KOHONEN SELF ORGANIZING FOR AUTOMATIC IDENTIFICATION OF CARTOGRAPHIC OBJECTS

J. Amini and M.R. Seradjian

Department of Surveying Engineering, Faculty of Engineering, Tehran University
Tehran, Iran, Fax: 8008837, jamini@geomatics.ut.ac.ir - mrsaradjian@ geomatics ut.ac.ir

(Received: April 24, 2000 – Accepted in Revised Form: January 30, 2002)

Abstract Automatic identification and localization of cartographic objects in aerial and satellite images have gained increasing attention in recent years in digital photogrammetry and remote sensing. Although the automatic extraction of man made objects in essence is still an unresolved issue, the man made objects can be extracted from aerial photos and satellite images. Recently, the high-resolution satellite images, typically at most 3 meters in panchromatic band ground sample distance (GSD) and up to four multispectral bands in the visible and near infrared spectrum, are suitable for detection and identification of objects. This paper presents a new algorithm for identification of cartographic objects based on Artificial Neural Network (ANN). The algorithm is divided in two modules: image simplification by the Wavelet transform, Mathematical Morphology (MM) operators, and identification of object by the Kohonen Self Organizing Map (KSOM) and split and merge method. The study area included two parts of an orthoimage from Kish, Iran.

Key Words Road, Classification, Extraction, Kohonen, Morphology, Wavelet, Self-Organizing

چکیده امروزه در فتوگرامتری و سنجش از دور استخراج و واضح نمودن عوارض کارتوگرافی از تصاویر هوایی و ماهواره ای مورد توجه است. هر چند استخراج کاملا اتوماتیک عوارض ساخت بشر هنوز به عنوان یک مساله غیرقابل حل مطرح است ولی این عوارض از عکسهای هوایی و تصاویر ماهواره ای می تواند استخراج شود. امروزه تصاویر ماهواره ای با قدرت تفکیک بالا در باند پانکروماتیک و حداقل چهار باند چند طیفی برای استخراج و آشکارسازی عوارض مناسب است. این مقاله یک الگوریتم جدید برای وضوح عوارض بر اساس شبکه های عصبی مصنوعی بیان می‌دارد. در این الگوریتم ابتدا با بکار بردن تبدیل ویولت (Wavelet Transform) و مورفولوژی ریاضی (Mathematical Morphology) تصویر ساده گشته و سپس با استفاده از شبکه خود سازمان یافته کوهونن (Kohonen Self Organizing Network) و روش جداسازی و ادغام عارضه واضح می گردد. منطقه مورد مطالعه قسمتهایی از تصویری مربوط به جزیره کیش در ایران بوده که در آن راهها به عنوان یک عارضه واضح گشته است.

1. INTRODUCTION

There are numerous procedures in photogrammetry and remote sensing that can be applied for the purpose of integration of image and map data for change detection, automatic object extraction and so on, from images and maps.

One of the most fascinating promises of digital photogrammetry is the highly automated acquisition and updating of spatial data from images. Remarkable progress has been made in areas involving image and template matching such as automatic interior orientation, relative orientation, digital terrain modeling (DTM) and orthoimage generation. Although the current level of automation on most digital photogrammetric stations is still fairly low, a number of these developments are already available on some commercial systems [5].

In general, object extraction from images consists of two main operations [5]:

i) Identification of an object within an image, and
ii) Tracking the object by precisely determining its outline.

It is well known that there exists no universal edge detection that could be applied to a digital image for identifying and tracking object outlines

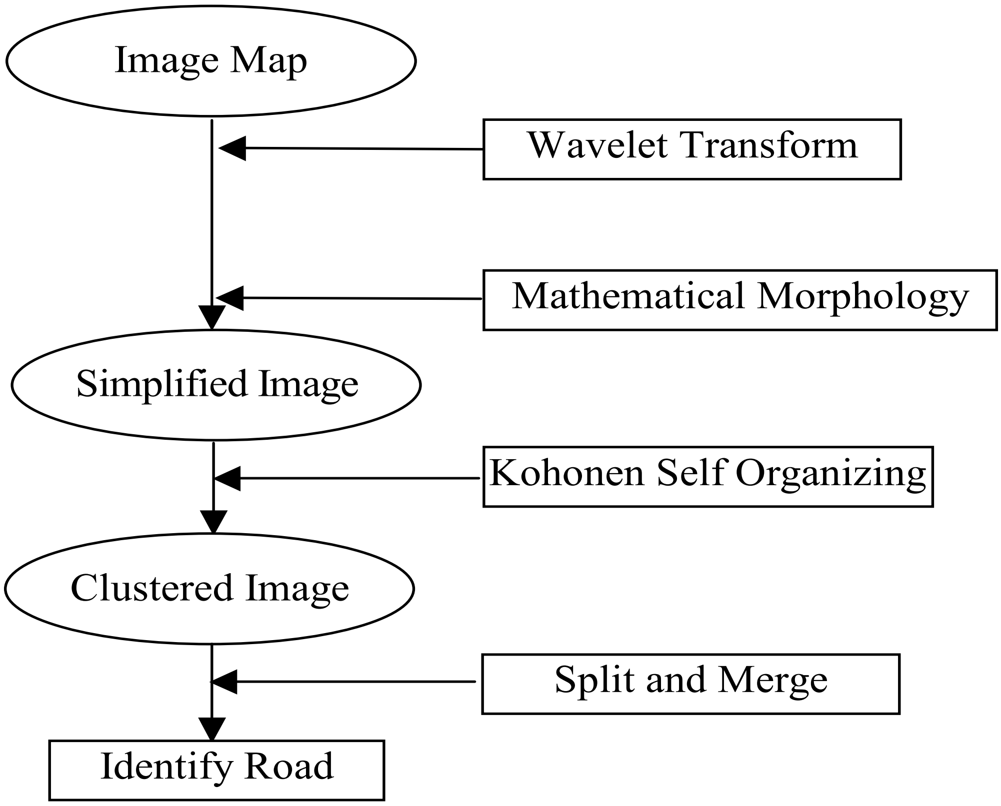

Figure 1. Strategy of object identification.

with a reasonable success. Instead, there is a tradeoff between qualitative accuracy associated with identification, and geometric accuracy associated with tracking. According to these measures, one can still classify existing operators into two broad categories [4].

Type *i* operators, which offer high accuracy for classes of objects without particularity dealing with precise outline determination, and Type *ii* operators, which don't aim at accurate identification, but offer high precision in detecting outlines when adequate approximations of the object location are available.

In an effort to optimize both measures, operators from these two classes can be combined in complex strategies for automatic object extraction [10].

According to our semi automatic object extraction philosophy, the identification task of a type *i* operator is performed manually on an image, while a special automatic digital module performs the tracking task of a type *ii* operators. In the automatic object extraction strategy, the identification is performed automatically on an image.

In this paper we used ANN architecture for object identification. Furthermore, since there are small objects such as trees, cars and shadows on the image, we used wavelet transforms and MM operators to resolve the problems. Figure 1 shows the proposed strategy used for object identification.

2. REMOVAL OF SMALL OBJECTS FROM IMAGE

In urban environments, there are many small objects (cars, trees, ...) on an image for automatic objects (such as roads and buildings) extraction.

Therefore, it is needed to simplify the image for further processing.

In this paper for removing small objects, we propose wavelet transforms and mathematical morphology. Firstly the wavelet transforms are used for decomposition of the image, then mathematical morphology operators wavelet remove small objects, and the second step used the coarse to fine strategy of wavelet transforms in order to reconstruct the simplified image.

2.1 The Wavelet Transform The Discrete Wavelet Transform (DWT) is a fast linear operation, which transforms data vectors by smoothing filters. After permuting, the result is stored in to a vector of the same length and transformed again. The wavelet transform W_h of a function *f(k)* is orthogonal and invertible:

$$W_h(f,m,n) = a_0^{-\frac{m}{2}} \sum_{k=-00}^{00} f(k)h(a_0^{-m}k - nb_0) \qquad (1)$$

In contrast to the Fast Fourier Transform (FFT), where the base functions are sine and cosine, the base functions *h(a,b,k)* of the DWT in Equation 1 are more complicated and are called "wavelets". Similar to the base functions of the FFT, the wavelets are alternating and thereby localized in frequency or (more precisely) in characteristic scale of $a = a_0^m$. Furthermore, unlike *sine* and *cosine,* wavelet functions are localized in space with translation. There are many possible bases for the wavelet space. One of these, discovered by I. Daubechies [3], has four coefficients and is often called DAUB4.

Generally, wavelets reveal similarities with short oscillating at high frequency and declining sharply at the edges. We decided to use the wavelet DAUB4 because it is very short and provides results with high accuracy. Peaks are exactly localized and the computations have good performance because only four calculations are necessary for each wavelet coefficient of $W_h(m,n)$. The DWT is a hierarchical operation based on dyadic scale, which transfers data from the signal space into the feature space. In this regard, the data vector with the length of integer power of two is correlated with the wavelet function by pushing the wavelet over the input vector with translation step of *b=2* and computing a wavelet coefficient for each step. After executing the whole input, the wavelet will be expanded by scale of *a=2* that is synonymous with compressing the data vector by the same scale. Using a smoothing filter effects this action, done by DWT.

Equation2 illustrates the principle of a hierarchical computation of $8=2^3$ wavelet coefficients which is called a pyramidal algorithm.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix} -> \begin{bmatrix} d_1 \\ s_1 \\ d_2 \\ s_2 \\ d_3 \\ s_3 \\ d_4 \\ s_4 \end{bmatrix} -> \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} -> \begin{bmatrix} D_1 \\ S_1 \\ D_2 \\ S_2 \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} -> \begin{bmatrix} S_1 \\ S_2 \\ D_1 \\ D_2 \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} \qquad (2)$$

The input vector $y = [y_1, y_2, ..., y_n]^T$ is transformed by smoothing and detail filters. After permuting the result, only the "smooth" components of low frequency, which represent the compressed input vector, are transformed again. This is repeated until only two "smooth" components remain. The other components are the wavelet (detail) components that represent position and scale of the peaks in the data vector [11].

In two dimensions, assuming the array dimension to be in power of two, an array or matrix Y_0 can be decomposed into a series of low frequency components Y_1^{LL}, Y_2^{LL}, ..., high frequency components Y_1^{HH}, Y_2^{HH}, ..., and mixed frequency components Y_1^{LH}, Y_2^{LH}, ..., and, Y_1^{HL}, Y_2^{HL}, ..., corresponding respectively to the low frequency operator L and high frequency operator H. The wavelet transforms of 2-dimensional arrays are often displayed graphically as shown in Figure 2 (with only the first three levels).

At any level, k = 0,1,2,..., the reconstruction of the corresponding image is simply the analogue of the 1-dim. case, i.e.

$$Y_{k+1}^{LL} \text{ with } Y_{k+1}^{LH}, Y_{k+1}^{HH} \text{ imply } Y_k^{LL}$$

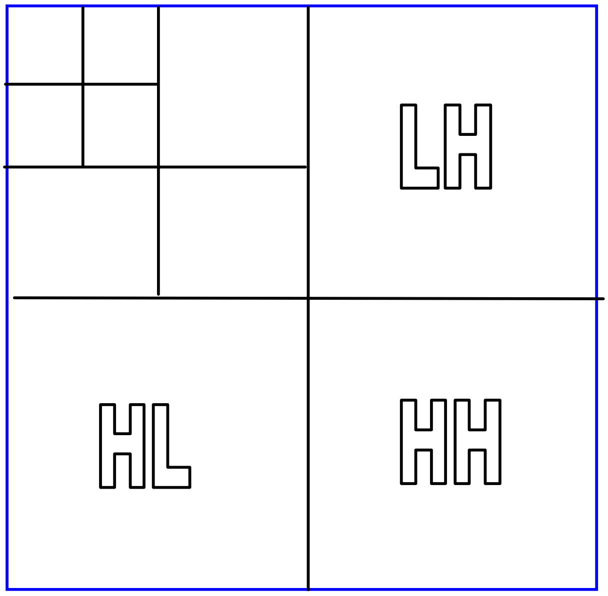

Figure 2. Quadtree structure for 2-WDT.

with the obvious difference that the four component arrays have to be combined to reconstruct the next finer level array. Each Y_k^{LL} can be used independently for some specific browsing and analysis at that scale, as they are the different level images of the image pyramid [2].

2.2 Mathematical Morphology Based on set theory, mathematical morphology provides an approach to the processing of digital image representing the geometrical structures of objects. Using appropriate sets, known as structuring elements, mathematical morphological operations can simplify image data while maintaining their shape characteristics and eliminating irrelevancies.

Maralick, et al [7] discussed the basic mathematical morphological operations and their relations in an N-dimensional properties of the basic binary (N=1) and multi-level morphological (N=2) operations with both 1-D and 2-D structuring elements. The four basic mathematical morphological operators are Dilation, Erosion, Opening and Closing that are defined as follows:

In Binary Case Consider a discrete binary image set X in an N-dimensional discrete grid Z^N. Let $T \in Z^N$ denote a structuring element, $\check{T} = \{-t / t \in T\}$ denote the symmetric set of T with respect to the origin, and ϕ denote the empty set. The translation of X bt a point $z \in Z^N$ is denoted by X_Z and defined as $X_z = \{x + z / x \in X\}$. Then the operators are defined as follows:

$$\begin{aligned}
&Dilation \quad X \oplus T = \{z / T\check{T}_z \cap X \neq \phi\} = \bigcup_{t\in T} X_t \\
&Erosion \quad X \Theta T = \{z / T_z \subseteq X\} = \bigcup_{t\in T} X_{-t} \\
&Opening \quad X \circ T = (X \Theta T) \oplus T \\
&Closing \quad X \bullet T = (X \oplus T) \Theta T
\end{aligned} \tag{3}$$

In Gray Scale Case The binary morphological operations can be extended to gray scale imagery by using the concept mentioned in Reference 7. For such images, the minimum and maximum values are computed within neighborhoods represented by the structuring element. Let F and T be the domains of the gray scale image f and the gray scale structuring elements t, respectively. The gray scale dilation and erosion can be computed by:

$$\begin{aligned}
&Dilation : (f \oplus t)(x, y) = \\
&\max_{(x-m, y-n)} \{f(x-m, y-n) + t(m, n)\} \\
&Erosion : (f \Theta t)(x, y) = \\
&\min_{(x+m, y+n)} \{f(x+m, y+n) + t(m, n)\}
\end{aligned} \tag{4}$$

In the above definition, dilation is used to fill small holes and narrow gaps in objects or expand image objects. Erosion shrinks the image objects. Opening is used to eliminate specific image details smaller than the structuring element while closing connects objects that are close to each other, fill up small holes, and smooth the object outline by filling up narrow gaps.

3. FEATURE EXTRACTION BY KOHONEN NETWORKS

The objective of the feature extraction phase is to identify the spectral classes present in the image and define the set of corresponding samples to be used in the classification phase afterwards.

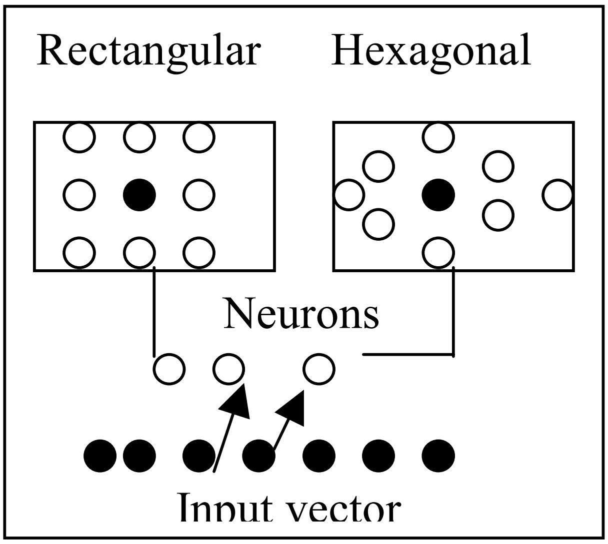

Figure 3. Geometrical representation of neurons for KSOM.

There is no well-developed theory for feature extraction, mostly features are application oriented and often found by heuristic methods and interactive data analysis.

An important basic principle is that the features must be independent of class membership because, by definition, at the feature extraction phase the membership in the classes is not yet known. This implies that any learning methods used for feature extraction should be unsupervised in the sense that the target class for each object is unknown [9].

One of the approaches is the Kohonen Self Organizing Map (KSOM) that uses competitive learning, which in turn results in data clustering [8]. The KSOM belongs to the class of unsupervised neural networks based on competitive learning, in which only one output neuron, or one per local group of neurons at a time gives the active response to the current input signal. The level of activity indicates the similarity between the input signal vector and its respective weight vector. A standard way of expressing similarity is through the Euclidean distance between these vectors.

Since the distance between the weight vector of a given neuron and the input data vector is minimal for all neurons in the network, a neuron together with a predefined set of neighbor neurons will have their weights automatically updated by the learning algorithm. The neighborhood for each neuron may be defined accordingly to the geometrical form, over which the neurons are arranged. Figure 3 depicts two examples of representation proposed by Kohonen [8]; a rectangular grid and a hexagonal grid.

A short description of the learning algorithm of KSOM is given below:

Step 1: select a training pattern $X = (x_1, x_2, x_3, ..., x_n)$ and present it as an input to the network.

Step 2: Compute distances 'd_i'between the input vector, and each j neuron weight vector, according to:

$$d_i = \sum_{j=1}^{n} (x_j(t) - w_{i,j}(t))^2 \quad (5)$$

where $x_j(t)$ is the *j*th input in a given interaction and $w_{i,j}(t)$ is the weight of neuron *j* from the input layer connected to neuron *i* from the output layer.

Step 3: Select neuron i^* with the smallest distance among all other neurons, and update the weight vector of i^* and its neighbors using the following expression:

$$\begin{aligned} &w_{i,j}(t+1) = w_{i,j}(t) + \alpha(t) * (x_j(t) - w_{i,j}(t)) \\ &for\ i \in n_{i^*}\ ; j = 1,2,...,n \end{aligned} \quad (6)$$

where n_{i^*} is a set that condition i^* and its neighbors, and $\alpha(t)$ is the learning rate, usually smaller than one. This procedure repeats until the weight updating is no longer significant. By the end of the learning process each neuron or group of neighboring neurons will represent a distinct pattern among the set of patterns presented as input to the network [6].

4. IDENTIFICATION OF OBJECTS

Using the split and merge method identifies the object after the image is clustered by KSOM in Section 2. There are various split and merge algorithms. We have shown that the "local standard deviation versus global mean algorithm" is suitable for this propose [1].

Figure 4. The original image.

5. EXPERIMENTAL RESULTS

In our study, we used a part of an image to identify roads by our algorithm.

The wavelet transform was used in the original image to construct a pyramid. The small objects such as cars, shadows, etc. on the road were taken away by using mathematical morphology operators (first opening and then closing applied) on the second coarse layer of pyramid. Then by using inverse wavelet transforms on this coarse layer, we reconstructed the fine layer that is a simplified image. Figure 5 shows a simplified image from Figure 4.

The KSOM was then applied to the image derived from Section 1. In our experimentation, the Kohonen network groups the image to 6 clusters. The structure of the output nodes is chosen as two-dimensional (2×3). Initial learning rate is 0.01; training set is randomly selected from the image and after 20,000 iterations, training is stopped. Then the network simulates the entire image and the image is clustered. The simplified image is depicted in Figure 5 and the clustered image is shown in Figure 6.

The road is identified by split and merge method on the clustered image (Figure 7).

The recommended algorithm is used for another part of this orthoimage (Figure 8) and the final result is shown in Figure 9.

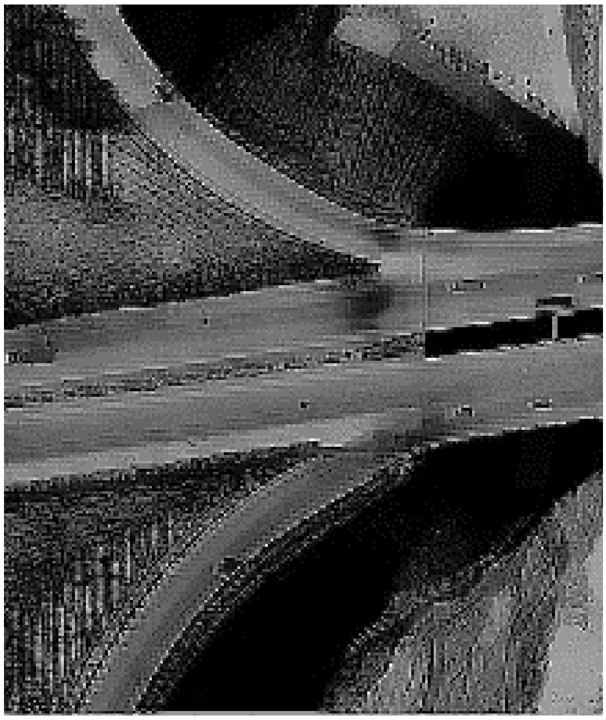

Figure 5. The simplified image.

Figure 6. The clustered image.

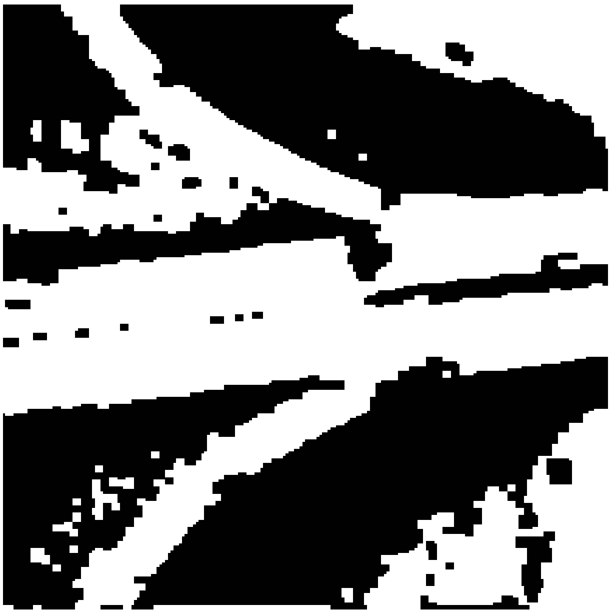

Figure 7. The road identification from the clustered image.

Figure 8. Part of an orthoimage from Kish, IRAN.

6. CONCLUSION

In this paper we presented an algorithm for road identification, which combines wavelet transform, mathematical morphology and Kohonen neural network.

First results show that with combination of wavelet transform and mathematical morphology, the cars and other small objects can be removed from large-scale image.

Second results show that by using Kohonen network and split and merge method, the simplified image is divided into clusters and objects are identifiable, for example, here we identify road.

In recent years, automatic extraction cartographic objects are important in digital photogrammetry. Since there are many objects in an image, it is recommended that first to use a non-linear strategy be used for objects identification and then tracking algorithms be used to extract desired objects.

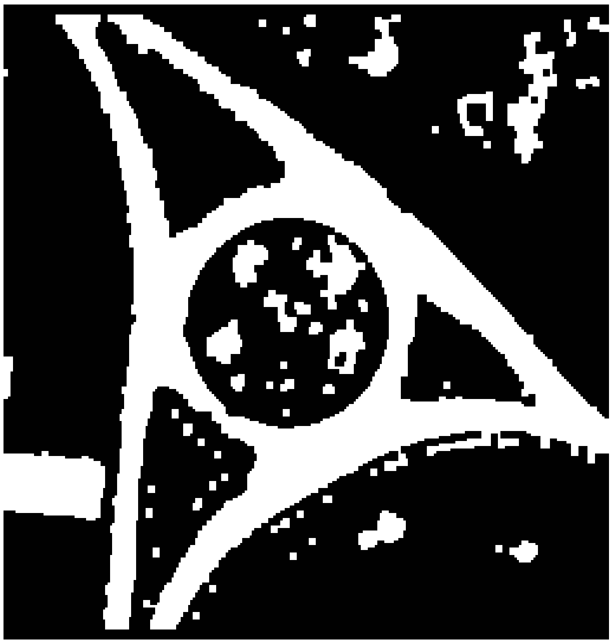

Figure 9. Identification of road.

REFERENCES

1. Amini, J., and Saradjian, M. R., 2000, "Imagemap Simplification by Using Mathematical Morphology", ***ISPRS***, xxxiii (B3), 36-47.
2. Blais, J. A. R., "Multiresolution Spatial Data Representation and Processing in GIS", ***Final Report to Department of Natural Resources, Canada***, (1996).
3. Daubechies, I., "Communication on Pure and Applied Mathematics", Vol. 41, (1988), 909-996.
4. Fischler, M. A., Tenenbaum, J. M. and Wolf, H. C.,

"Detection of Roads and Linear Structures in Low-Resolution Aerial Imagery Uses a Multisource Knowledge Integration Technique", ***CGIP***, Vol. 15, No. 3, (1981), 201-223.

5. Gruen, A., Agouris, P., and Li, H., "Linear Feature Extraction with Dynamic Programming and Globally Enforced LSM", ***IAPRS***, Part 3/1, (1994), 316-323.
6. Goncalves, L., Andrade Netto, L., "A Modular Neural Architecture for Image Classification Using Kohonen Feature Extraction", ***IAPRS***, Vol. xxxi, Part B2, (1996), 117-122.
7. Haralick, R. M., Sternberg, S. R., and Suing, X., "Image Analysis Using Mathematical Morphology", ***IEEE Trans. Pattern Analysis and Machine Intell.***, V. PAMI 9, No. 4, (1987), 532-550.
8. Kohonen, T., "Self Organization and Associative Memory", Springer-Verlay, NewYork, (1988).
9. Oja, E., Lampinen, J., "Unsupervised Learning for Feature Extraction, Computational Intelligence: Imitating Life", ***IEEE Press,*** (1994).
10. Suetens, P., Fua, P. and Hanson, A. J., "Computational Strategies for Object Recognition", ***ACM Computing Surveys***, Vol. 24, No. 1, (1992), 5-61
11. Wirth, G., Mlynski, A., "A Fuzzy Classification System for Analysis of Polymer Spectra Using Wavelet Transforms", ***IEEE Trans***., (1996), 947-950.