Change Detection of Water-Body in Synthetic Aperture Radar Images

Sree Sharmila T sreesharmilat@ssn.edu.in
Assistant Professor
SSN College of Engineering
Chennai, India

Ramar K kramar_nec@rediffmail.com
Einstein College of Engineering
Tirunelveli, India

Vidhusha S vidhushas@ssn.edu.in
Assistant Professor
SSN College of Engineering
Chennai, India

Abstract

Change detection is the art of quantifying the changes in Synthetic Aperture Radar (SAR) images occurring over a period of time. Remote sensing has been instrumental in performing change detection analysis. The impact of applying the combination of texture features for classification techniques to separate water bodies from land masses is empirically investigated in this paper. First, the images are classified using unsupervised Principle Component Analysis (PCA) based K-means clustering for dimension reduction. Then, texture features like Energy, Entropy, Contrast, Inverse Difference Moment, Directional Moment and Median are extracted using Gray Level Co-occurrence Matrix (GLCM) and these features are utilized in Learning Vector Quantization (LVQ) and Support Vector Machine (SVM) classifiers. This paper aims to apply a combination of the texture features in order to significantly improve the accuracy of detection. The utility of detection analysis influences management and policy decision making for long-term construction projects by predicting preventable losses.

Keywords: Change Detection, Classification, Learning Vector Quantization, Support Vector Machine, Texture Features

1. INTRODUCTION

Remote sensing and related techniques such as geographic information systems have a persistent impact on the conduct of realistic work. With satellite instruments, it is possible to observe a target repeatedly, thereby contributing effectively to perform change detection in areas of interest.

This work mainly focuses on change detection, which happens because of many possible environmental and human actions. SAR images enable direct observation of the land surface at repeated intervals, allowing mapping, monitoring and assessment. Change detection analysis is effective in long-term planning. PCA based K-means clustering is a statistical technique widely used for dimensional reduction. Usually the clustering methods are developed for different purposes. Clustering algorithms used for unsupervised classification of remote sensing data vary according to the efficiency with which clustering takes place. The unsupervised clustering provides the cluster information about the water body in a relatively quick manner [5]. This lacks complete information about the region of interest and particularly subtle variations therein. To avoid unexpected groupings, supervised classification is recommended.

The mapping of classes is much more accurate in supervised classification but is heavily dependent on the input given. The researchers compared and analyzed SVD (Singular Value Decomposition) and Gray Level Co-occurrence Matrix (GLCM) as two methods for feature extraction [4, 2]. The comparison result shows that GLCM gives a better result than SVD. Hence, the texture features from the training data sets were extracted by using GLCM and then subjected to the LVQ.

The classes of interest that is the coastal and non-coastal areas and the learning rate are determined to optimize the classification accuracy of the SAR images [1]. This paper aims to further improve the accuracy of classification by combining the six texture features instead of using a single or different combination of texture features. Like PCA, LVQ is also used in feature vector dimension reduction. Here is the chance to optimize training data through reducing the number of samples for this analysis. After achieving the desired accuracy by LVQ, the features were subjected onto SVM. In SVM, the two classes were identified and differentiated which helped in finding out the classes in target image. Once the images were classified using supervised classification the change map is constructed to depict the changes [8,13,14].

2. METHODOLOGIES USED FOR CHANGE DETECTION

Figure 1 and Figure 2 depict the major steps in the nomenclature of water bodies and land masses region using unsupervised and supervised classification.

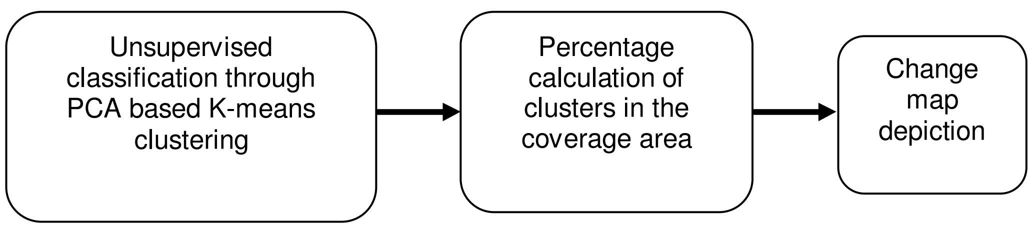


FIGURE 1: Modules of Clustering

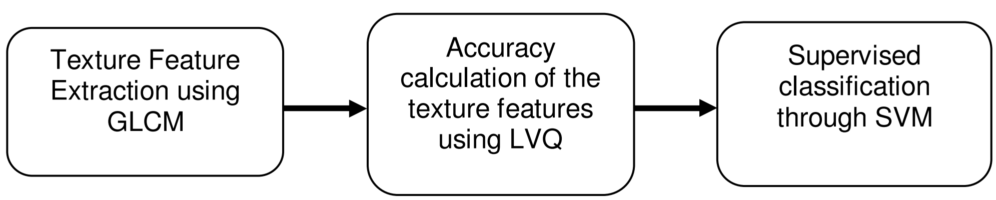


FIGURE 2: Feature Extraction and Supervised Classification

2.1 PCA based K-Means Clustering

The most important aspect of image classification is finding groups in data [4, 5]. Pixels with similar intensities forms clusters in images. Gray values are specifically used for this purpose. The K-means algorithm has some further refinements for its applications on change detection, by splitting and merging of clusters. This is illustrated in Figure 1. First, the given image values are preprocessed by using PCA. Using the most important components of PCA, the image information is mapped into the new feature space. Then, the K-means algorithm is applied to the data in the feature space. The final objective is to distinguish the different clusters using eigen values. Clusters are grouped if the cluster standard deviation exceeds a threshold and the number of pixels is twice for the minimum number of pixels. The main intention of K-means algorithm is to reduce variability within clusters [5]. The objective function is the sum of square distances between cluster centre and its assigned pixel value.

$$F = \sum_{i=0}^{n} \left[x_i - C\left(x_i\right)\right]^2 \tag{1}$$

where, x_i is the pixel value assigned to mean value of the cluster $C(x_i)$. Next the Mean Squared Error (MSE) is determined; minimizing the error is equivalent to minimizing the sum of squared distances.

$$\text{Error} = \frac{\sum_{i=0}^{n} \left[x_i - C\left(x_i\right)\right]^2}{(N - C)} \tag{2}$$

where, N indicates how many number of pixels, C specifies the expected number of clusters. K-means is very responsive to initial values. It is often not obvious that the clustering with the lesser MSE is truthfully the better classification. Thus, the results obtained have aided in attaining this objective and thereby PCA based K-means classification has been efficiently applied for the change detection analysis.

The unsupervised classification provides the cluster information about the water body in a relatively quicker manner [12]. This clustering lacks complete information about the region of interest and particularly subtle variations therein. Supervised classification gives the better results compared to the unsupervised classification.

2.2 Texture Feature Extraction

Texture portrays a rich source of data about the natural landscape. The ways of extracting the texture features have been performed through GLCM. It is a perfect and an efficient tool to perform the extraction of texture features [4, 6]. The basis of GLCM is assigning the relationship between two neighbouring pixels in one offset as the second order texture. The gray value relations in a target image are transformed into the co-occurrence matrix by a given kernel mask 3×3. In the transformation from the image onto the co-occurrence matrix, the neighbouring pixels in 0° direction can be used. It contains information of the position of pixels having similar gray level values.

The texture features extracted for classification are Energy (E), Entropy (Ent), Contrast (Con), Inverse Difference Moment (IDM), Directional Moment (DM) and the Median (M). Energy can be defined as the measure of the extent of pixel pair repetitions. Entropy is the measure of randomness that is used to characterize the texture of the input image. Its value will be maximum when all the elements of the co-occurrence matrix are the same. The contrast is a measure of intensity of a pixel and its neighbour over the image. Contrast is 0 for a constant image. Inverse Difference Moment is a measure of image texture. IDM has a range of values so as to determine whether the image is textured or non-textured. Directional moment is a textural property of the image computed by considering the alignment of the image as a measure in terms of the angle [15-17]. There are four orientations namely 0°, 45°, 90° and 180°. The following defines the texture features subjected for classification.

$$E = \sqrt{\sum_{i=0}^{M-1} \sum_{j=o}^{M-1} M^2(i, j)} \tag{3}$$

$$Ent = \sum_{i=0}^{M-1} \sum_{j=o}^{M-1} M(i, j)(-\ln(M(i, j))) \tag{4}$$

$$Con = \sum_{i=0}^{M-1} \sum_{j=o}^{M-1} (i-j)^2 M(i, j) \tag{5}$$

$$IDM = \sum_{i=0}^{M-1} \sum_{j=o}^{M-1} \frac{1}{1+(i-j)^2} M(i, j) \tag{6}$$

$$DM = \sum_{i=0}^{M-1} \sum_{j=o}^{M-1} M(i, j) \left| i-j \right| \tag{7}$$

$$M = \sum_{i=0}^{M-1} \sum_{j=o}^{M-1} \frac{M(i, j)+1}{2} \tag{8}$$

where i and j are the coefficients of co-occurrence matrix, $M(i, j)$ is the element in the co-occurrence matrix at the coordinates i and j, M is dimension of the co-occurrence matrix.

2.3 Learning Vector Quantization

LVQ is a variation of the well-known Self Organizing Map (SOM) architecture. This has a feed-forward network structure with a single layer of neurons arranged in rows and columns. Each neuron is fully connected to input layer source units. The different stages of LVQ are:

Step 1: Initialization – Assign the initial weight vector w_j by selecting the random values.
Step 2: Sampling – Describe a sample input vector x from the input space. Here the network chooses five texture features as input to classify the two classes.
Step 3: Matching – Find weight vector closest to the input vector named as winning neuron I(x), which is defined as,

$$\sum_{i=1}^{n} (x_i - w_{ji})^2 \tag{9}$$

Step 4: Updating – Apply the following weight update function to update the weights

$$\Delta w_{ji} = \eta \, T_{j,I(x)} \left(x_i - w_{ji} \right) \tag{10}$$

where $T_j, I_{(x)}$ is a Gaussian neighbourhood and η is the learning rate.

Step 5: Continuation – Repeat the steps 2 to 5 to reach the best accuracy of input space.

2.4 Support Vector Machine

A support vector machine is a set of related supervised learning algorithms that analyze the inputs and recognize the classes, used for classification and regression. SVM takes a set of input data; for each given input, it determines whether the input is a member of which of two classes. This makes SVM as a linear classifier [7]. The training algorithm of SVM constructs a model that assigns new data into one category or the other of two classes. Figure 3 depicts basic building blocks of SVM as follows,

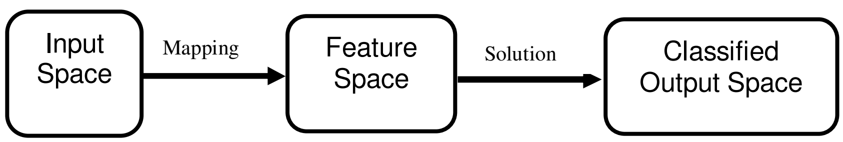


FIGURE 3: The working of SVM algorithm

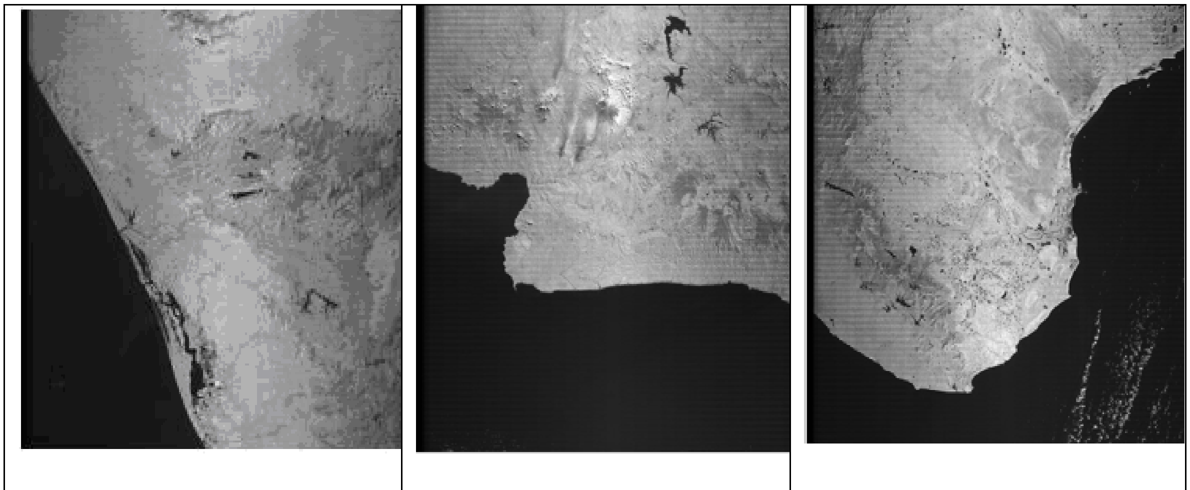
The goal is to separate the two classes without loss of generality by a function which is induced from available training data sets. The task is to produce a classifier that will work in a generalized manner [3, 9]. The experiment for handling the data set is tested on a change detection problem with a large set of data points [11]. The application of SVM for the desired problem is minimizing the error through maximizing the margin which means that it maximizes the distance between it and the nearest data point of each class [10]. Since SVM are known to generalize well even in high dimensional spaces under small training sample. This linear classifier is termed as the optimal separating hyper plane. Cover's theorem states that if the transformation is non-linear and the dimensionality of the feature space may be transformed into a new feature space is high enough, then the input space may be transformed into a new feature space where the patterns are linearly separable with high probability. This non-linear transformation is performed in an implicit way is called kernel function [9,10].

SVM is known to generalize well even in high dimensional spaces under small training sample conditions and have shown to be superior to traditional neural networks. The experiment for handling the data set is tested on a change detection problem with a large set of data points. The support vector machine classifier uses large set of training and testing data for classification of coastal areas.

3. RESULTS AND DISCUSSIONS

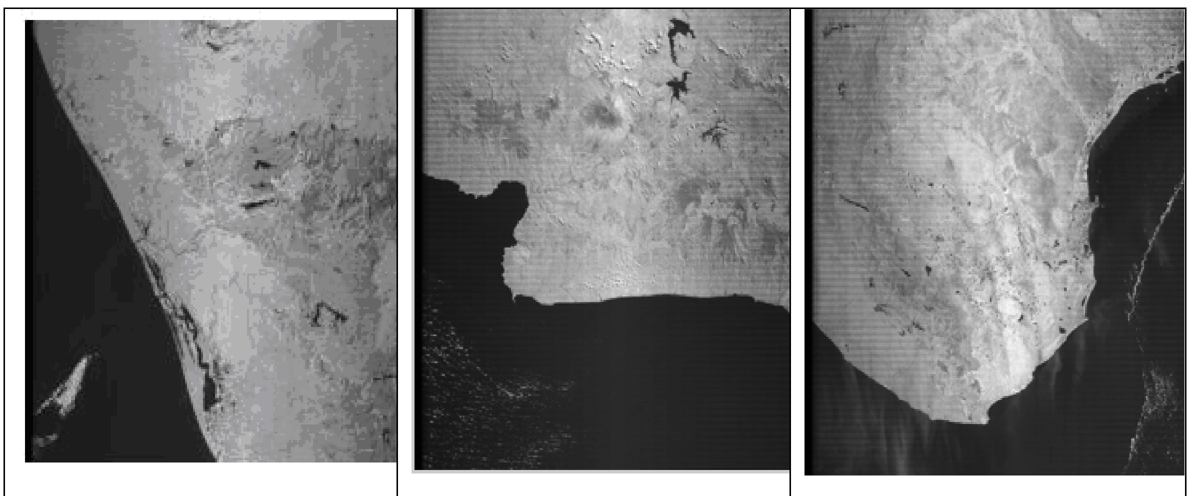
3.1 Data Sets

This paper deals with the Land Remote-Sensing Satellite (LANDSAT) images taken from different time frames of the coastal regions of the world. Some of the sample input study area imageries are shown in Figure 4 and 5.



Kochi Region	Indonesia Region	Kanyakumari Region

FIGURE 4: Input Landsat images of coastal landscape taken during 2005



Kochi Region	Indonesia Region	Kanyakumari Region

FIGURE 5: Input Landsat images of coastal landscape taken during 2010

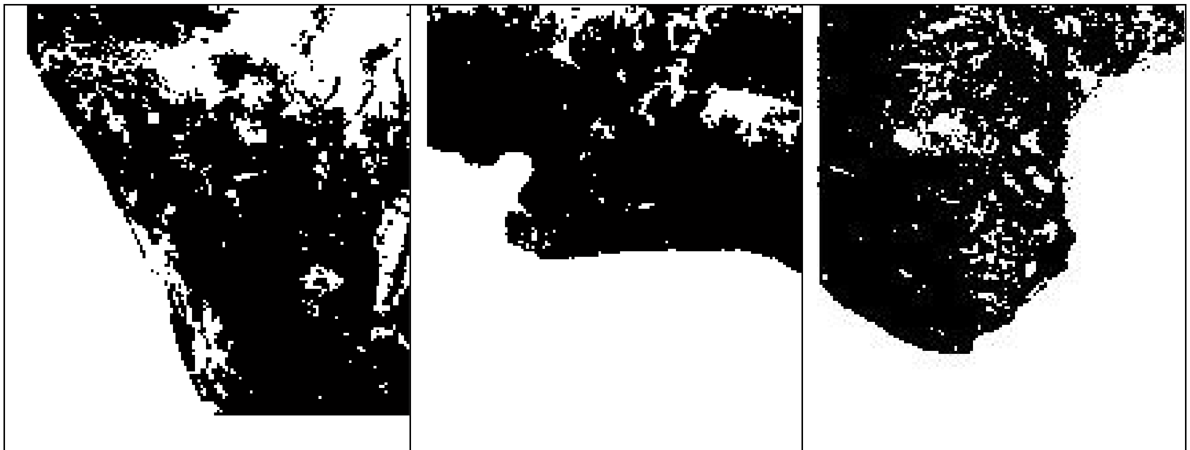
3.2 Results of Classification Through PCA Based K-means Clustering

The input SAR images have been subjected to unsupervised classification performed using a PCA based K-means classifier. Figure 6 and Figure 7 depict the result of classification which facilitates differentiation of water bodies from land masses. These results have been further utilized in calculating the percentage of coverage area to detect changes over a period of time.

3.3 Results of Supervised Classification

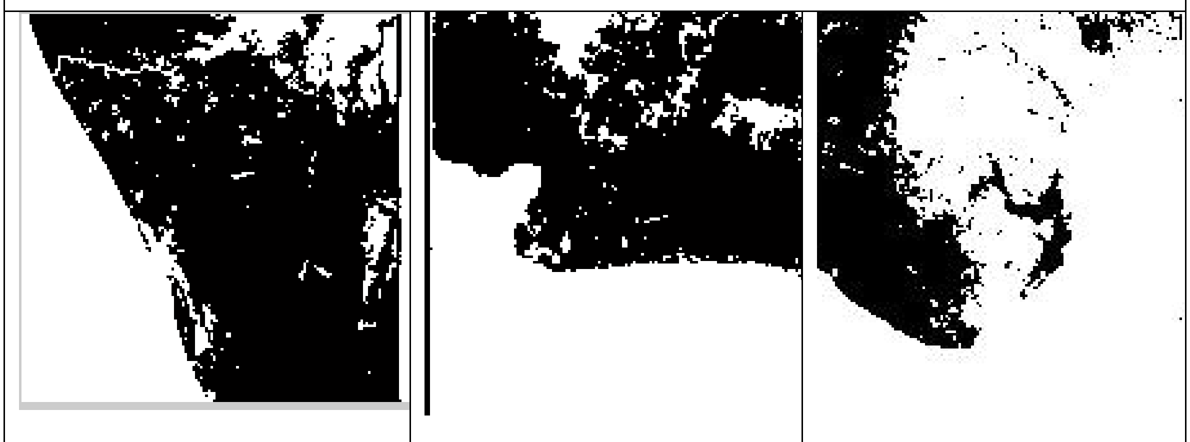
The six texture features are extracted using GLCM and subjected to LVQ to satisfy the necessary and sufficient conditions to achieve maximum accuracy. The training data has led to results with 96% accuracy.

Subjecting the combination of extracted texture features onto the SVM classifier resulted in 98% accuracy for classifying input data. The classified results are depicted in Figure 8.



Kochi Region	Indonesia Region	Kanyakumari Region

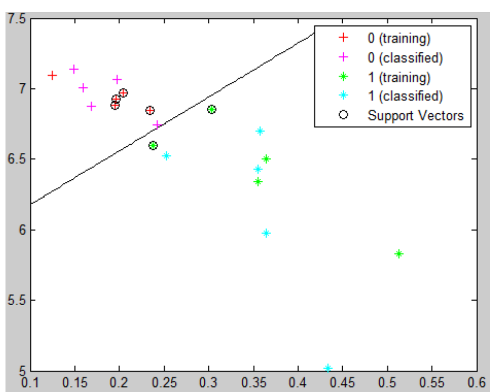
FIGURE 6: PCA based K-means clustering performed on 2005 images



Kochi Region	Indonesia Region	Kanyakumari Region

FIGURE 7: PCA based K-means clustering performed on 2010 images

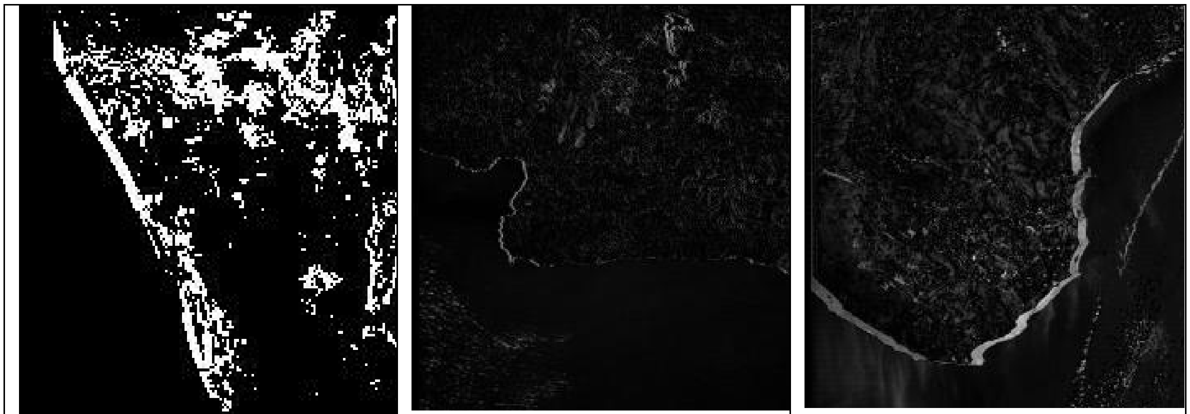
Non-Water body



Water body

FIGURE 8: The Classification Using SVM

So it is possible to quantifiably depict the amount of increase of water body in 2010 when compared to 2005 as shown in Figure 9.



Kochi Region	Indonesia Region	Kanyakumari Region

FIGURE 9: Change depicted over the Five Years

Area	2005	2005	2010	2010	Changes detected in the period of 5 years
	Water	Non-water	water	Non-water	
Kochi	26.1200	73.88	29.2	70.8	3.08%
Kanyakumari	46.4966	53.5034	44.2856	55.7144	0.16%
Indonesia	46.6476	53.3524	46.6599	53.3401	0.01%
Kolkata	27.4445	72.555	31.7780	68.2220	4.333%
Vishakhapatnam	62.8494	37.1506	63.5956	36.4044	0.74%
Sydney	56.3477	43.6523	60.0342	39.9658	3.68%
Bhubaneswar	62.8494	37.1506	63.5956	36.4044	0.74%

TABLE 1: Coverage areas of water body and non-water body of different landscapes in percentage

The percentage of changes in water bodies and land masses can be interpreted from the results shown in Figure 9. The count has been materialised by taking pixel values of the classified image into account as the input, and then pixels of the water region are counted first which leads to the count of the land mass. The change in coverage area of water body has recorded a significant increase over five years in the study region, as shown in Table 1. The increase in the coverage areas of water bodies due to global warming has been substantially proven through these results.

4. CONCLUSION

The images are classified using the unsupervised algorithms which do not need pre determined cluster definitions, namely the PCA based K-Means Clustering. In order to support these findings the supervised mechanisms, LVQ and SVM are employed. Instead of applying the texture features separately, the combination of textures from the training data sets are subjected to the LVQ to yield better results. After achieving the required accuracy, the features are subjected to SVM. In SVM, the two classes are identified and differentiated, which helped in finding out the classes in the target image accurately. This work will intend to define a mechanism to aid long-term planning by predicting preventable losses and to generalise this scheme for supporting crisis management due to global warming.

5. REFERENCES

[1] Kaichang Di F., Ruijin Ma, A., Jue Wang, W., Ron Li, U., “Coastal mapping and change detection using high-resolution IKONOS satellite imagery”, Digital Government Society of North America, pp.1-4, 2003.

[2] Krishnan, V., Swoboda, J., Yarman, C. E., and Yazici, B., “Multistatic Synthetic Aperture Radar Image Formation”, IEEE transactions on image processing, volume: 19, no: 5, pp.1290-1306, 2010.

[3] Kwang In Kim, Keechul Jung Se, Hyun Park and Hang Joon Kim., “Support Vector Machines for Texture Classification”, IEEE Transaction on Pattern Analysis and Machine Intelligence, volume: 24, no: 11, pp. 1542-1550, 2002.

[4] Nath, K., and Deb S. K., “Water-Body Area Extraction from High Resolution Satellite Images - An Introduction, Review, and Comparison”, International Journal of Image Processing (IJIP), volume: 3, Issue: 6, pp. 353-372, 2010.

[5] Ravichandran, K. S., and Ananthi, B., “Color Skin Segmentation Using K-Means Cluster”, International Journal of Computational and Applied Mathematics, volume: 4, no: 2, pp.153-157, 2009.

[6] Rowman, J. F., and George, P. T., “Geological Units Classification of Multispectral Images by Using Support Vector Machines”, International Conference on Intelligent Networking and Collaborative Systems, pp.267-272, 2009.

[7] Sun, B. X., and Huang, D.S., “Texture Classification Based on Support Vector Machine and Wavelet Transform”, Proceeding of the Fifth World Congress on Intelligent Control and Automation, pp. 15-19, 2004.

[8] Peng Xu, Min Dai, Andrew K. Chan., “Texture Classification Using Optimized Support Vector Machines”, IEEE (C), pp.544-547, 2004.

[9] Zhang Zhaohui, T., Veronique Prinet R. and Songde, M.A., “Water body extraction from multisource satellite images”, Proc. IEEE Geoscience and remote Sensing Symposium, volume: 6, pp.3970-3972, 2003.

[10] Hyeran Byun and Seong-Whan Lee., “Application of Support Vector machines for Pattern Recognition: A Survey”, SVM 2002, LNCS 2388, pp.213-236, 2002.

[11] Guosheng Wang., “A Survey on Training Algorithms for Support Vector Machine Classifiers”, Fourth International Conference on Networked Computing and Advanced Information Management, IEEE, pp. 3-128, 2008. DOI 10.1109/NCM.2008.103.

[12] Gonzalez, C. Woods, R.E., “Digital Image Processing” (Publishing House of Electronics industry, 2002, 2nd edn.).

[13] Bagli, S and Soille, P., “Morphological automatic extraction of coastline from landsat images”, In Proceedings of the Fifth International Symposium on GIS and Computer Cartography for Coastal Zone Management, volume 3, pp. 58–59, Genova, 2009.

[14] Clausi, D.A, and Yue, B., In Proceedings of the Fifth International Symposium on “GIS and Computer Cartography for Coastal Zone Management”, volume 3, pp. 58–59, Genova, 2009.

[15] Christina Corbane, Nicolas Baghdadi, Xavier Descombes, Geraldo Wilson Junior, Nicolas Villeneuve and Michel petit., "Comparative Study on the Performance of Multiparameter SAR Data for Operational Urban Areas Extraction Using Textural Features", IEEE Geoscience and Remote sensing Letters, vol.6, no.4, pp.728-732, 2009.

[16] Chethan, Raghavendra and Hemantha Kumar., "Texture based Approach for Cloud Classification using SVM", International conference on advances in technologies in Communication and Computing, pp.688-690, 2009. DOI: 10.1109/ARTCom.2009.43.

[17] Zhi-Zhong Wang and Jun-Hai Yong., "Texture Analysis and Classification with Linear Regression Model Based on Wavelet Transform", IEEE Transactions on Image Processing, vol.17, no.8, pp. 1421-1430, 2008.