Image Segmentation Using Wavelet Packet Frames and Neuro-fuzzy Tools

Mausumi Acharyya and Malay K. Kundu

***Abstract*— The present article describes a image segmentation technique using M-band wavelet packet frames features. Those wavelet features are evaluated and selected using an efficient neuro-fuzzy feature evaluation techniqu. Both the feature extraction and neuro-fuzzy feature evaluation processes are of unsupervised type. They do not require the knowledge of number and distribution of classes corresponding to various landcovers in remotely sensed images. The effectiveness of the methodology is demonstrated on IRS-1A and SPOT images.**

***Index Terms*— M-band wavelet packet frames, adaptive basis selection, fuzzy feature evaluation index, neural networks, remotely sensed image**

I. Introduction

A WAVELET means small wave (oscillation), which has its energy concentrated in time(space). Wavelet transform is a tool for the analysis of transient, non-stationary, or time varying event. It is a two-dimensional expansion set for any one-dimensional signal. The wavelet expansion gives time-frequency localization, which means most of the energy of the signal is well represented by a few expansion coefficients. Wavelet system also satisfies the multiresolution conditions. This means that the lower resolution (coarser) representation of a signal can be calculated from its higher resolution information. Typically, the wavelet transform maps an image onto a low-resolution image and a series of detail images. The lower resolution image is obtained by iteratively blurring the original image; the detail images contain the information lost during this operation. The energy or mean deviation of the detail images are most commonly used features for texture based image analysis.

The major aim of the image processing /analysis research is to develop better tools, that may extract different perspectives on the same image, to understand not only its content, but also its meaning and significance. However, no image processing system can compete with the human visual system in terms of accuracy, but it can easily outperform the latter on the observational consistency, and ability to carry out detailed mathematical operations. With the passage of time, image processing research has broadened its approach from basic pixel based low-level operation to high-level analysis, to have a better semantic understanding of images. This is based on the relationship between image components and their context. What comprises an image, must be identified clearly before any further attempt is made for its detail analysis.

Segmentation is the technique for partitioning an image space into a finite number of non-overlapping and meaningful regions. The segmentation of different landcover regions precisely, in a remotely sensed image. This has been recognized as a complex problem for a long time. Remotely sensed images usually have inferior illumination quality and are mainly due to different type of environmental distributions. Spatial resolution of these images are also comparatively low. The natural scene mostly contains many objects (regions), *e.g.*, vegetation, water bodies, habitation, concrete structures, open spaces etc., but these regions are not very well separated because of low spatial resolution (spatial ambiguities). Moreover, generally the gray value assigned to a pixel is due to the average reflectance of different types of landcovers present in an area that corresponds to a pixel. Assigning unique class levels with certainty for all the pixel is thus a genuine problem of remotely sensed images. Fuzzy set theory provides a way of handling this uncertainties in a better way.

Texture is a concept used to indicate some spatial properties of image regions. Most naturally occurring patterns and natural surfaces exhibit texture. It is a fundamental characteristic of an image and plays an important role in the human visual system for recognition and interpretation of images. Despite of its pivotal role in the analysis of image data, there exist neither a formal /precise definition of texture nor an obvious quantitative measure to characterize it. Image texture can be qualitatively expressed in terms of coarseness, fineness, granularity, lineation, randomness and smoothness. The analysis of image texture content is extremely important in image analysis. It requires the under-standing of how humans discriminate between different texture types and how to model algorithms to perform image analysis task in a best possible way. Another important aspect of texture is scale. The importance of scale in texture descriptions is clear from the fact there is a change in appearance of most textures when viewed at different resolutions, and also during the empirical division from macro to micro textures. Texture can also be defined as a local statistical distribution of pixel pattern (micro region) in observer's domain. Psychovisual studies reveal that the human visual system processes images in multiple scales The visual cortex has separate cells that decomposes images into filtered images of various band of frequencies and orientation. Texture is especially suited for this type of multiresolution analysis,

using both frequency and spatial information, because of its inherent characteristics. During the past two decades, wavelet analysis has become an important paradigm for multiresolution analysis, and have found important applications in various real life applications, ranging from seismology to image analysis and compression.

Remotely sensed images may contain information over a large range of scales and the spatial frequency structure also changes over different regions (*i.e.,* non-periodic signal). In remote sensing perspective, the resolution of the imagery may be different in many cases, and so it is important to understand how information changes over different scales of imagery. These reasons justify the use of multiresolution type analysis for this purpose and is most effective using wavelets. Moreover, wavelet theory is well suited in this area of study where signals are complex and non-periodic. Furthermore, wavelets are particularly good in describing a scene in terms of the scale of the textures in it.

Texture is an important property of all reflective natural surfaces which helps human visual perception system to segment and classify different objects in a digital image. In a remotely sensed image,texture is considered to be the visual impression of coarseness or smoothness caused by the variability or uniformity of image tone. These textural properties of a remotely sensed image are likely to provide valuable information for analysis (classification/segmentation), where different object regions are treated as different texture classes, *i.e.,* a multitexture segmentation problem. Note that segmentation of these images is necessary in order to identify regions of vegetation, habitation, water bodies, city area etc.

Effective classification and segmentation of images based on textural features is of key importance to many applications like image analysis, remote sensing, robot vision, query by content in large image data bases and many others. The wide variety of texture analysis methods that have been developed over the past two decades are reviewed in [42], [52]. Earlier approaches focused on first-order and second-order statistics of textures [21], [12], description using texture primitives and system rules [20]. There has been an extensive study on model based approaches like Markov random fields [13], [19], [24] and local linear transforms [27].

Haralick *et al.* [21] have used gray level co-occurrence features to analyze remotely sensed images. They have computed gray level co-occurrence matrices for a distance of one with four directions($0^0, 45^0, 90^0$ and 135^0). For a seven class problem they have achieved 80% classification accuracy. Rignot and Kwok [45] have analyzed SAR (Synthetic Aperture Radar) images using texture features computed from gray level co-occurrence matrices. However, they supplement these features with knowledge about the properties of SAR images. The use of various texture features have been studied for analyzing SAR images by Du [17]. He used the Gabor filters for extracting texture features and successfully segmented SAR images into categories of water, new forming ice, older ice,and multi-year ice.

All these approaches are restricted to the analysis of spatial interactions over relatively small neighborhoods, and are best suited for analysis of micro textures. In general, natural phenomena do not have a simple mathematical representation and do not even obey the restrictions imposed by several methods in order to use them suitably.

Psychovisual studies reveals that the human visual system processes images by decomposing them into filtered images of various frequencies and orientation at different scales that is capable of preserving both local and global information. This multiscale processing of the human visual system is the strong motivation for using methods based on these concepts for texture analysis [23], [11], [9], [53]. An extensive study has been made which certainly reveals the superiority of these multiscale processing over the more traditional ones.

Multiresolution techniques intend to transform images into representation in which both frequency and spatial information are present. Wavelet theory provides a more formal, precise and unified approach to multiresolution representations [15], [30]. The importance of scale in texture descriptions is clear from the change in appearance of most textures when viewed in different resolutions, and from the empirical division into micro and macro textures. These recent findings have motivated several important studies for texture analysis [32], [43], [7], [10], [55], [54].

The work of Mecocci *et al.* [33] have presented a wavelet-based algorithm combined with a *fuzzy c-means* classifier. Lindsay *et al.* [28] have used the 1D discrete wavelet transform (DWT) based on Daubechies wavelet filter. A wavelet-based texture feature set is derived in [18]. It consists of the energy of subimages obtained by the overcomplete wavelet decomposition of local areas in SAR images, where the downsampling between wavelet levels is omitted.

Simard *et al.* [47] studied the use of a decision tree classifier and multiscale texture measures to extract thematic information on the tropical vegetation cover from the Global Rain Forest Mapping (GRFM). The aim of the work by [35] is to show how coastline can be derived from SAR images by using wavelet and active contour methods. In a first step an edge detection method suggested by Mallat *et al.* [31] is applied to SAR images to detect all edges above a certain threshold. A block-tracing algorithm (BA) then determines the boundary area between land and water. Several other wavelet-based segmentation for geoscience and remote sensing applications have also been reported in the literature [51], [49].

Other approaches to segmentation of remotely sensed images include various fuzzy thresholding techniques reported in [38]. Genetic algorithm as a classifier has been investigated in the domain of satellite imagery for partitioning different land cover regions from satellite images, having complex/overlapping class boundaries in [6]. Muchoney and Williamson [34] have shown neural network classifiers to provide supervised classification results that significantly improve on traditional classification algorithms such as the Bayesian (maximum likelihood [ML])classifier.

All of these above methods use supervised classification where *a priori* knowledge about the images are essential. We apply a methodology to carry out this segmentation where no *a priori* knowledge about the image is available.

Applications of octave band wavelet decomposition scheme for texture segmentation to remotely sensed images have been

studied in [33], [28], [18], [47], [35] as mentioned earlier. The octave band wavelet decomposition [30], [54] provides a logarithmic frequency resolution and are not suitable for the analysis of high frequency signals with relatively narrow bandwidth. Therefore, the main motivation of the present work is to utilize the decomposition scheme based on M-band ($M > 2$) wavelets, which, unlike the standard wavelet, provides a mixture of logarithmic and linear frequency resolution [48] [5] and hence can characterize a texture more efficiently. We conjecture that M-band wavelet decomposition would give improved segmentation results than the methods mentioned earlier. But the use of M-band wavelet decomposition gives rise to a large number of features, which incurs redundancy and confusion. Therefore, selection of the appropriate features using some feature selection algorithms is required.

The proposed methodology for segmenting a remotely sensed satellite image has two parts. The first part deals with extraction of texture features using M-band wavelet packet frame, followed by their neuro-fuzzy evaluation for selecting an optimal set of features. Note that the M-band ($M > 2$) wavelet transform is a tool for viewing signals at different scales and decomposes a signal by projecting it onto a family of functions generated from a single wavelet basis via its dilations and translations [48] [5]. Neuro-fuzzy computing [39] [37] which integrates the merits of fuzzy set theory and artificial neural networks (ANN's), enables the feature selection process artificially more intelligent. Incorporation of fuzzy set theory, as described above, helps one to deal with uncertainties in remotely sensed images in an efficient manner. ANN is used here for the task of optimization in an adaptive manner.

It may be noted that neuro-fuzzy hybridization is a widely used tool of soft computing paradigm. Soft computing is a consortium of methodologies which work synergistically and provides, in one form or another, flexible information processing capabilities for handling real life ambiguous situations. Its aim is to tolerate the imprecision, uncertainty, approximate reasoning and partial truth in order to achieve tractability, robustness, low solution cost and close resemblance with human like decision making. At this juncture, fuzzy logic (FL) and artificial neural networks (ANN's) and genetic algorithm (GA) are the three principal components where FL provides algorithms dealing with imprecision and uncertainty, ANN is used as the machinery for learning and adaptation, and GA is used for optimization and searching [57] [37].

The article is organized as follows. Section II presents the mathematical framework of M-band wavelets. Section III gives a brief overview of neuro-fuzzy hybridization. Section IV discusses about filtering technique used in our investigation and the extraction of features. Section V provides the neuro-fuzzy feature selection algorithm. Segmentation and the quality measure is discussed in Section VI while Section VII analyzes experimental results and the article concludes with Section VIII.

II. M-band Wavelet Transform

A. M-band wavelets

The standard dyadic (2-band) wavelets are not suitable for the analysis of high-frequency signals with relatively narrow bandwidth. To resolve this problem, M-band orthonormal wavelets [48] were developed as a direct generalization of the 2-band orthogonal wavelets of Daubechies [14]. These M-band wavelets are able to zoom in onto narrowband high frequency components of a signal and have been found to provide better energy compaction than 2-band wavelets [22].

The scaling function $\phi(x)$ is given by [48]

$$\phi(x) = \sum_k h(k)\sqrt{M}\phi(Mx - k). \tag{1}$$

Additionally there are$(M-1)$ wavelets which are given by [48]

$$\psi_l(x) = \sum_k g_l(k)\sqrt{M}\psi(Mx - k). \tag{2}$$

In discrete form, these functions can be indexed by scale parameter j and translation parameter k, and is written as [48]

$$\phi_{j,k}(x) = \sum_k M^{j/2}\phi(M^j x - k) \tag{3}$$

and

$$\psi_{j,k,l}(x) = \sum_k M^{j/2}\psi_l(M^j x - k), l = 1, \ldots, M-1. \tag{4}$$

The subspaces spanned by the functions $\phi_{j,k}(x)$ and $\psi_{j,k,l}(x)$ be respectively defined as, $\mathcal{V}_j = \overline{span_k \phi_{j,k}}; \forall k \in Z$, and $\mathcal{W}_{j,l} = \overline{span_k \psi_{j,k,l}}; \forall k \in Z$ [48]. It follows from equation (1) that the subspaces $\mathcal{V}_j$ have a nested property. If the scaling and the wavelet functions satisfy the orthonormality condition, the subspaces $\{\mathcal{W}_{j,l}\}$ form an orthogonal decomposition of the l^2 function space and are related to the nested subspaces $\mathcal{V}_j$ by

$$\mathcal{V}_j = \mathcal{V}_{j+1} \oplus [\oplus_{l=1}^{M-1} \mathcal{W}_{j+1,l}] \tag{5}$$

Thus a function $f(x) \in l^2$ can be expressed in terms of the sum of projections onto subspaces $\mathcal{V}_j$ and $\mathcal{W}_{j,l}$ as

$$f(x) = \sum_k c(k)\phi_{j,k}(x) + \sum_{l=1}^{M-1}\sum_k d_l(k)\psi_{j,k,l}(x) \tag{6}$$

where $\oplus$ is the orthogonal plus. This is the discrete M-band wavelet transform (DMbWT). The expansion coefficients can be expressed as $c(k) = \langle f, \phi_{j,k}\rangle$ and $d_l(k) = \langle f, \psi^{(j)}_{j,k}\rangle, \quad l = 1, \ldots M-1$, where $\langle \alpha, \beta\rangle$ represents the inner product of α and β.

We can further extend our discussion in defining wavelet packets as a generalization of orthonormal and compactly supported wavelets [14]. From the subband filtering point of view, the difference between wavelet packet transform (DWPT) and standard wavelet transform (DWT) is that the former recursively decomposes the high frequency components as well, unlike the other, thus resulting in a tree structured multiband extension of the wavelet transform.

III. Neuro-fuzzy hybridization

The theory of fuzzy set has been introduced in 1965 by Zadeh [56] as a new way of representing uncertainties in everyday life. This theory provides an approximate and yet effective means for describing the characteristics of a system which is too complex or ill-defined to admit precise mathematical analysis. It is reputed to handle, to a reasonable extent, uncertainties (arising from deficiencies of information) in various applications particularly in decision making models under different kinds of risks, subjective judgment, uncertainties and ambiguity. The deficiencies may result from various reasons, *viz.*, incomplete, imprecise, not fully reliable, vague or contradictory information depending on the problem. Since this theory is a generalization of the classical set theory, it has greater flexibility to capture various aspects of incompleteness or imperfection in information about a situation.

Artificial neural networks (ANN) [16], [29] are signal processing systems that try to emulate the behavior of biological nervous systems, by providing a mathematical model of combination of numerous neurons connected in a network. These can be formally defined as *massively parallel interconnections of simple (usually adaptive) processing elements (called neurons) that interact with objects of the real world in a manner similar to biological systems*. The benefit of neural nets lies in the high computation rate provided by their inherent massive parallelism. This allows real-time processing of huge data sets with proper hardware backing. All information is stored distributed among the various connection weights. The redundancy of interconnections produces a high degree of robustness resulting in a *graceful degradation* of performance in case of noise or damage to a few nodes/links. Neural network models have been studied for many years particularly in the field of pattern recognition and image processing.

We see that fuzzy set theoretic models try to mimic human reasoning and the capability of handling uncertainty, whereas the neural network models attempt to emulate the architecture and information representation schemes of the human brain. Integration of the merits of fuzzy set theory and neural network theory therefore promises to provide, to a great extent, more intelligent systems (in terms of parallelism, fault tolerance, adaptivity and uncertainty management) to handle real life decision making problems. For the last ten years or so, there have been several attempts [39] [37] by researchers over the world in making a fusion of the merits of these theories under the heading *neuro-fuzzy computing* for improving the performance in decision making systems.

IV. Wavelet feature extraction

The notion of M-band wavelet, as described in Section II-A, is used here to extract multiscale wavelet features of remotely sensed images. The methodology involves M-band wavelet packet filtering of an input image followed by adaptive basis selection. Subsequently, features are computed from this set of selected basis by using a nonlinear operator and smoothing filter. These features are then evaluated and selected using a neuro-fuzzy algorithm (described in Section V).

A. M-band wavelet packet filters and adaptive basis selection

The objective of filtering is to transform the edges in a texture image into detectable discontinuities. The filter bank in essence is a set of bandpass filters which select frequency and orientation. In the filtering stage, we make use of orthogonal and linear phase M-band (M=4) wavelet following [5].

The motivation for a larger $M(M > 2)$ comes from the desire to have a more flexible tiling of the time-frequency (scale-space) plane than that resulting form 2-band wavelet. It also enables to have some regions of uniform bandwidths rather than the logarithmic spacing of the frequency responses. Although the M-band wavelet decomposition results in a combination of linear and logarithmic frequency (scale) resolution, we conjecture that a further recursive decomposition of the high frequency regions would characterize textures better.

The M-band wavelet decomposition can be interpreted as signal decomposition in a set of independent, spatially oriented frequency channels. The discrete normalized scaling $(\phi_{j,k})$ and wavelet $(\psi_{j,k,l})$ basis functions are defined in terms of their filter responses as,

$$\begin{aligned} \phi_{j,k}(x) &= M^{j/2} h_j(M^j x - k) \\ &\text{and} \\ \psi_{j,k,l}(x) &= M^{j/2} g_{j,l}(M^j x - k) \end{aligned} \tag{7}$$

where j and k are the dilation and translation parameters and $l = 1, \ldots, M-1$ is the number of wavelet functions. h_j and $g_{j,l}$ are respectively the sequence of lowpass and bandpass filters of increasing width indexed by j, which are expanded by inserting an appropriate number of zeros between taps and satisfy the quadrature mirror filter (QMF) condition and are called the analysis (synthesis filters).

The standard wavelet decomposition method require a downsampling by a factor M at each scale. But this decomposition is not translation invariant which is desirable for image analysis tasks. A possible solution is achieved by using an overcomplete wavelet decomposition called a discrete wavelet frame (DWF). In the following we give a discrete M-band wavelet frame (DMbWF) decomposition, which is similar to discrete M-band wavelet transform (DMbWT), except that no downsampling is done between scales (levels of decomposition). It is mention worthy that there are other alternatives to alleviate this problem of shift (translation) variance by using complex wavelets [25].

Let $I(x, y) \in l^2(R)$ be an image in 2-D. Its DMbWF decomposition is given by,

$$\begin{aligned} c_j(x,y) &= [h_{j,x} * [h_{j,y} * c_{j-1}]](x,y) \\ d^{s_l^1}_{j,l}(x,y) &= [h_{j,x} * [g_{j,l,y} * c_{j-1}]](x,y) \\ d^{s_l^2}_{j,l}(x,y) &= [g_{j,l,x} * [h_{i,y} * c_{i-1}]](x,y) \\ d^{s_l^3}_{j,l}(x,y) &= [g_{j,l,x} * [g_{j,l,y} * c_{i-1}]](x,y) \qquad \text{for } l = 1,2,3 \end{aligned} \tag{8}$$

where, $*$ denotes the convolution operator, and $c_0 = I(x, y)$ the original 2D signal. $h_{j,x}$ $(g_{j,l,x})$ and $h_{j,y}$ $(g_{j,l,y})$ are the lowpass (bandpass) filtering along x and y directions, respectively, corresponding to different scales.The number of subbands or frequency channels resulting from such decomposition (8) is found to be 16 (considering all possible combinations of l and the subband corresponding to c_j).

$c_j(x,y)$ corresponds to the lower frequencies, the $d^{s^i}_{j,l}$ are obtained by bandpass filtering in a specific direction and thus contain the detail information at scale j.

We extend our discussion in defining discrete M-band wavelet packet transform (DMbWPT) as a generalization of DMbWT [14]. As mentioned earlier the difference between DMbWPT and DMbWT as described above, is that the former recursively decomposes the high frequency components as well, unlike the other, thus resulting in a tree structured multiband extension of the wavelet transform. At scale $j = J$, the image is first decomposed into $M \times M$ channels using all the filters h_j and $g_{j,l}$ with $l = 1, 2, 3$, and without downsampling. The process is repeated for each of the subbands for subsequent scales.

For extraction of textural features of a remotely sensed image, it is appropriate to detect the most significant frequency bands contained in the image. This leads naturally to a tree structured wavelet transform of the image. An M-band wavelet packet decomposition gives M^{2^k} number of bases, for a decomposition depth of k. It is usually redundant to decompose all the subbands in each scale to achieve the full tree of decomposition. Also it is quite evident that an exhaustive search to determine the optimal basis from this large set is computationally expensive.

In order to find out a suitable basis without going for a full decomposition, we make use of an adaptive decomposition algorithm using a maximum criterion of textural measures extracted from each of the subbands [3] [1]. Then the most significant subbands are identified and it is decided whether further decomposition of the particular channel would generate more information or not. This search is computationally efficient and enables one to zoom into any desired frequency channel for further decomposition [4] [2] [26].

After decomposition of the image into $M \times M$ channels, as described above energy for each subband is evaluated. Among these subbands, those for which energy values exceed $\epsilon_1\%$ of the energy of the parent band, are considered and decomposed further. We have further decomposed a subband if its energy value is more than some $\epsilon_2\%$ of the total energy of all the subbands at the current scale. This step results in a set of 4-band wavelet packet bases. These bases corresponding to different resolutions are assumed to capture and characterize effectively different scales of texture of the input image. Empirically we have seen that a value of $\epsilon_1 = 2\%$ and $\epsilon_2 = 10\%$ are good choice for the images we have considered here.

B. Computation of features

After selection of the significant bases, a local estimator which constitutes a nonlinear operator followed by a smoothing filter, is applied to each subbands. This estimates a textural feature of a subband image in a local region around each pixel. A nonlinearity is needed in order to discriminate texture pairs with identical mean brightness and second-order statistics. There are a wide variety of textural measures. In this study we have used *energy* as the textural measures available. Since the magnitude of the correlation between the wavelet function and the image is all that is important, we have used *absolute values* of the wavelet coefficients as a generalized energy definition.

For a subband image $F_s(x,y)$ of subband number s, where $0 \le x \le M-1,\ 0 \le y \le N-1$, the local energy $Eng_s(x,y)$ around the $(x,y)^{th}$ pixel can be formally defined as

$$Eng_s(x,y) = \mid (F_s(m,n)) \mid, \tag{9}$$

This step is succeeded by a Gaussian low pass (smoothing) filter $h_G(x,y)$ to get a feature image. Formally, the feature image $Feat_s(x,y)$ corresponding to subband image $F_s(x,y)$ is given by

$$Feat_s(x,y) = \sum_{(a,b)\in G_{xy}} \Gamma(F_s(a,b)h_G(x-a, y-b)),$$

where $\Gamma(\cdot)$ gives the energy measure and G_{xy} is a $G \times G$ window centered at pixel with coordinates (x,y). This step results in a set of feature images $Feat_s(x,y)$, from which a set of feature vectors are derived. These feature vectors corresponding to the decomposed images at different resolutions are assumed to capture and characterize effectively different scales of texture of the remotely sensed image.

V. Neuro-fuzzy feature evaluation

The wavelet features extracted in Section IV are now evaluated in a neuro-fuzzy framework under unsupervised learning. The method is a modification of an earlier one [36]. This modification enables one to handle large data sets in an efficient manner. Note that for an image a large number of pattern vectors are generated as described in Section IV. These selected features are then used for the purpose of segmenting the different regions in remotely sensed images.

A. Fuzzy feature evaluation index and membership function

The feature evaluation index for a set of transformed features is defined as

$$E = \frac{2}{s(s-1)} \sum_{p} \sum_{q \ne p} \frac{1}{2} [\mu^T_{pq}(1-\mu^O_{pq}) + \mu^O_{pq}(1-\mu^T_{pq})]. \tag{10}$$

Here μ^O_{pq} and μ^T_{pq} are the degree that both the pth and qth patterns belong to the same cluster in the n-dimensional original feature space, and in the n'-dimensional ($n' \le n$) transformed feature space respectively. μ values determine how similar a pair of patterns are in the respective features spaces. s is the number of samples on which the feature evaluation index is computed.

The feature evaluation index decreases as the membership value representing the degree of belonging of pth and qth patterns to the same cluster in the transformed feature space tends to either 0 (when $\mu^O < 0.5$) or 1 (when $\mu^O > 0.5$). In other words, the index decreases as the decision on the similarity between a pair of patterns (*i.e.*, whether they lie in the same cluster or not) becomes more and more crisp. This means, if the intercluster/intracluster distances in the transformed space increase/decrease, the feature evaluation index of the corresponding set of features decreases. Therefore, our objective is to select those features for which the evaluation index becomes minimum; thereby optimizing the decision

on the similarity of a pair of patterns with respect to their belonging to a cluster.

The membership function μ_{pq} in a feature space, satisfying the characteristics of E (10), may be defined as [36]

$$\begin{aligned} \mu_{pq} &= 1 - \frac{d_{pq}}{D} \quad if\ d_{pq} \leq D, \\ &= 0, \qquad\quad otherwise, \end{aligned} \tag{11}$$

where the distance d_{pq} between the pth and qth patterns can be written as

$$\begin{aligned} d_{pq} &= [\sum_i w_i^2 (x_{pi} - x_{qi})^2]^{\frac{1}{2}}, \\ &= [\sum_i w_i^2 \chi_i^2]^{\frac{1}{2}}. \quad \chi_i = (x_{pi} - x_{qi}), \end{aligned} \tag{12}$$

The term $w_i \in [0, 1]$ represents weighting coefficient corresponding to ith feature, and x_{pi} and x_{qi} are values of ith feature (in the corresponding feature space) of pth and qth patterns, respectively. Note that, the higher the value of d_{pq}, the lower is the similarity between pth and qth patterns, and *vice versa*. D is a parameter which reflects the minimum separation between a pair of patterns belonging to two different clusters. When $d_{pq} = 0$ and $d_{pq} = D$, we have $\mu_{pq} = 1$ and 0, respectively. If $d_{pq} = \frac{D}{2}$, $\mu_{pq} = 0.5$. That is, when the distance between the patterns is just half the value of D, the difficulty in making a decision, whether both the patterns are in the same cluster or not, becomes maximum; thereby making the situation most ambiguous. Note that the computation of μ_{pq} in (11) does not require the information on class label of the patterns.

The term D in (11) may be expressed as $D = \alpha d_{max}$, where d_{max} is the maximum separation between a pair of patterns in the entire feature space, and $0 < \alpha \leq 1$ is a user defined constant. α determines the degree of flattening of the membership function (11). The higher the value of α, more will be the degree, and *vice-versa*. d_{max} is defined as

$$d_{max} = [\sum_i (x_{maxi} - x_{mini})^2]^{\frac{-1}{2}}, \tag{13}$$

where x_{maxi} and x_{mini} are the maximum and minimum values of the ith feature in the corresponding feature space.

The membership value μ_{pq} is dependent on w_i. The values of w_i (< 1) make the μ_{pq} function of (11) flattened along the axis of d_{pq}. The lower the value of w_i, the higher is extent of flattening. In the extreme case, when $w_i = 0$, $\forall i$, $d_{pq} = 0$ and $\mu_{pq} = 1$ for all pair of patterns, *i.e.*, all the patterns lie on the same point making them indiscriminable.

The weight w_i in (12) reflects the relative importance of the feature x_i in measuring the similarity of a pair of patterns. The higher the value of w_i, the more is the importance of x_i in characterizing a cluster or discriminating various clusters. $w_i = 1$ (0) indicates most (least) importance of x_i.

As mentioned earlier, our objective is to minimize the evaluation index E (10) which involves the terms μ^O and μ^T. The computation of μ^T requires (11)–(13), while μ^O needs these equations with $w_i = 1$, $\forall i$. Therefore, the evaluation index E (10) is a function of $\mathbf{w}$, if we consider ranking of n features in a set. The problem of feature selection/ranking thus reduces to finding a set of w_is for which E becomes minimum;

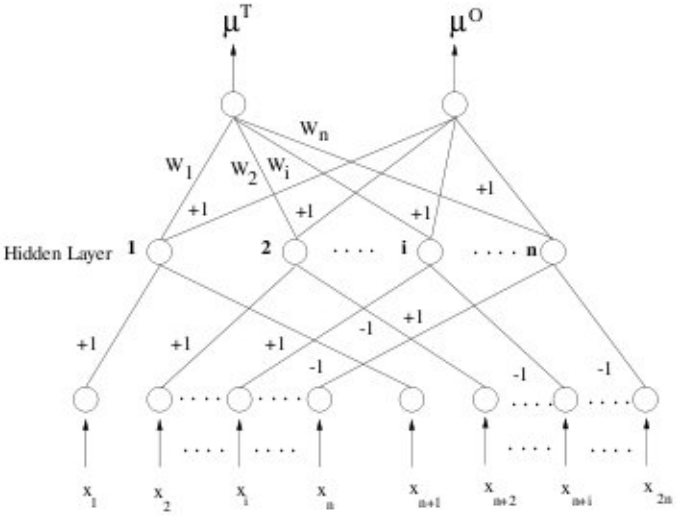

Fig. 1. A schematic diagram of the neural network model for feature selection.

w_is indicating the relative importance of x_is. The task of minimization, as in [36], is performed using gradient-descent technique in a connectionist framework under unsupervised mode. This is described below.

B. Connectionist model

The network (Figure 1) consists of an input, a hidden and an output layer [36]. The input layer consists of a pair of nodes corresponding to each feature, *i.e.*, the number of nodes in the input layer is $2n$, for n-dimensional (original) feature space. The hidden layer consists of n number of nodes which compute the part χ_i^2 of Eqn. (12) for each pair of patterns. The output layer consists of two nodes. One of them computes μ^O, and the other μ^T. The feature evaluation index E (10) is computed from these μ-values off the network.

Input nodes receive activations corresponding to feature values of each pair of patterns. A jth node in the hidden layer is connected only to an ith and $(i+n)$th input nodes via connection weights $+1$ and -1, respectively, where $j, i = 1, 2, \ldots, n$ and $j = i$. The output node computing μ^T-values is connected to a jth node in the hidden layer via connection weight W_j ($= w_j^2$), whereas that computing μ^O-values is connected to all the nodes in the hidden layer via connection weights $+1$ each.

During learning, each pair of patterns are presented at the input layer and the evaluation index is computed. The weights W_js are updated in order to minimize the index E. The task of minimization of E (10) with respect to $\mathbf{W}$ is performed using gradient-descent technique, where the change in W_j ($\triangle W_j$) is computed as

$$\triangle W_j = -\eta \frac{\partial E}{\partial W_j}, \forall j, \tag{14}$$

η is the learning rate. Note that, d_{max} is directly computed from the unlabeled training set. The values of d_{max} and α are stored in both the output nodes for the computation of D. For details concerning the operation of the network, one may refer to [36].

C. Modified neuro-fuzzy algorithm for handling large data

As we have seen in Section IV, the number of patterns generated for an $N \times N$ image is N^2, *i.e.*, $s = N^2$. Each of these patterns corresponds to a pixel and has all the multiscale wavelet feature extracted in Section IV. Therefore, for selecting an optimal set of features out of them, the number of patterns to be presented to the connectionist system in one epoch, during its training, is $\frac{s(s-1)}{2} = \frac{N^2(N^2-1)}{2}$, which is a very large quantity. This incurs a very high computational time. In order to avoid this situation, *i.e.*, in order to make the neuro-fuzzy algorithm computationally more efficient, we, first of all, apply a clustering algorithm (e.g., k-means clustering algorithm) on the entire feature space, for grouping the data, and the cluster centers $\boldsymbol{cen}_q$'s are noted. Then two sets of samples, namely, $S = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_p, \cdots, \boldsymbol{x}_{N^2}\}$ and $S_c = \{\boldsymbol{cen}_1, \boldsymbol{cen}_2, \cdots, \boldsymbol{cen}_c\}$ are formed. That is, S is the entire training set, and S_c is the set of c cluster centers (for c clusters) obtained by the clustering algorithm. Now the similarity between the patterns and these cluster centers are computed, instead of computing it for every pair of patterns. These cluster centers are considered as representatives (prototypes) of all the points belonging to the respective clusters. Thus, the number of patterns to be presented to the network in one epoch becomes $\frac{s(s_c-1)}{2} = \frac{N^2(c-1)}{2}$, where $s = |S|$ and $s_c = |S_c| << s$.

Note that the clustering algorithm at this stage provides a first hand knowledge about the various landcovers in a remotely sensed image. Based on this knowledge, neuro-fuzzy feature selection algorithm is applied, for removing the redundant and confusing features, and thereby improving the segmentation result. Thus the algorithm for finding optimal $\boldsymbol{W}$ is as follows.

Algorithm for learning $\boldsymbol{W}$

- Cluster the patterns using an appropriate algorithm (*e.g. k-means* clustering algorithm) and the cluster centers $\boldsymbol{cen}_q$'s are noted.
- Calculate d_{max} from the unlabeled training set and store it in both the output nodes. Store β (user specified) in both the output nodes.
- Initialize W_j with small random values in $[0, 1]$.
- Repeat until convergence, *i.e.*, until the value of E becomes less than or equal to certain predefined small quantity, or number of iterations attains certain predefined number of iterations:
  - For each pair of patterns $\boldsymbol{x}_p$ and $\boldsymbol{cen}_q, \forall p, q$:
    - Present the pattern pair to the input layer.
    - Compute $\triangle W_j$ for each j using the updating rule in (14).
  - Update W_j for each j with $\triangle W_j$ averaged over all the pattern pairs.

E (10), after convergence, attains a local minimum and then the weights ($W_j = w_j^2$) of the links connecting hidden nodes and the output node computing μ^T-values, indicate the order of importance of the features. Note that this unsupervised method performs the task of feature selection without clustering the feature space explicitly and does not need to know the number of clusters present in the feature space.

VI. Segmentation and Quality measure

Having selected the features, the main task is to integrate these features to produce a segmentation. We define a scale-space signature as the vector of features at different scales taken at a single pixel in an image,

$$\boldsymbol{Feat}(i,j) = [Feat_0(i,j), Feat_1(i,j), Feat_2(i,j), \ldots, Feat_k(i,j)]$$

These scale-space signatures are considered as feature vectors in a feature space.

Let us assume that there are M texture categories, $C_1, \ldots, C_c$, present in the image. If our texture features that have already been obtained are capable of discriminating these categories then the patterns belonging to each category will form a cluster in the feature space which is compact and isolated from clusters corresponding to other texture categories. Segmentation algorithm accept as input a set of features and put a class labeling for each pixel. Fundamentally this can be considered a multidimensional data clustering problem. We have used a traditional *k-means* clustering algorithm [50].

Quantitative performance measure :

A quantitative index β [38] is used to evaluate the segmentation results. β is defined as the ratio of the total variation and within class variation. Let n_i be the number of pixels in the ith ($i = 1, 2, \ldots c$) region. X_{ij} is the gray value of the jth pixel ($j = 1, 2, \ldots n_i$) in the ith region, and $\overline{X_i}$ is the mean of n_i gray values in the ith region. Then β is defined as

$$\beta = \frac{\sum_{i=1}^{c} \sum_{j=1}^{n_i} (X_{ij} - \overline{X})^2}{\sum_{i=1}^{c} \sum_{j=1}^{n_i} (X_{ij} - \overline{X_i})^2} \tag{15}$$

where, n is the size of the image and $\overline{X}$ is the mean gray value of the image.

The numerator is constant for a given image and number of classes present in the image. Therefore, the value of β is dependent only on the denominator. The denominator, on the other hand, decreases with homogeneity of a region. So higher the value of β, better is the segmentation.

VII. Segmentation of *IRS* and *SPOT* images

In this section, the results on implementation of our algorithm on several remotely sensed images are provided. The images considered are two IRS-1A images and one SPOT image.

A. Data sets

IRS images :

The IRS-1A images (Figures 2 and 6) were obtained from Indian Remote Sensing Satellite which is a circular sun-synchronous satellite, rotating around the earth at the rate of 14 orbits per day, at an altitude of 904 km and a repetition cycle of 22 days (NRSA 1986). This satellite is equipped with two different sensors LISS (Linear Imaging Self Scanner)-I and LISS-II. LISS-I has a spatial resolution of 72.5m $\times$ 72.5m while that for LISS-II is 36.25m $\times$ 36.25m. The IRS-1A images used for this work were taken using the scanner LISS-II in the wavelength range $0.45\mu m$ - $0.86\mu m$. The whole spectrum range is decomposed into four spectral bands namely,

band 1 - blue band of wavelength 0.45μm - 0.52 μm,
band 2 - green band of wavelength 0.52μm - 0.59 μm,
band 3 - red band of wavelength 0.62μm - 0.68 μm and
band 4 - near infra red wavelength 0.77μm - 0.86 μm.

The images in Figures 2(a)-2(d) cover an area around the city of Calcutta in these four bands. In all these figures the prominent black stretch across them is the river *Hoogly*. There is a prominent light patch on the bottom right corner, which is the *Salt Lake stadium* and the black patches nearby are the *fisheries*. In the upper right part of the images, there is a distinct line structure corresponding to the *airport runway*. In total there are five major classes in which the regions of Calcutta IRS images can be classified. These are *water bodies* (WB), *vegetation* (VEG), *habitation* (HAB), *city area* (CA) and *open spaces* (OS).

Figures 6(a)-6(d) show a part of the city of Bombay in the four aforesaid bands respectively. The elongated city area is surrounded by the Arabian sea. There is a *concrete structure* (on the right side top corner) connecting Bombay to New Bombay. On the southern part of the city, there are several islands including the famous *Elephanta islands*. The *dockyard* is situated on the south-eastern part of Bombay, which can be seen as a set of three finger like structure. On the upper part of the images, towards left, there is a distinct crisscrossed structure, which corresponds to the *Santa Cruz airport*. In total the regions of Bombay images can be classified into six major classes, *viz*., *turbid water 1* (TW1), *turbid water 2* (TW2), *concrete* (CONCR), *habitation* (HAB), *vegetation* (VEG) and *open spaces* (OS). The sea water is decomposed into two classes TW1 and TW2 for better classification since they have somewhat different reflectance properties due to variation in sea water density (as seen in Figure 6(d)).

SPOT image :

The SPOT image was taken by the French Satellites SPOT (Systems Probataire d' Observation de la Terre) [44], launched in 1986 and 1990. They carry two imaging devices that consist of a linear array of charge coupled device (CCD). Two imaging modes are possible, the multispectral and panchromatic modes. The image considered here has three bands in the multispectral mode. These are:
band 1 - green band of wavelength 0.50μm - 0.59 μm,
band 2 - red band of wavelength 0.61μm - 0.68 μm and
band 3 - near infra red band of wavelength 0.79μm - 0.89 mum.

Figures 10(a)-10(c) are SPOT images of Calcutta and were taken in these three bands with a spatial resolution of $20.0m \times 20.0m$. In all these figures the prominent black stretch across them corresponds to the river *Hoogly*. There are two distinct black, elongated patches below the river, on the left side of the images. These are water bodies, the one to the left being *Garden Reach Lake* and the one to the right being the *Khiderpore dockyard*. Just to the right of these water bodies, there is a thin line starting from the right bank of the river and going to the bottom edge of the images, which corresponds to *Talis nala*. On the right side of the images, there is a triangular patch which is the *race course*. On the top right hand side of the images, there is a thin line, stretching from the top edge and ending in the middle of the image, which corresponds to the *Beleghata canal*. A bridge called the *Rabindra setu*, cuts the river near the top of the images. This image has seven major classes, which are *turbid water* (TW), *pond water* (PW), *concrete* (CONCR), *vegetation* (VEG), *habitation* (HAB), *open space* (OS) and *roads/bridges* (B/R).

All the images considered in this investigation are of size 512×512. Due to poor illumination, the actual object classes present in the input images are not visible clearly. So we have presented the histogram equalized images in Figures 2(a)-2(d), 6(a)-6(d) and 10(a)-10(c), which highlight the different landcover regions. But the algorithms were implemented on the actual inputs.

B. Results

Here we demonstrate the effectiveness of our methodology over several remotely sensed images as described in Section VII-A. In order to validate the importance of neuro-fuzzy feature evaluation, we show how the feature dimensionality can be greatly reduced after feature evaluation. The test images have several fine (line) structures (roads and bridges). In order to detect these structures the window size has been kept small (3×3).

The total number of features considering all the bands are found to be 10 using the proposed feature extraction methodology described in Section IV. Figure 3 shows the segmented output of the IRS-1A image using all these 10 features considering the number of classes to be $c = 5$. The β index is found to be 3.65887. The neuro-fuzzy feature evaluation algorithm reduces the number of features to only one for which the segmentation result is given in Figure 4. The value of β index is found to be 3.83431 signifying an improvement in segmentation quality. The *stadium* and *fisheries* as well as the *airport runway* are quite distinctly discernible in the segmented output (Figure 4) as compared to Figure 3. We can comment on this finding that since we have considered all the 4 bands for feature extraction, most of the features so extracted contain very less information and hence can be discarded. It can also be inferred that out of these 4 bands only one band is significant and contains most of the image information. It is also evident from the Figures 2(a)-2(d) that the image corresponding to the near infra red band (Figure 2(d)) is the most significant band. The feature with which we have obtained the above segmentation result after neuro-fuzzy evaluation corresponds to band-4 and validates our comment. So we can easily disregard all the other bands and keep only the band-4 image for feature extraction. Comparing Figures 3 and 4 we find that we get better segmentation output in the *saltlake stadium* area, also the compactness of the various classes present in the IRS-1A Calcutta image increases after the feature evaluation step. For a comparative study we present the segmentation result (Figure 5) obtained by using *k-means* clustering algorithm with average and busyness computed over 3×3 neighborhood [46] incorporating local information, as features. The figure shows clear improvement in segmentation output using our methodology both quantitatively (β value) and qualitatively (structural details).

The segmentation result of the IRS-1A image of Bombay is shown in Figure 7 considering number of classes to be

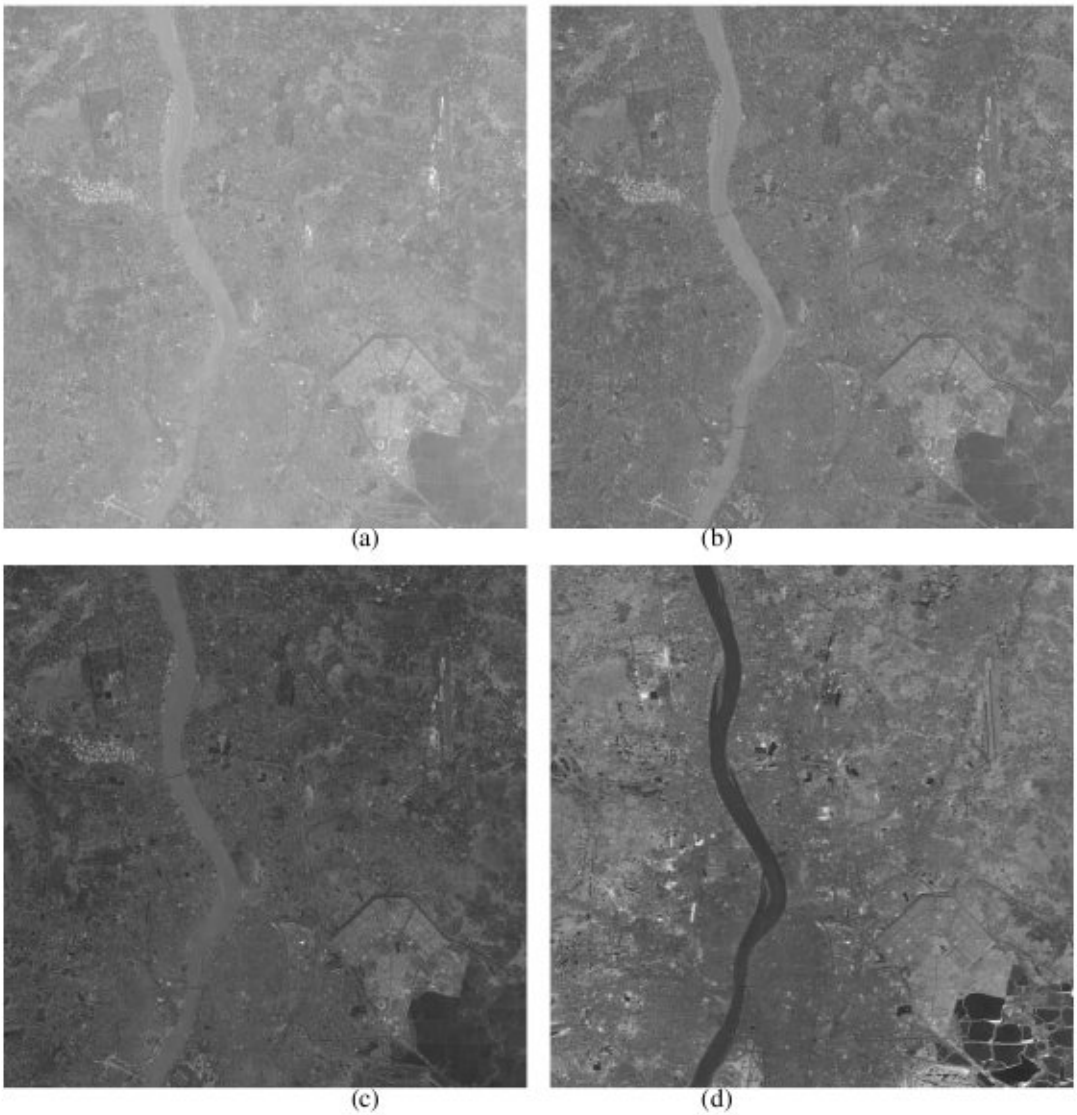

Fig. 2. IRS-1A image of Calcutta a) band-1 b) band-2 c) band-3 d) band-4

c=6, and the β value is found to be 8.01462. The number of extracted features from all the 4 bands is found to be 12 as obtained by Section IV. The feature dimensionality is reduced to one after neuro-fuzzy feature evaluation. As in the previous case, β index (8.21309) shows better segmentation quality. In the segmented output after the feature evaluation step (Figure 8), it can be seen that the *dockyard*, the *concrete structure* connecting Bombay and New Bombay and *Santa Cruz airport* are very well detected. Even for this image of Bombay we can reiterate the above comments. Here also the feature obtained after the feature evaluation step corresponds to band-4 image and contains most of the image information. Comparing Figures 7 and 8 we find that various objects and regions (*santa cruz airport, dockyard, roads and bridges* etc.) are identified more prominently. Moreover, the feature evaluation step increases the compactness of the various classes similar to the case of IRS-1A Calcutta image. In this case also a comparative study is presented. The segmentation result in Figure 9 as mentioned above obtained by using *k-means* clustering algorithm with average and busyness as features. Here also we can find an improvement in segmentation.

In the case of SPOT image of Calcutta (Figures 10(a)-10(c)), we have initially considered all the 10 features extracted by the wavelet based method. The segmentation result is given in Figure 11 with c=7 and a β value of 3.27832. After the neuro-fuzzy feature evaluation the number of features reduces to three. The segmentation result is given in Figure 12, and β value (3.45361) indicates better segmentation quality. It is to be noted that not only the *race course* but also a triangular outline (which is an open space) corresponding to the *track* of the *race course* are also detected. Here the *Rabindra Setu* has also been detected along with the *Talis nala* and *Beleghata canal*. Moreover, the *Khiderpore dock* and *Garden Reach Lake* are quite discernible. In this case all the three bands are important and furnish some information. The features obtained after the feature evaluation step correspond to each of the three bands. Comparing Figures 11 and 12, it can be seen that the *race course* can be identified quite vividly (Figure 12). In this case also the compactness of the various classes increases after the feature evaluation step. The segmentation result along with

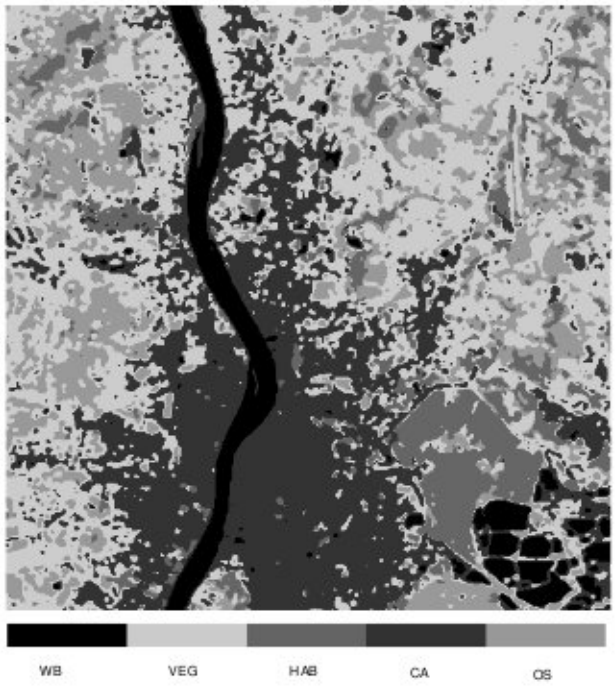

Fig. 3. Segmented output of Calcutta IRS-1A image with $c = 5$ and $\beta = 3.65887$

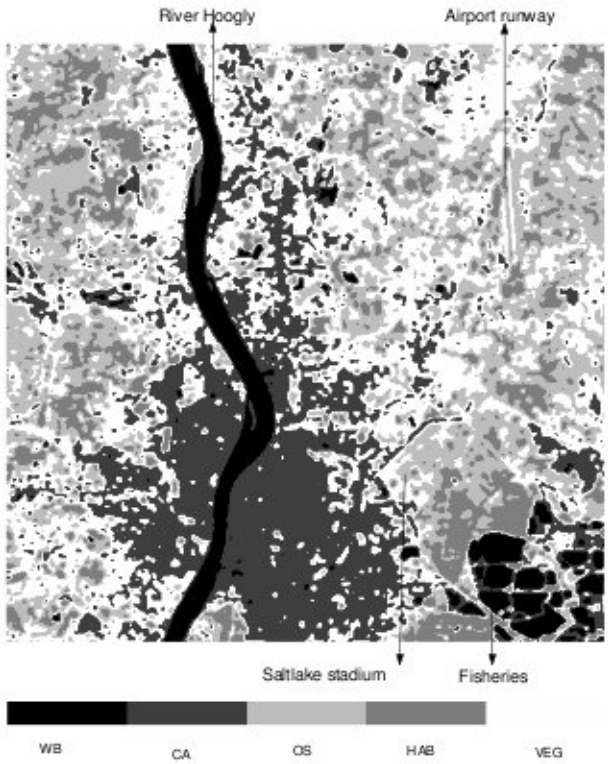

Fig. 4. Segmented output of Calcutta IRS-1A image with $c = 5$ and $\beta = 3.84578$

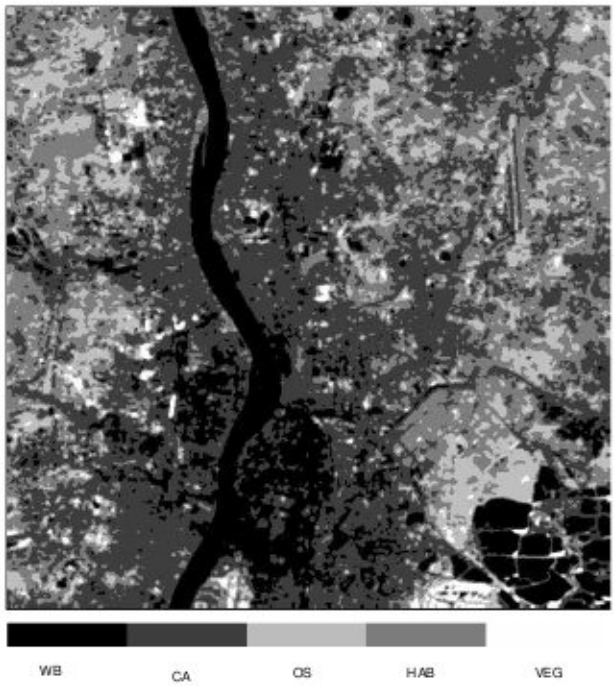

Fig. 5. Segmented output of Calcutta IRS-1A image using average and busyness as features with $c = 5$ ($\beta = 2.84578$)

the β value obtained by using *k-means* clustering algorithm with average and busyness as features (as mentioned above) is given in Figure 13 for a comparative study validating that the features extracted are characterize the remotely sensed images in a better way.

In order to prove the efficacy of our algorithm we have tested it on a synthetic 10-texture mosaic of size 256×512 comprising of ten different Brodazt textures (Figure 14(a). The segmented output is presented in Fig. 14(b). Note that although the image contains 10 different Brodazt textures, some of them are not distinctly identifiable visually as depicted in Fig. 14(a). Interestingly, the proposed methodology has been able to identify more or less all the classes. Here the percentage of correctly classified pixels has been found to be 79.55%.

As part of the investigation, an extensive comparison has been made to show the superiority of the proposed methodology over various existing related algorithms. Note that several approaches to multichannel filtering for texture segmentation have been studied and a comparative performance evaluation is presented by Randen and Husϕy in [41].

TABLE I
PERFORMANCE EVALUATION OF $Nat10a$

Methods/filters		Classification	Number of features
Proposed method M-band wavelet	without feature evaluation	71.7%	19
	with feature evaluation	79.5%	7
Heuristic Gabor filter bank (d) [40]		67.7%	20
Optimized Optimized representation Gabor filter [9] [8]		64.1%	15
Full rate QMF FIR filter $f8a$ (d) [40]		60.3%	40
Critically sampled QMF FIR $f8a$ (d) [40]		57.7%	13

VIII. CONCLUSION

In this article, we have presented a scheme to show how wavelet theory and neuro-fuzzy hybridization together can be applied in the domain of remotely sensed imagery for segmentation purpose. The use of wavelet theory via M-band wavelet decomposition of remotely sensed images provides an efficient representation of these images in terms of frequencies in different directions and orientations at different resolutions. This representation is obtained by segmenting some wavelet features from multispectral remotely sensed images. This facilitates an improved segmentation of the different class regions. The feature extraction method splits the lower as well as the higher frequency bands, and results in a tree structure. This enables the system to extract their characteristic features in the lower as well as the higher frequency bands of remotely sensed images. The neuro-fuzzy feature evaluation method helps in searching for important features efficiently from a remotely sensed image where the various classes are overlapping in nature. Both the feature extraction and neuro-fuzzy feature evaluation schemes are unsupervised and do not require any *a priori* knowledge about the number of classes, and spatial relationship of different regions in these images. Note that the neuro-fuzzy feature evaluation method presented here is a modified version of an existing one described in [36]. This modification helps in dealing with large data set efficiently. We have tested our algorithm on synthetic data comprising of natural textures. The results validates that our methodology is indeed superior to some of the related ones existing in the literature. Also the results presented using average and busyness as features and using *k-means* clustering algorithm for a comparative study show that our methodology is indeed

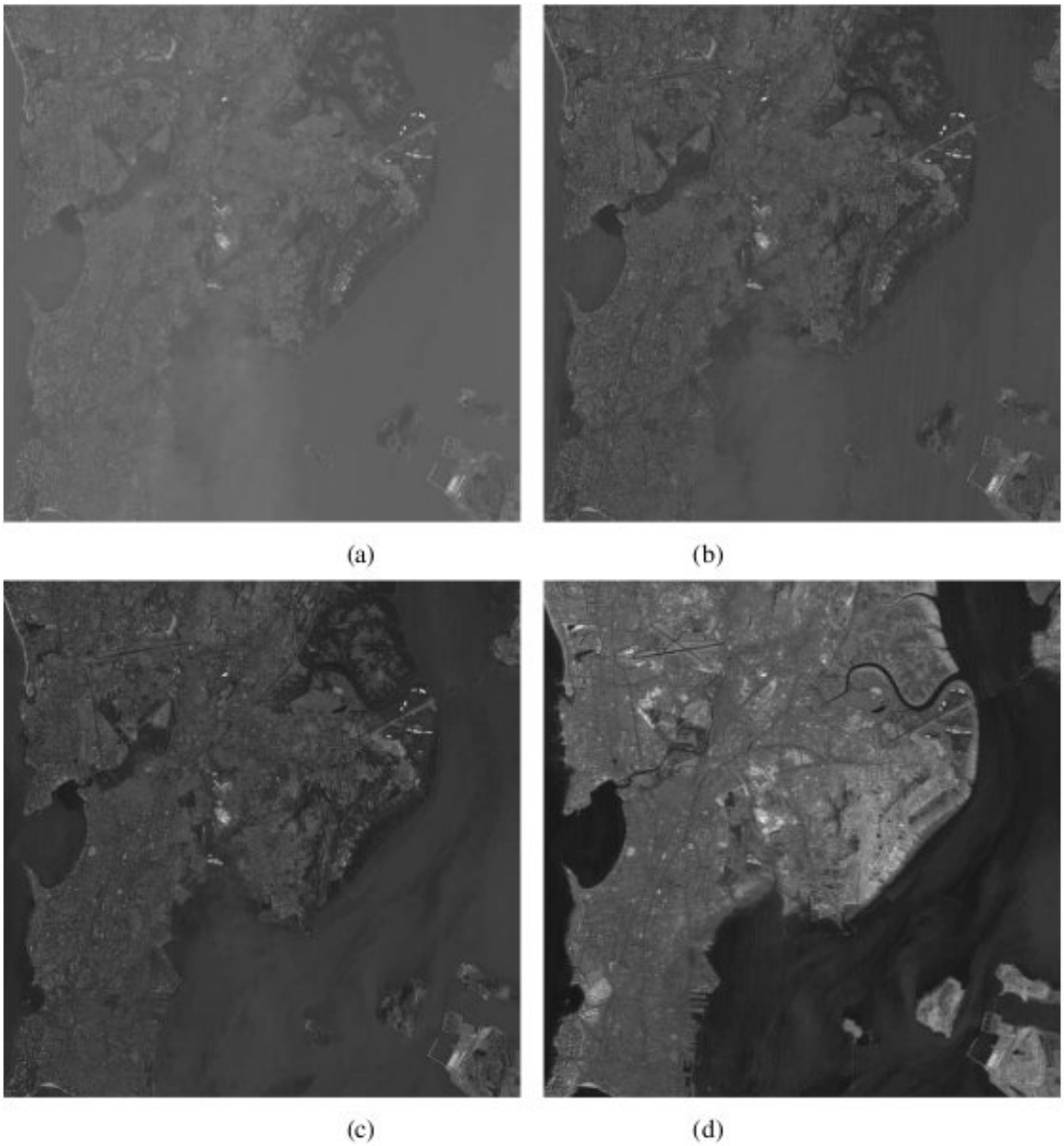

Fig. 6. IRS -1A image of Bombay a) band-1 b) band-2 c) band-3 d) band-4

effective in characterizing remotely sensed images in a better way.

The features obtained by the feature extraction method have been able to segment the remotely sensed images satisfactorily. It is been found that for the IRS-1A images that ultimately one feature is selected which corresponds to the band 4 of these images. However, for the SPOT image three features are selected. Almost all the desired classes are obtained in the segmented output satisfactorily. Moreover, the compactness of the various classes increases after the neuro-fuzzy feature evaluation step.

REFERENCES

[1] M. Acharyya, R.K. De, and M. K. Kundu. Extraction of features using m-band wavelet packet frame and their neuro-fuzzy feature evaluation for multitexture segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(12), 2003.

[2] M. Acharyya, R.K. De, and M. K. Kundu. Segmentation of remotely sensed images using wavelet features and their evaluation in soft computing framework. *IEEE Trans. on Geoscience and Remote Sensing*, 41(12), 2003.

[3] M. Acharyya and M. K. Kundu. Adaptive basis selection for multitexture segmentation by m-band wavelet packet frame. In *Proc. 2001 Int. Conf. Image Processing*, pages 622–625, Sept. 2001.

[4] M. Acharyya and M. K. Kundu. Wavelet-based texture segmentation of remotely sensed images. In *Proc. 2001 Int. Conf. Image Analysis and Processing*, pages 69–74, Sept. 2001.

[5] O. Alkin and H. Caglar. Design of efficient m-band coders with linear phase and perfect reconstruction properties. *IEEE Trans. on Signal Processing*, 43(7):1579–1590, 1995.

[6] S. Bandyopadhyay and S. K. Pal. Pixel classification using variable string genetic algorithms with chromosome differentiation. *IEEE Trans. on Geoscience and Remote Sensing*, 39(2):303–308, 2001.

[7] J. R. Bergen and M. S. Landy. Computational modeling of visual texture segregation. In M. S. Landy and J. A. Movshan, editors, *Computational Models of Visual Processing*, pages 472–481. The MIT Press, Cambridge, MA, 1991.

[8] A. C. Bovik. Analysis of multichannel narrow-band filters for image texture segmentation. *IEEE Trans. Signal Processing*, 39:2025–2043, Sept. 1991.

[9] A. C. Bovik, M. Clark, and W. S. Geisler. Multichannel texture analysis

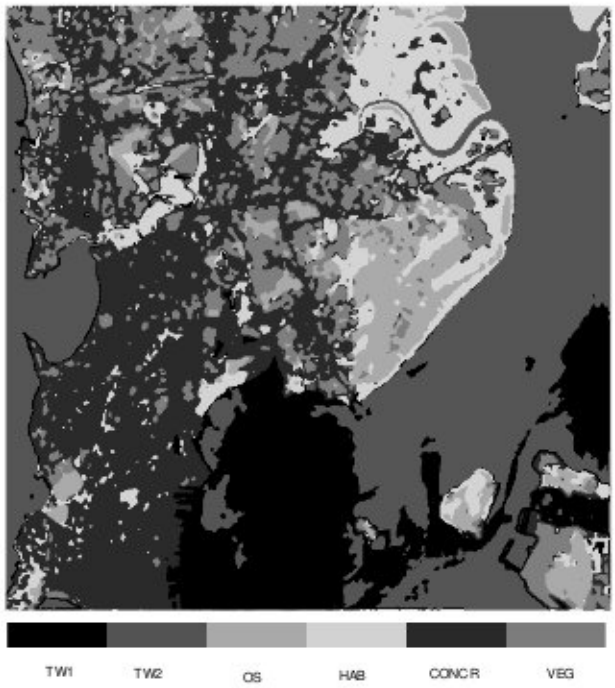

Fig. 7. Segmented output with $c = 6$ and $\beta = 8.01462$

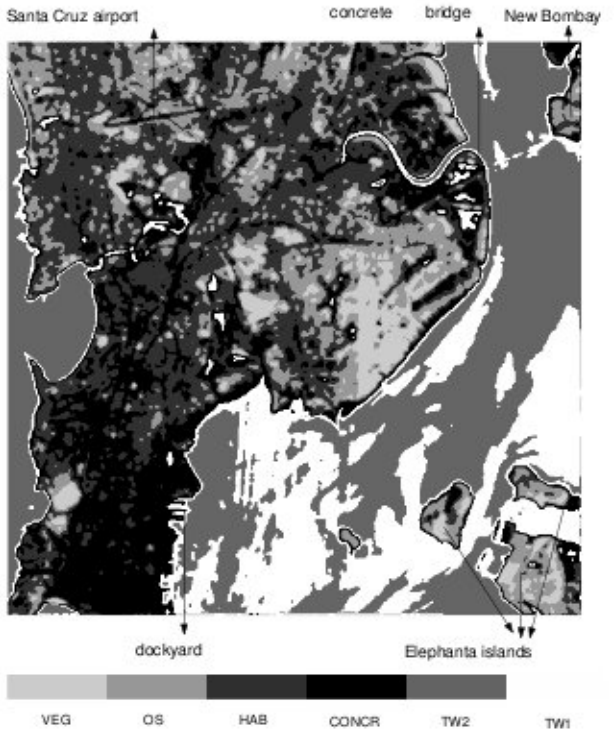

Fig. 8. Segmented output with $c = 6$ and $\beta = 8.21309$

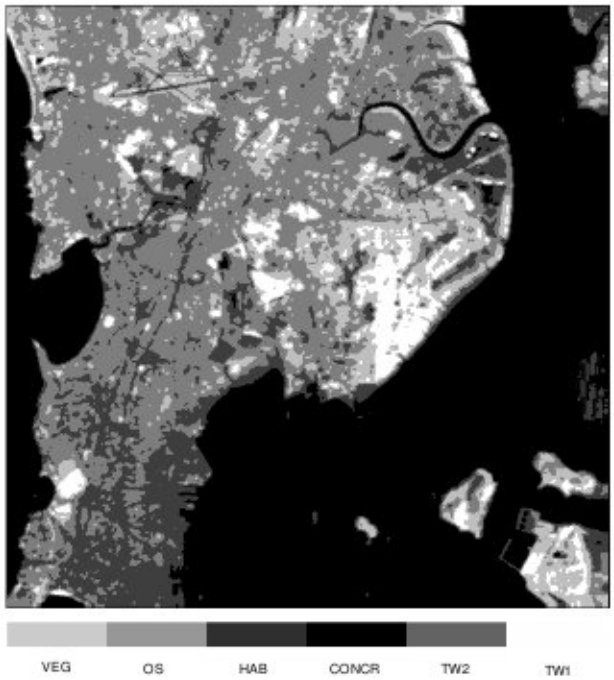

Fig. 9. Segmented output with $c = 6$ with average and busyness as features ($\beta = 7.21309$)

using localized spatial filters. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12:55–73, 1990.

[10] P. H. Carter. Texture discrimination using wavelets. In *SPIE Applications of Digital Image Processing XIV*, volume 1567, pages 432–438, 1991.

[11] C. C. Chen and D. C. Chen. Multiresolution gabor filters in texture analysis. *Pattern Recognition Letters*, 17(10):1069–1076, 1996.

[12] P. C. Chen and T. Pavlidis. Segmentation by texture using correlation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 5:64–69, Jan 1983.

[13] G. C. Cross and A. K. Jain. Markov random field texture models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 5:25–39, 1983.

[14] I. Daubechies. Orthogonal bases for compactly supported wavelets. *Comm. Pure Appl. Math*, 41:909–996, 1988.

[15] I. Daubechies. *Ten Lectures on Wavelets*. Soc. Ind. Applied Math, Philadelphia, 1992.

[16] Judith E. Dayhoff. *Neural Network Architectures An Introduction*. Van Nostrand Reinhold, New York, 1990.

[17] L. J. Du. Texture segmentation of sar images using localized spatial filtering. In *Proc. of Intl. Geoscience and Remote Sensing Symposium*, pages 1983–1986, 1990.

[18] S. Fukuda and H. Hirosawa. A wavelet-based texture feature set applied to classification of multifrequency polarimetric sar images. *IEEE Trans. on Geoscience and Remote Sensing*, 37(5):2282–2286, 1999.

[19] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Pattern Anal. and Machine Intell.*, 6:721–741, 1984.

[20] R. M. Haralick. Statistical and structural approaches to texture. *Proc. IEEE*, 67:768–804, 1979.

[21] R. M. Haralick, K. Shanmugam, and I. Dinstein. Texture feature for image classification. *IEEE Trans. on Systems, Man and Cybernatics*, 8(6):610–621, 1973.

[22] H.Zou and A. H. Tewfik. Discrete orthogonal, m-band wavelet decomposition. In *Proc. Intl. Conf. on Acoustics, Speech, and Signal Processing*, volume 4, pages IV–605–IV–608, 1992.

[23] A. K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, 24(12):1167–1186, 1991.

[24] R. L. Kashyap and A. Khotanzad. A model-based method for rotation invariant texture classification. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(4):472–481, 1986.

[25] N. G. Kingsbury. Complex wavelets and shift invariance. In *Proc. IEE Colloquium on Time-Scale and Time-Frequency Analysis and Applications*, 2000.

[26] M. K. Kundu and M. Acharyya. M-band wavelets : Application to texture segmentation for real life image analysis. *International Journal of Wavelets, Multiresolution and Information Processing*, 1(1):115–149, 2003.

[27] K. L. Laws. Rapid texture identification. In *Proc. SPIE*, volume 238, pages 376–380, 1980.

[28] R. W. Lindsay, D. B. Percival, and D. A. Rothrock. The discrete wavelet transform and the scale analysis of the surface properties of sea ice. *IEEE Trans. on Geoscience and Remote Sensing*, 34(3):771–787, 1996.

[29] Richard P. Lippmann. An introduction to computing with neural nets. *IEEE ASSP Magazine*, 4(2):4–22, 1987.

[30] S. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[31] S. Mallat and Sifen Zhong. Characterization of signals from multiscale edges. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(7):710–732, 1992.

[32] B. S. Manjunath and R. Chellappa. A computational approach to boundary detection. *Proc. CVPR*, pages 358–363, 1991.

[33] A. Mecocci, P. Gamba, A. Marazzi, and M. Barni. Texture segmentation in remote sensing images by means of packet wavelets and fuzzy clustering. In *Proc. of the European Symposium on Satellite and Remote Sensing II*, volume SPIE 2584, pages 142–157, Sept. 1995.

[34] D. Muchoney and J. Williamson. A gaussian adaptive resonance theory neural network classification algorithm applied to supervised land cover mapping using multitemporal vegetation index data. *IEEE Trans. on Geoscience and Remote Sensing*, 39:1969–1977, 2001.

[35] A. Niedermeier, E. Romaneesen, and S. Lehner. Detection of coastline sar images using wavelet methods. *IEEE Trans. on Geoscience and Remote Sensing*, 38(5):2270–2281, 2000.

[36] S. K. Pal, R. K. De, and J. Basak. Unsupervised feature evaluation : A neuro-fuzzy approach. *IEEE Trans. on Neural Networks*, 11:366–376, 2000.

[37] S. K. Pal, A. Ghosh, and M. K. Kundu, editors. *Soft Computing for Image Processing*. Physica Verlag, Heidelberg, 2000.

[38] S. K. Pal, A. Ghosh, and B. Uma Shankar. Segmentation with remotely sensed images with fuzzy thresholding, and quantitative evaluation. *Int. Journal of Remote Sensing*, 21(11):2269–2300, 2000.

[39] S. K. Pal and S. Mitra. *Neuro-fuzzy Pattern Recognition: Methods in Soft Computing*. John Wiley, New York, 1999.

[40] T. Randen. *Filter and Filter Bank Design for Image Texture Recognition*. PhD thesis, Norwegian Univ. of Science and Technology, Trondheim, Norway, 1997.

[41] T. Randen and J. H. Husøy. Filtering for texture classification : a comparative study. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21:291–310, April 1999.

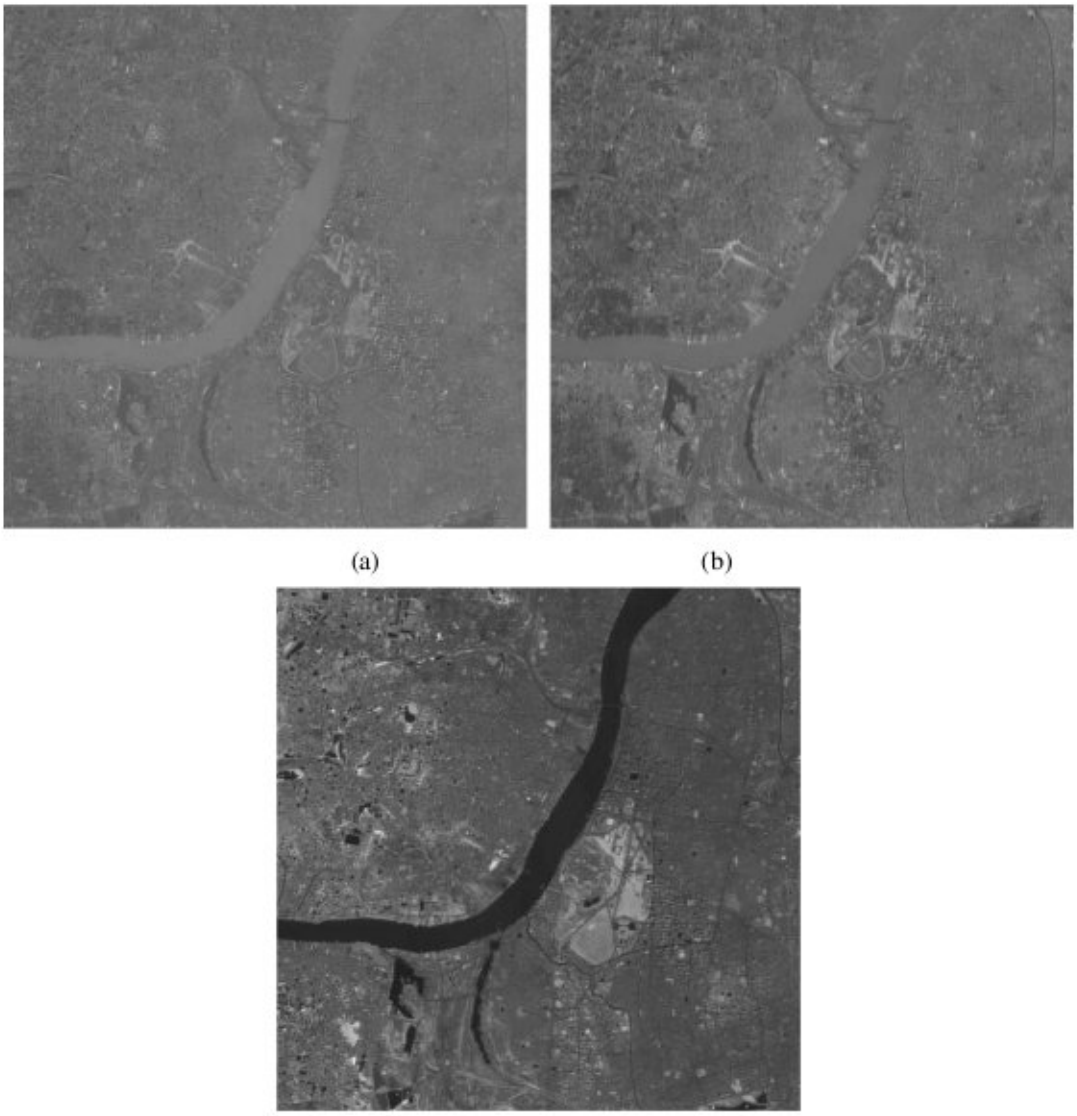

Fig. 10. SPOT image of Calcutta a) band-1 b) band-2 c) band-3

[42] T. R. Reed and J. M. H. du Buf. A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image Understanding*, 57(3):359–372, 1993.

[43] T. R. Reed and H. Wechsler. Segmentation of textured images and gestalt organization using spatial/ spatial frequency representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12:1–12, Jan 1990.

[44] J. A. Richards. *Remote Sensing Digital Image Analysis: An Introduction*. Springer-Verlag, New York, 1993.

[45] E. Rignot and R. Kowk. Extraction of textural features in sar images: Statistical model and sensitivity. In *Proc. of IEEE Geoscience and Remote Sensing Symposium*, May 1990.

[46] A. Rosenfeld and A. C. Kak. *Digital Image Processing*. Academic Press, New York, 2nd edition, 1982.

[47] M. Simard, S. S. Saatchi, and G. De Grandi. The use of decision tree and multiscale texture for classification of jers-1 sar data over tropical forest. *IEEE Trans. on Geoscience and Remote Sensing*, 38(5):2310–2321, 2000.

[48] P. Steffen, P. N. Heller, R. A. Gopinath, and C. S. Burrus. Theory of regular m-band wavelet bases. *IEEE Trans. on Signal Processing*, 41(12):3497–3510, 1993.

[49] P. Thitimajshima. Multiresolution fuzzy clustering for sar image segmentation. In *Proc. of the Intl. Geoscience and Remote Sensing Symposium, IGARSS '99*, volume 5, pages 2507–2509, 1999.

[50] J. T. Tou and R. C. Gonzales. *Pattern Recognition Principles*. Adision-Wesley, Reading, 1974.

[51] D. C. Tseng, H. M. Tsai, and C. C. Lai. Unsupervised texture segmentation for multispectral remote-sensing images. In *Proc. of the Intl. Conf. on Pattern Recognition, ICPR'98*, pages 1630–1632, 1998.

[52] M. Tuceryan and A. K. Jain. Texture analysis. In C. H. Chen, L. F. Pau, and P. S. P. Wang, editors, *Handbook of Pattern Recognition and Computer vision*, chapter chapter 2, pages 235–276. World Scientific, Singapore, 1993.

[53] M. R. Turner. Texture discrimination by gabor function. *Biological Cybernetics*, 55:71–82, 1986.

[54] M. Unser. Texture classification and segmentation using wavelet frames. *IEEE Trans. on Image Processing*, 4(11):1549–1560, 1995.

[55] M. Unser and M. Eden. Multiresolution feature extraction and selection for texture segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11:717–728, 1989.

[56] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[57] L.A. Zadeh. Fuzzy logic, neural networks, and soft computing. *Communications of the ACM*, 37:77–84, 1994.

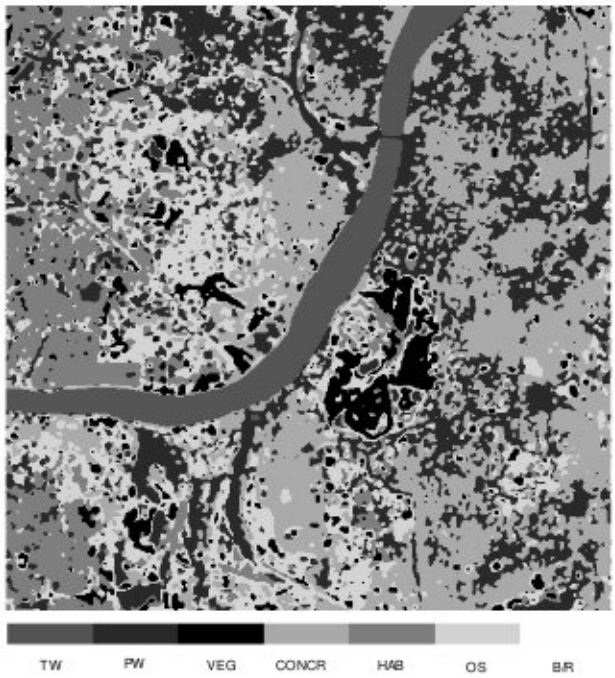

Fig. 11. Segmented output with $c = 7$ and $\beta = 3.27832$

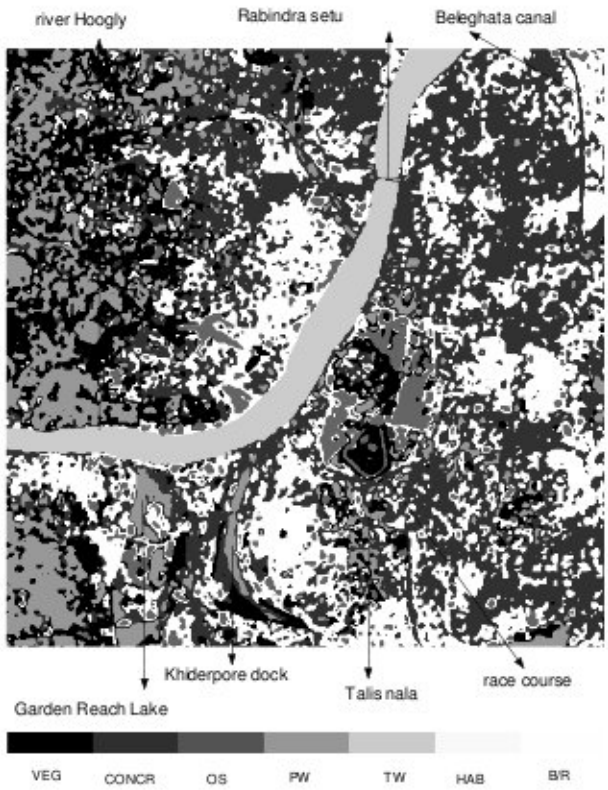

Fig. 12. Segmented output with $c = 7$ and $\beta = 3.45631$

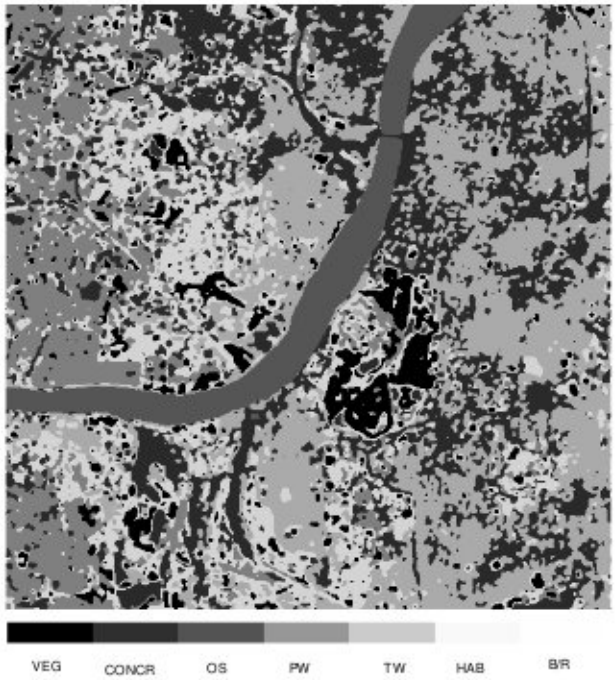

Fig. 13. Segmented output with $c = 7$ using average and busyness as features ($\beta = 2.45631$)

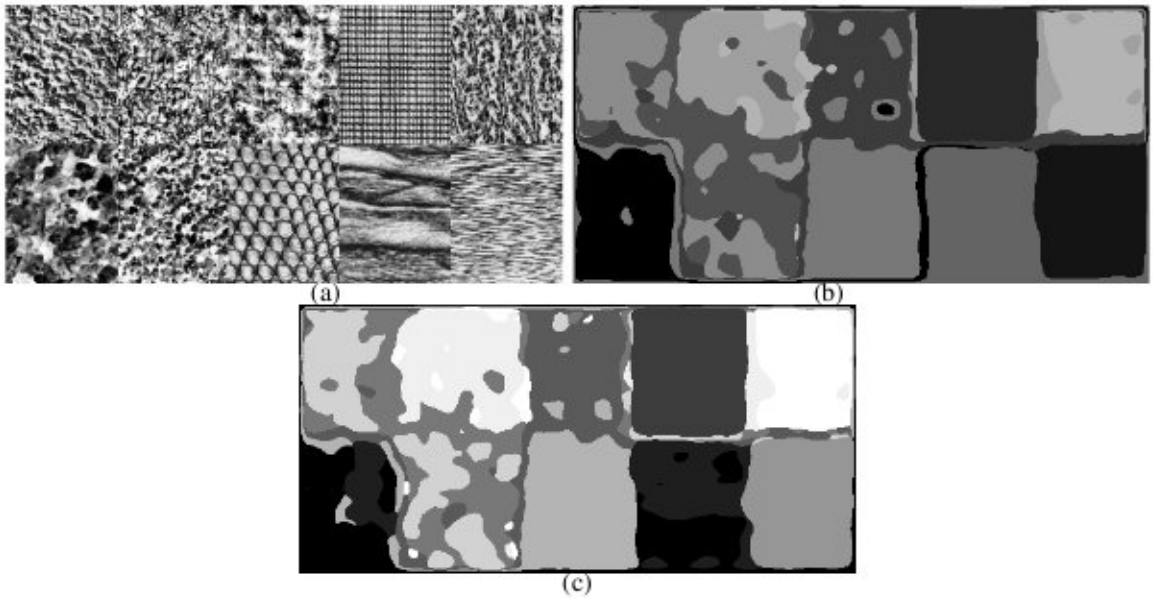

Fig. 14. (a) $Nat10a$, segmented output (b) after and (c) before neuro-fuzzy feature evaluation