An Unsupervised Artificial Neural Network Method for Satellite Image Segmentation

Mohamad Awad
National Council for Scientific Research, Center for Remote Sensing, Lebanon

Abstract: *Image segmentation is an essential step in image processing. The goal of segmentation is to simplify and/or to change the representation of an image into a form easier to analyze. Many image segmentation methods are available but most of these methods are not suitable for satellite images and they require a priori knowledge. In order to overcome these obstacles, a new satellite image segmentation method is developed using an unsupervised artificial neural network method called Kohonen's self-organizing map and a threshold technique. Self-organizing map is used to organize pixels according to grey level values of multiple bands into groups then a threshold technique is used to cluster the image into disjoint regions, this new method is called TSOM. Experiments performed on two different satellite images confirm the stability, homogeneity, and the efficiency (speed wise) of TSOM method with comparison to the iterative self-organizing data analysis method. The stability and homogeneity of both methods are determined using a procedure selected from the functional model.*

Keywords: *Artificial neural network, segmentation, unsupervised, remote sensing, satellite image.*

Received August 14, 2008; accepted September 25, 2008

1. Introduction

Image segmentation is the process of image division into regions with similar attributes [10]. It is an important step in image analysis chain with applications to pattern recognition, object detection, *etc.* Until recently, most of the segmentation methods and approaches are supervised such as MAximum Posteriori (MAP) [8] or Maximum Likelihood (ML) [3] with an average efficiency rate of about 85% [9], [16]. In the supervised methods a priori knowledge is needed to get a successful segmentation process and sometime the required information may not be available.

Satellite images are important source of information which is used in many environmental assessments and monitoring of agriculture, meteorology...*etc.,* Compared to 1972, when Landsat Multi Spectral Scanner System (MSS) was launched satellite image systems now exhibit extraordinary diversity. There are operational satellite systems that sample all available parts of the electromagnetic spectrum with spatial resolution from 0.61 to 1000 m. Satellite image segmentation continues to be an important area in the research, which led to the publication of several methods in this domain.

Self Organizing Map (SOM) or Kohonen's Map is an unsupervised neural network method. Although the use of SOM in image segmentation is well reported in the literature, such as segmentation of printed fabric images [12], or in sonar images [13], their application in satellite image segmentation is not widely known. One can cite the work of [1] which was used in the segmentation of Indian Remote Sensing (IRS) satellite image. Another promising work is the cooperation between SOM and hybrid genetic algorithm [2]. This approach shows good results when applied to different types of satellite images. However, the problem with this approach is the slow speed. The cooperative segmentation approach between K-means and SOM [17] is a recent work where the role of K-means is to segment the image in the coarser scale, and then SOM will re-segment the image in the fine scale. In This method K-means requires the pre-determination of cluster numbers and this kind of work can be taken care by SOM itself. In addition, this type of cooperation requires longer time to converge to an optimal solution.

A known research [6] uses SOM incrementally to segment a Landsat image by first extracting features and then SOM is trained to segment these types of images. This method was attempted on one type of satellite images and for this method to provide accurate results a field survey must be conducted. The correct samples are provided to the SOM network for further training in a supervised manner. In this paper we present a new unsupervised method for segmenting satellite images. The proposed method uses both SOM and a developed technique known as T-Cluster. This new method is compared to another known commercial method the Iterative Self Organizing Data Analysis (ISODATA) and the results of both methods are verified for their stability homogeneity, and accuracy using a procedure implemented as part of the Functional Model (FM) [18].

The paper is structured as follows. In section 2, an overview of SOM and how it is implemented as part of the new segmentation method. In section 3, threshold clustering technique (T-Cluster) is presented. In section 4 TSOM segmentation method is explained. Section 5 covers ISODATA algorithm. Section 6 is devoted to experimental results. Finally, section 7 gives the conclusions.

2. Feature Extraction Using Self Organizing Map

Self Organizing Map (SOM) [5] is an unsupervised neural network method. SOM converts patterns of arbitrary dimensionality into the responses of two dimensional arrays of neurons. One of the important characteristics of SOM is that the feature map preserves neighborhood relations of the input pattern. A typical SOM structure is given in Figure 1. It consists of two layers: an input layer and an output layer. The number of input neurons is equal to the dimensions of the input data. The output neurons are, however, arranged in a two-dimensional array.

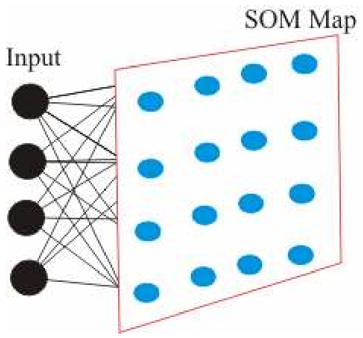


Figure 1. Self-organizing map.

Colors are one of the most important features considered in biological visual systems, since it is used to separate objects and patterns, even in conditions of equi-luminance [7]. SOM is used to map patterns in a three-dimensional color (multi-bands) space to a two-dimensional space. In SOM, the input signals are n-tuples and there is a set of m cluster units (automatic empirical determination with respect to the size of the satellite image). Each input is fully connected to all units. The initial weights are random and small, and their contribution for the final state decreases with the decrease of the number of samples [14]. The network is composed of an orthogonal grid of cluster units (neurons), each is associated with three internal weights for the three layers of the satellite image. At each step in the training phase, the cluster unit with weights that best match the input pattern is elected as the winner usually by using minimum euclidean distance as in equation 1.

$$\|\mathbf{x} - W_l^{\lfloor k \rfloor}\| = \min_i \|\mathbf{x} - W_i^{\lfloor k \rfloor}\| \tag{1}$$

where x is the input vector, $W_l^{\lfloor k \rfloor}$ is the weight of the wining unit l at iteration k, and $W_i^{\lfloor k \rfloor}$ is the weight for neuron i at iteration k. The winning unit and a neighborhood around it are then updated in such a way that their internal weights be closer to the presented input. All the neurons within a certain neighborhood around the leader participate in the weight-update process. This learning process can be described by the iterative procedure as in equation 2.

$$W_i^{\lfloor k+1 \rfloor} = W_i^{\lfloor k \rfloor} + H_{li}^{\lfloor k \rfloor}(\mathbf{x} - W_i^{\lfloor k \rfloor}) \tag{2}$$

where $H_{li}^{\lfloor k \rfloor}$ is a smoothing kernel defined over winning neuron. This kernel can be written in terms of the Gaussian function as in equation 3.

$$H_{li}^{\lfloor k \rfloor} = \alpha^{\lfloor k \rfloor} \exp\left(-\frac{d^2(l,i)}{2(\sigma^{\lfloor k \rfloor})^2}\right) \tag{3}$$

$H_{li}^{\lfloor k \rfloor} \to 0$ when $k \rightarrow To$, where To is the total number of iterations. $\alpha^{\lfloor 0 \rfloor}$ is the initial learning rate and it is equal to 0.1. The learning rate is updated every iteration. $\sigma^{\lfloor k \rfloor}$ is the search distance at iteration k, initially can be half the length of the network or the maximum of either the width or length of the image divided by two. $d(l,i)$ is the distance between the leader neuron l and its neighbor i. As learning proceeds, the size of the neighborhood should be diminished until it encompasses only a single unit.

After SOM neural network converges to a balance state, the original image is mapped from a high color space to a smaller color space. The number of colors in this space is equal to the number of neurons of SOM network. The final weights vectors in the map as the new sample space. This new data set is used for clustering, and allows determining a set of cluster centers.

3. Threshold Technique (T-Cluster)

In order to eliminate small clusters (clusters with few pixels) and to reduce over segmentation problem the following technique (T-Cluster) is implemented. This technique consists of several steps as follow:

- After obtaining cluster centers by SOM the process of clustering starts by calculating the distance between the values of the cluster centers representing the sum of the three bands.
- Two clusters are combined if the distance between their centers is less than a predefined threshold T.
- If step two is correct, the minimum number of pixels is considered in the combination procedure, where the cluster with smaller number of pixels is merged with the larger one. Figure 2 shows more

details of the merging procedure and equation 4 explains the procedure of distance calculation.

$$d(V(P_i),V(P_j)) <= T \tag{4}$$

where T is a predefined threshold and $V(P_i)$ is the value of the three bands of the cluster center P_i which is the sum of the resultant 3 weights obtained from running SOM each weight is multiplied by 255. $V(P_j)$ is the value of the three bands of another cluster center P_j. These two cluster centers are combined together if the distance value is less than a predefined threshold T. The value of the final cluster is the cluster with higher number of pixels.

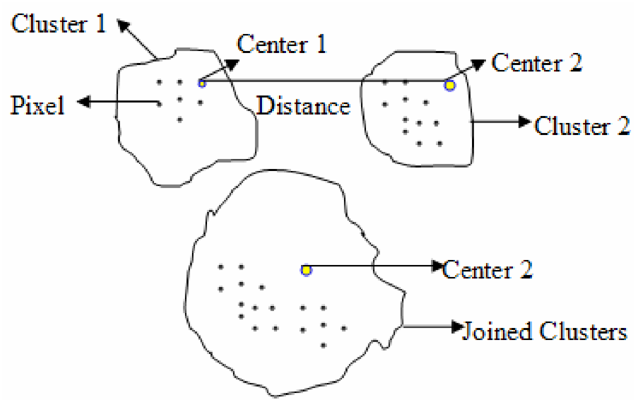


Figure 2. Merging process according to the distance between cluster centers.

4. TSOM Segmentation Method

SOM and the threshold technique (T-Cluster) work sequentially in order to complete the segmentation process. In other words, working separately cannot complete the job correctly as shown in Figure 3. SOM uses satellite image features to organize pixels in group. The highest peaks of the histogram are used as cluster centers and are provided to T-Cluster to deliver the final solution in the image segmentation process.

This method starts by reading a satellite image than it is provided to SOM to organize pixels in groups. The organized pixels are used by T-Cluster to obtain the final number of cluster centers (no under or over segmentation). TSOM fixes the problem of under and over segmentation which are caused by using SOM separately. After that a procedure selected from the FM [18] is used to verify the stability and the homogeneity of the segmented image.

The core of the selected procedure is the SO, which is composed of five elementary blocks that are named measure, criterion, control, modification and Stop as shown in Figure 4. The segmentation process is achieved through one or more iterations of these blocks. In the measure block sample windows $W\times W$ are selected from each segmented image and then the variance and the skewness are computed. This is considered as part of checking region homogeneity.

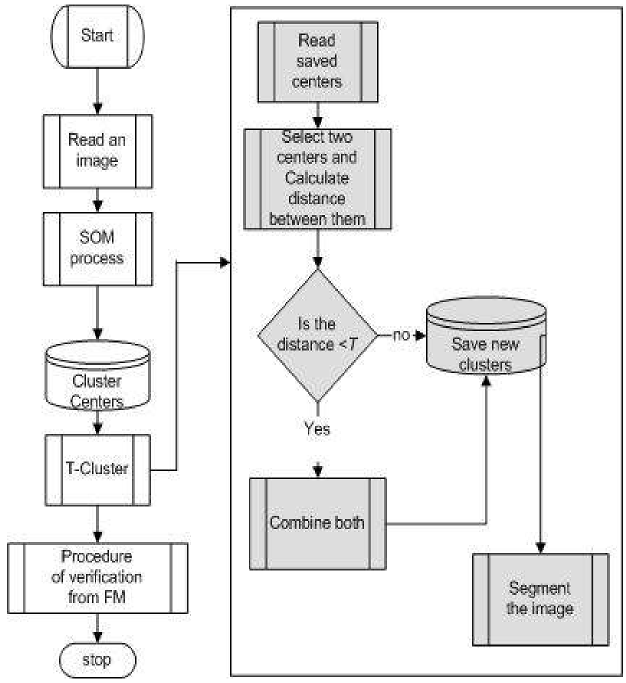


Figure 3. SOM and T-cluster sequential process.

The criterion block receives all the measures from the measure block and build a scalar criterion $C^k(n) = f(F^k(Orr,n))$ for each window n. $F^k(Orr,n)$ represents the measure block of window n at iteration k while Orr is the numerical order of the statistical operations (1-variance, 2-skewness, 3-mean, 4-kurtosis...*etc.*). The following equation 5 is used in criterion computation.

$$C^k(n) = \sum_{i=1}^{M} w^k(Orr,n)F^k(Orr,n) \tag{5}$$

A set of M scalar measures and n windows at iteration k (here M is equal to two because it represents two different statistical operations). The weight w^k (Orr, n) in this evaluation has two values 0.45 and 0.55 for the two statistical operations. These values are taken randomly and can be changed depending on how many statistical operations are used as well as the importance of each operation. Here Orr is equal 1 and 2 the numerical order of the variance and skewness statistical operations. The scalar Criterion is used to identify significant changes in the segmentation result. The Control block takes as input Criterion values and produces the Control value $E^k(n)$ as shown in equation 6.

$$E^k(n) = \frac{C^k(n) - C^{k-1}(n)}{C^k(n) + C^{k-1}(n)} \tag{6}$$

The control value is normalized between -1 and 1 where a positive or negative number means that an over- or an under-segmentation problem exists in the segmentation results, and 0 means that no modification is needed and the results are acceptable.

In the case of SOM and ISODATA the modification can take place in the variation of number of iterations.

In addition the variation of the threshold value can improve or worsen the segmentation results.

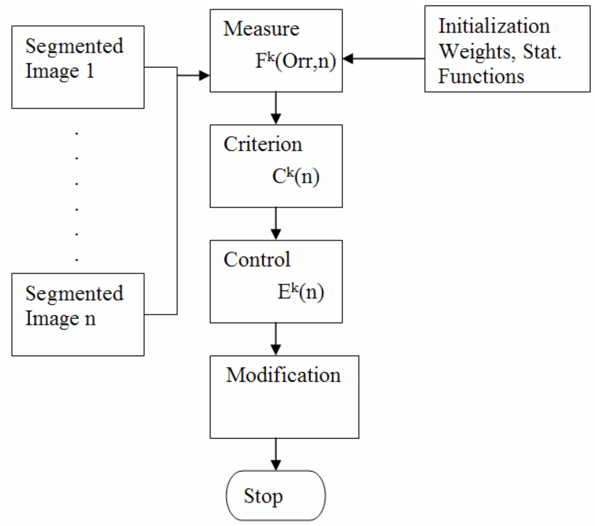


Figure 4. The implemented procedure from the functional model.

5. Iterative Self Organizing Data

Iterative Self Organizing Data (ISODATA) [4] clustering is iterative in that it repeatedly performs an entire classification (outputting a thematic raster layer) and recalculates statistics. Self-organizing refers to the way in which it locates clusters with minimum user input. The ISODATA method uses minimum spectral distance to assign a cluster for each candidate pixel. The process begins with a specified number of arbitrary cluster means or the means of existing signatures (saved cluster centers), and then it processes repetitively, so that those means shift to the means of the clusters in the data. To perform ISODATA clustering, the following must be specified:

- The maximum number of clusters to be considered. Since each cluster is the basis for a class, this number becomes the maximum number of classes to be formed. The ISODATA process begins by determining *N* arbitrary cluster means. Some clusters with few pixels can be eliminated, leaving less than *N* clusters.
- A convergence threshold, which is the maximum percentage of pixels whose class values are allowed to be unchanged between iterations.
- The maximum number of iterations to be performed.

ISODATA is selected to be compared with SOM because it is an unsupervised segmentation method which has many similarities to SOM. In addition, ISODATA is used with ERDAS IMAGINE™ and it is available in our laboratory.

6. Experimental Results

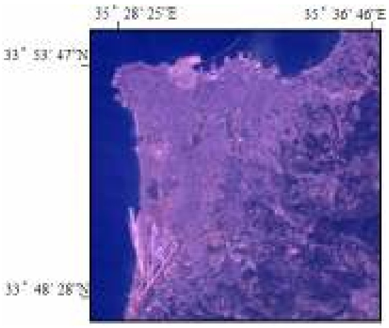
The proposed image segmentation method is implemented using C++ language on an Intel Centrino -1.7 GHz computer. The maximum number of iterations (*To*) for TSOM is 1000 and maximum threshold (*T*) is 60. Two experiments are conducted using two types of satellite images high and medium resolution images (SPOT V and Landsat 7 ETM+). These experiments are carried out to demonstrate the accuracy, and efficiency of TSOM segmentation method compared to ISODATA. At the end of the segmentation processes, the verification is done using the procedure implemented from the FM. Finally, In order to be confident with the results, the segmented images are converted to vector format and the number of labels per cluster is counted.

6.1. Experiment 1

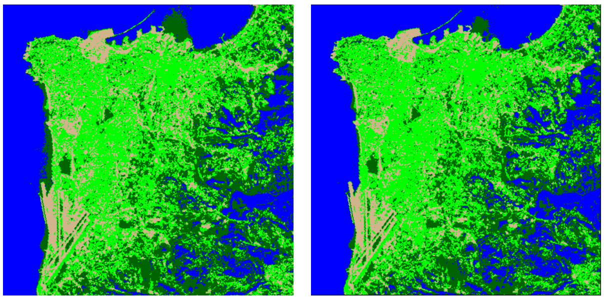
The first test image is Landsat 7 ETM+ with a size of 450×450, the resolution is 30 meter and it consists of three visible bands (1, 2, and 3) as shown in Figure (5a). The results obtained by TSOM segmentation process with threshold value equal to 75 and iteration number equal to 1000 is shown in Figure (5b) (these values are selected to segment the Landsat image after several tests with different threshold and iteration values).

The same image is segmented by ISODATA algorithm and the result is shown in Figure (5c). In order to verify the results of TSOM and ISODATA segmentation methods, the procedure which is part of the FM is used and the results show that TSOM is more stable and homogeneous with value = 0.22 while ISODATA converged after 51 iterations with threshold equal to 1.0 is less stable and homogeneous with value = 0.31.

To increase the confidence in the obtained results, the two segmented images are converted to vector format. There are many raster to vector conversion methods such as the one by Suzuki [11] and by Zenzo *et al.*, [15]. These methods are available in many commercial software such as ArcInfo™. In this research the conversion process is done using the Vector module provided by ERDAS Imagine™. This Vector module is compatible with ESRI's vector formats such as ArcInfo™ and Shapefile™. The vector layer elements are labeled according to each cluster number and the count of labels for each cluster is shown in Table 1. This proves again that ISODATA provides results which suffer from the over-segmentation problem. In this case, more labels are found in ISODATA results than TSOM results. This is clear in the urban classification where TSOM gives one homogeneous cluster while ISODATA mixes the urban cluster with other clusters. In this experiment TSOM took almost 25 seconds to segment the image compared to 45 seconds for ISODATA. this means that TSOM is almost two times faster than ISODATA.



(a) Original image Landsat 7 ETM+.



(b) TSOM result. (c) ISODATA result.

Figure 5. The result for the 1st experiment.

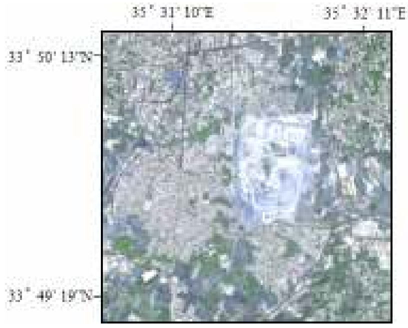
Table 1. Number of labels for each cluster in the segmented Landsat image.

	TSOM	ISODATA
Cluster I	1289	1404
Cluster II	3241	3715
Cluster III	3602	3361
Cluster IV	3839	3538
Total	11971	12020

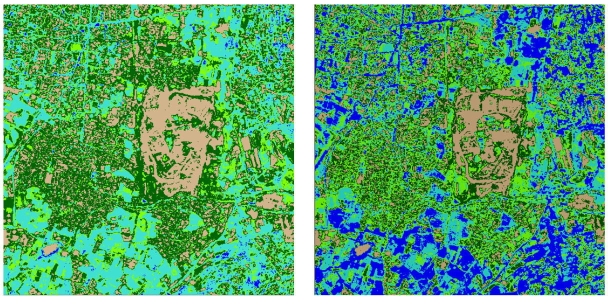
6.2. Experiment 2

The second experiment is performed on SPOT V image with size 540×540 and resolution 5 meters. It consists of 3 bands (1, 2, and 3) pan sharpened with the panchromatic band as shown in Figure 6(a). TSOM is used with threshold equal to 45 and number of iterations equal to 1000, the result is shown in Figure 6(b). The initial number of iterations in ISODATA method is equal to 1000, but it converges after 31 iterations with convergence threshold equal to 1. The result of this segmentation method is shown in Figure 6(c). The segmented image by TSOM has a stability and homogeneity value equal to -0.04 which is very close to zero. On the other hand, the segmented image by ISODATA method has a value equal to 0.6. Comparing the two segmented images one can see clearly that the green cover and the urban area is over segmented by ISODATA algorithm.

The same process is applied to the segmentation results of both methods by converting the raster data to vector data and by counting the number of labels for each cluster. The results in Table 2 show that TSOM does not have the problem of over segmentation. Similar to the previous experiment, TSOM is almost one and half faster than ISODATA. It took TSOM almost 33 seconds to segment the image compared to 50 seconds for ISODATA



(a) Original image SPOT XS V image.



(b) TSOM result. (c) ISODATA result.

Figure 6. The result for the 2nd experiment.

Table 2. Number of labels for each cluster in the segmented SPOT image.

	TSOM	ISODATA
Cluster I	972	3238
Cluster II	3344	9340
Cluster III	12684	10590
Cluster IV	4192	7300
Cluster V	3298	2734
Total	24490	33282

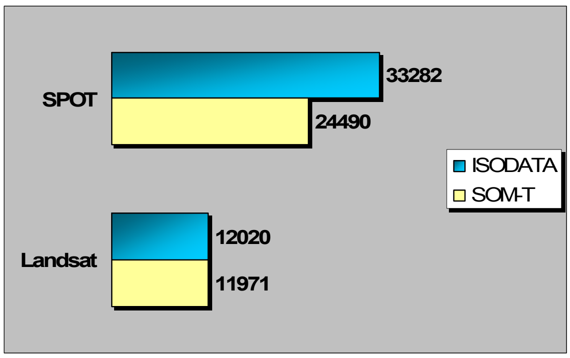


Figure 7. A comparison between the two different segmentation methods.

Finally one can see clearly in the graph as shown in Figure 7 that ISODATA cannot rise above the problem of over segmentation, because of the dependability of ISODATA on the computation of statistical parameters which can be affected by the noise in the satellite image.

7. Conclusion

Segmentation is an important step in image processing. The lack of an unsupervised Artificial Neural Network method to segment satellite images efficiently led us to implement a sequential method for satellite image segmentation using SOM and threshold technique we called it TSOM. The efficiency of this new implemented method depends on the selected number of iterations and threshold values. However, these values can be determined easily for each type of image, and they can be used permanently in segmenting these images. TSOM is compared to ISODATA method which has some similarities, especially in the need for determining the number of iterations and threshold value. The comparison covered two different types of satellite images and the measurement of the accuracy of the results is achieved by using a procedure implemented from the FM which measures the stability and homogeneity of the results and determine the quality of the under and over segmentation on a scale form -1 to 1.

In addition, a process for counting the labels or the polygons is used after converting the segmented raster image into vector format. TSOM gives better results than ISODATA and the speed of TSOM is almost two times faster if compared to the extreme case of running ISODATA where the threshold value of ISODATA is 1.0.

References

[1] Aria E., Saradjian M., Amini J., and Lucas C., “Generalized Cooccurence Matrix to Classify IRS-1d Images Using Neural Networks,” *in Proceedings of ISPRS Congress*, Turkey, pp. 117-123, 2004.

[2] Awad M., Chehdi K., and Nasri A., “Multi Component Image Segmentation Using Genetic Algorithm and Artificial Neural Network,” *Computer Journal of Geosciences and Remote Sensing Letters*, vol. 4, no. 4, pp. 571-575, 2007.

[3] Benediktsson J., Swain P., Ersoy O., and Hong D., “Neural Network Approaches Versus Statistical Methods in Classification of Multisource Remote Sensing Data,” *Computer Journal of Institute of Electrical and Electronics Engineers,* vol. 28, no. 4, pp. 540-551, 1990.

[4] Jensen J., *Introductory Digital Image Processing: A Remote Sensing Perspective*, Prentice-Hall, New Jersey, 1996.

[5] Kohonen T., “Self-Organizing Maps,” *Computer Journal of Springer Series in Information Sciences*, vol. 30, no. 3, pp. 501-501, 2001.

[6] Kurnaz M., Dokur Z., and Olmez T., “Segmentation of Remote Sensing Images by Incremental Neural Network,” *Computer Journal of Pattern Recognition Letters*, vol. 26, no. 8, pp. 1096-1104, 2005.

[7] Levine M., *Vision in Man and Machine*, McGraw Hill Publication, New York, 1985.

[8] Lopes A., Nezry E., Touzi R., and Laur H., “Maximum A Posteriori Speckle Filtering and First Order Textural Models in SAR Images,” *in Proceedings of International Geoscience and Remote Sensing Symposium*, Maryland, pp. 2409-2412, 1990.

[9] Perkins S., Theiler J., Brumby S., Harvey N., Porter R., Szymansk J., and Bloch J., “GENIE: A Hybrid Genetic Algorithm for Feature Classification in Multi Spectral Images,” *in Proceedings of SPIE Applications and Science of Neural Networks, Fuzzy Systems and Evolutionary Computation III 4120*, USA, pp. 52-62, 2000.

[10] Pratt W., *Digital Image Processing*, John Wiley & Sons, New York, 1991.

[11] Suzuki S., “Graph Based Vectorization Method for Line Patterns,” *in Proceedings of Computer Vision and Pattern Recognition Computer Society Conference*, USA, pp. 616-621, 1988.

[12] Xu B. and Lin S., “Automatic Color Identification in Printed Fabric Images by a Fuzzy Neural Network,” *Computer Journal of AATICC Review*, vol. 2, no. 9, pp. 42-45, 2002.

[13] Yao K., Mignotte M., Collet C., Galerne P., and Burel G., “Unsupervised Segmentation Using a Self Organizing Map and a Noise Model Estimation in Sonar Imagery,” *Computer Journal of Pattern Recognition Letters,* vol. 33, no. 9, pp. 1575-1584, 2000.

[14] Yin H. and Allinson N., “On the Distribution and Convergence of Feature Space in Self-Organizing Maps,” *Computer Journal of Neural Computation*, vol. 7, no. 6, pp. 1178-1187, 1995.

[15] Zenzo D., Cinque L., and Levialdi S., “Run Based Algorithms for Binary Image Analysis and Processing,” *Computer Journal of IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 1, pp. 83-89, 1996.

[16] Zhang P., Verma B., and Kumar K., “Neural vs Statistical Classifier in Conjunction with Genetic Algorithm Feature Selection in Digital Mammography,” *in Proceedings of IEEE Congress on Evolutionary Computation CEC*, Australia, pp. 634-639, 2003.

[17] Zhou Z., Wei S., Zhang X., and Zhao X., “Remote Sensing Image Segmentation Based on Self Organizing Map at Multiple Scale,” *in Proceedings of SPIE Geoinformatics: Remotely Sensed Data and Information*, USA, pp. 122-126, 2007.

[18] Zouagui H., Benoit-Cattin H., and Odet C., "Image Segmentation Model," *Computer Journal of Pattern Recognition*, vol. 37, no. 9, pp. 1785-1795, 2004.

Mohamad Awad received his BS and MS degrees from the Lebanese American University in computer science in 1988 and 2001, respectively. He worked for AT&T, Canada from 1989 to 1994 as a programmer. Then he started in early 1996 as research assistant in the Center for Remote Sensing, Lebanese National Council for Scientific Research. He earned his PhD from the University of Rennes I, Brittany, France.