang, X, 2004. Automated 2-D building footprint extraction from high-resolution satellite multispectral imagery, International Geoscience and Remote Sensing Symposium v. 3, p.1996–1999.

[10] Couloigner, I., Ranchin, T., 2000, Mapping of urban areas: A multi-resolution modeling approach for semiautomatic extraction of streets, Photogrammetry Engineering and Remote Sensing, v. 66, no.7, p. 867–874.

[11] Duda R.O, Hart P.E., Stork D.G., 2001, Pattern Classification 2nd edition, John Wiley & Sons.

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

[13] De Silva C.R., Ranganath S., and De Silva L.C., 2008, “Cloud Basis Function Neural Network: A modified RBF network architecture for holistic for holistic Facial Expression Recognition”, Elsevier Pattern Recognition 41, pp-1241-1253.

[14] Eo, J., and Kim, H., 1997, A Detail Extraction Technique For Image Coding Using morphological Laplacian Operator, IEEE TENCON, Korea, p.140-147.

[15] Meyer, F., and Beucher, S., 1990, “Morphological Segmentation,” Journal of Visual Communication and Image Representation, v. 11, p. 21–46.

[16] Costa L and Cesar R., 2009, Shape Classification and Analysis, 2nd Edition, CRC Press, p.259-310.