

# Delineation of Lineaments from Satellite Data Based on Efficient Neural Network and Pattern Recognition Techniques

Nikolaos Vassilas<sup>1</sup>, Stavros Perantonis<sup>2</sup>, Eleni Charou<sup>2</sup>, Theocharis Tsenoglou<sup>2</sup>,  
Martha Stefouli<sup>3</sup> and Stavros Varoufakis<sup>2</sup>

<sup>1</sup>Technological Educational Institution of Athens, Department of Computer Science,  
Ag. Spyridonos St., 122 10 Egaleo, Greece  
nvas@teiath.gr

<sup>2</sup>National Research Center "Demokritos", Institute of Informatics and Telecommunications,  
153 10 Agia Paraskevi, Greece  
{sper, exarou, varou}@iit.demokritos.gr, ttheocharis@hotmail.com

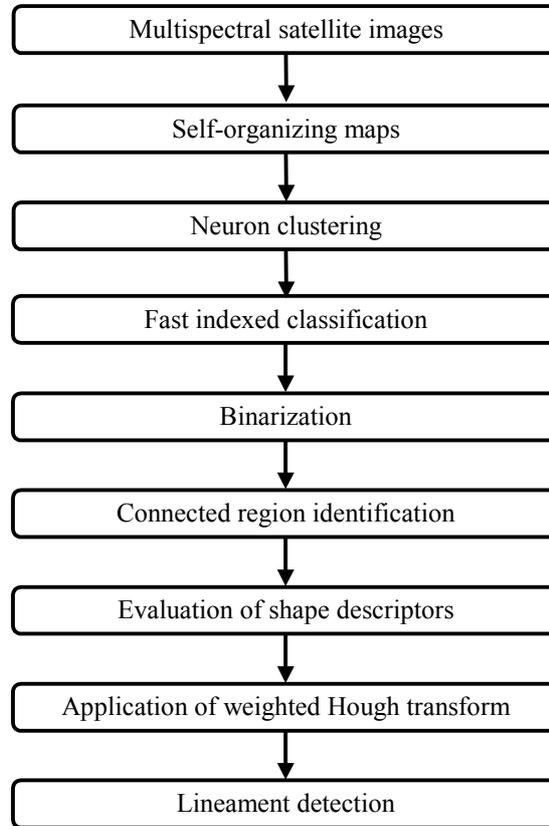
<sup>3</sup>Institute of Geology and Mineral Exploration, Department of Geographical Information  
Systems  
and Satellite Image Processing, 70 Messoghion St., 11527 Athens, Greece  
mst99@otenet.gr

**Abstract:** An automated lineament detection method based on a modified Hough transform is presented. The method first performs an efficient data clustering using Kohonen's self-organizing maps then binarizes the classification result and finally applies the modified Hough transform in order to identify lineaments. The capabilities of the method are described using Landsat TM satellite data from the Vermion area in Greece. The results of the automated analysis show major geological faults in the selected area.

## 1 Introduction

Lineaments are line features or patterns on earth's surface which reflect geological structure. Detection and mapping of lineaments is an important operation in the exploration for mineral deposits, in the investigation of active fault patterns, water resources, etc. On satellite images, lineaments usually appear as lines or linear formations whose pixels are either lighter or darker (or have a different color) than the background pixels.

Most lineament mapping is done visually using enhanced images. However, visual analysis is time consuming and subjective. As an aid to save time and improve objectivity of lineament analysis we developed a novel method based on self-organizing maps for fast and efficient unsupervised classification followed by a modified implementation of the Hough transform for lineament detection. The whole procedure for the delineation of lineaments is outlined in Fig. 1. The major steps of this approach are analyzed in the subsequent sections.



**Fig. 1.** Procedure for the delineation of lineaments.

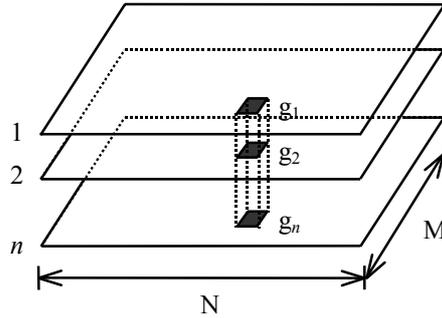
Previous work [1] has shown that the above technique can successfully been applied to the delineation of lineaments from airborne geophysical data such as grid measurements of magnetic and electromagnetic fields. However, since airborne data are much more costly than satellite images, this paper aims at showing that significant results can be obtained from multispectral satellite images as well.

## **2 Unsupervised Classification Using Self-organizing Maps**

A time and memory saving methodology for efficient clustering and automatic classification of images or grided spatial data has been presented in [2],[3]. Significant clustering and classification speedup was achieved with no significant loss in terms of final performance by: a) using a self-organizing map (SOM) for quantizing the input space, b) clustering the neurons of the map instead of the pixels of the original image

and c) using fast indexing techniques for efficient generation of the classification result.

In the sequel, we assume  $M \times N$  multispectral satellite images of  $n$  channels (bands). The input space  $\mathbf{R}^n$  is used to represent the image as a set of  $M \times N$  points (also called *spectral signatures*) whose coordinates are the gray levels of each band (see Fig. 2 for an explanation).



**Fig. 2.** Representation of multispectral satellite images of  $M \times N$  pixels and  $n$  bands. The spectral signature of a pixel is represented by a point  $(g_1, g_2, \dots, g_n) \in \mathbf{R}^n$  where  $g_i$  is to the gray level of the  $i^{\text{th}}$  spectral band,  $i = 1, 2, \dots, n$ .

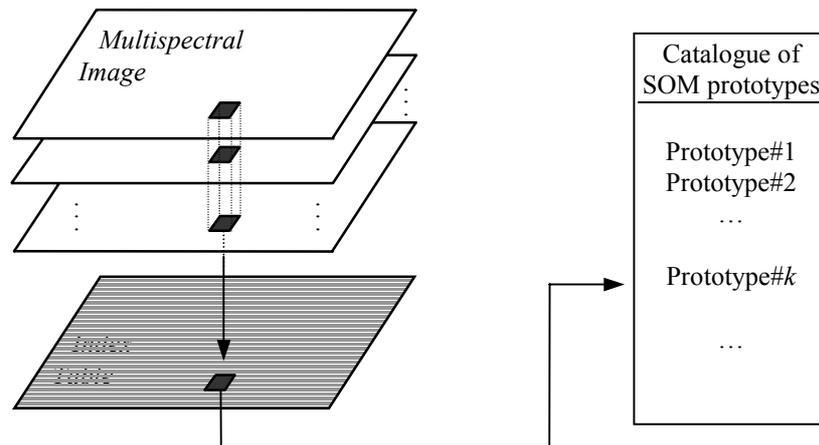
The first stage of the proposed methodology involves vector quantization of the input space using a 2-D lattice of neurons trained with Kohonen's SOM algorithm. Following a random presentation of the  $n$ -dimensional points, the result is to obtain a catalogue of prototypes (the asymptotic weights of the neurons) that quantize the satellite image.

Next, we use indexing techniques for mapping the pixels of the original image to their corresponding prototypes. To this end, a  $M \times N$  *index table* is constructed to store pointers from pixels to their closest prototypes as shown in Fig. 3. The replacement of the original image with the catalogue of prototypes and the index table constitutes the indexed representation of the multispectral satellite image and results in both: a) data compression, and b) significant speedup during clustering and classification.

In general, as the number of neurons of the map grows larger, the approximation of the original data space will be more accurate due to the smaller quantization distortion (provided that the map self-organizes). However, according to experience, map sizes of no more than  $16 \times 16$  neurons should suffice in most applications. In the case of large volumes of multispectral data from  $n$  bands with 8 bits/gray, compression ratios of approximately  $n:1$ , when 256 prototypes are used, are readily attainable.

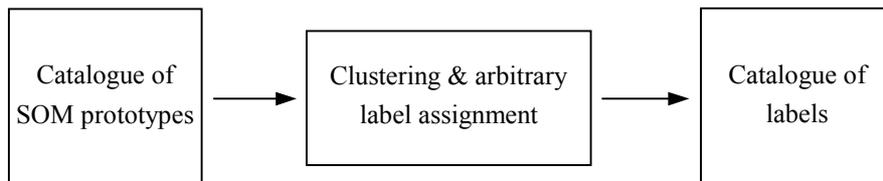
Typically, automatic classification involves clustering of the data space followed by label assignment. However, due to the large number of data points, clustering performed on the original image data is both memory and time inefficient. On the other hand, if clustering is performed on neuron prototypes, a speedup of several

orders of magnitude can be achieved allowing even the use of the most computationally complex clustering algorithms.



**Fig. 3.** Representation of multispectral satellite images by the index table and SOM prototypes. The index table stores pointers from pixels to their nearest prototypes.

Next, labels are assigned to each cluster. These clusters along with their labels represent the automatic classification categories (see Fig. 4).



**Fig. 4.** Creation of the catalogue of labels.

The final classification result is obtained by: a) classifying the prototypes to obtain a corresponding catalogue of labels (as in Fig. 4) and, b) fast indirect addressing through the index table, since its pointers to the prototypes also point to their labels, as shown in Fig. 5. As was observed, by selecting the number of classification categories similar to the number of possible land cover and lithological classes, most of the linear formations present in Landsat TM satellite images were preserved. On the other hand, if one uses SAR (*Synthetic Aperture Radar*) images, the number of categories should

be similar to the lithological classes since, as is well known, radar penetrates vegetation.

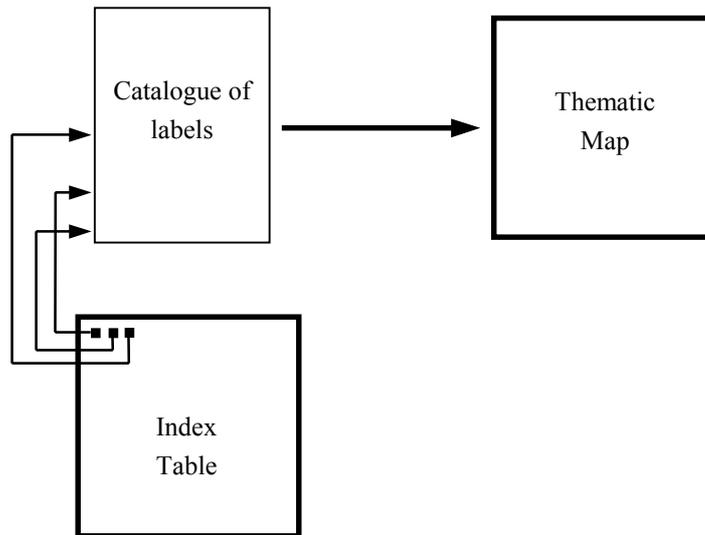


Fig. 5. Fast indexed classification using the index table and the catalogue of labels.

### 3 Lineament Extraction Using Pattern Recognition Techniques

In order to apply the pattern recognition techniques described in this section we first need to convert the classification result to a binary image. Binarization is performed interactively as follows: from the available categories (gray levels) of the classification result the user selects those that will be considered as foreground with the remaining categories considered as background. At this stage, usually many linear formations are present, along with regions that do not correspond to linear formations.

#### 3.1 Connected Region Identification and Shape Descriptor Evaluation

The connected regions of the binary image were specified by using the fast (one pass) label assignment procedure outlined in [4] (p. 198), whereby a linked list is constructed for storing simultaneously: a) labels belonging to different regions using vertical pointers, and b) equivalence classes using horizontal pointers.

For each connected region we then calculate three shape descriptors, all of which can be evaluated in terms of the central moments  $m_{ij}$  of the region:

- a) the area  $A = m_{00}$ ,

- b) the angle  $\phi = (1/2) \tan^{-1} [2 m_{11}/(m_{20} - m_{02})]$  of the principal axis relative to the x-axis, and,
- c) the elongation  $\varepsilon = |I_{max}/I_{min}|^{1/2}$  of the region by finding the ellipse that best fits the region (in the sense that it has the same moments of inertia). The elongation is defined as the ratio of major to minor axis length (so that  $\varepsilon > 1$ ) where  $I_{min} = m_{20}\sin^2\phi + m_{02}\cos^2\phi - m_{11}\sin 2\phi$  and  $I_{max} = m_{20}\cos^2\phi + m_{02}\sin^2\phi + m_{11}\sin 2\phi$ .

### 3.2 Modified Hough Transform

The Hough transform is a popular and powerful method for detecting parametrically described shapes in images. However, even in its simplest application, which is the detection of straight lines, the original Hough transform is susceptible to the presence of both random and correlated noise that may give rise to spurious maxima in the accumulator array [5]. Variants of Hough transform have been applied with some success to the delineation of lineaments [6],[7]. However, we noted that in most data sets, interference from unwanted pixels was very prominent.

It is thus highly desirable to have methods of hindering irrelevant pixels from contributing to the accumulator array. One such method has been recently presented by the authors [8] in order to prevent the foreground pixels to contribute the same amount to all accumulator array points that correspond to lines passing through them. This is achieved by allowing preferential weighting of certain pixels through an appropriate voting kernel that depends on the shape descriptors of the connected regions of the image described in the previous section.

Given a binary image, we are interested in the detection of straight lines whose points  $(x, y)$  are parameterized by  $r = x \cos\theta + y \sin\theta$ , where  $r$  is the distance of the origin from a particular line and  $\theta$  is the angle formed by the normal to the line and the x-axis, as shown in Fig. 6.

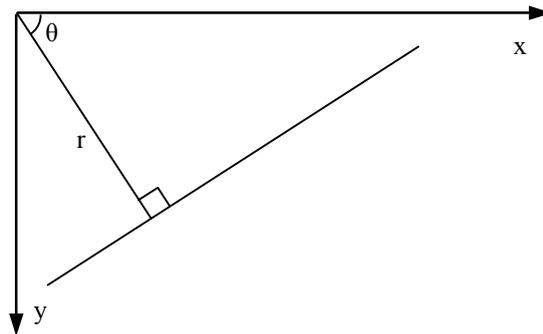


Fig. 6. The Hough transform parameterization of lines ( $r = x \cos\theta + y \sin\theta$ ).

For each connected region, we consider its geometrical center and increment for various values of  $\theta$  the corresponding cells in the accumulator array. To evaluate the contribution of each region to various cells in the accumulator array we utilize its related shape descriptors and express the dependence formally by introducing a voting kernel.

The voting kernel is a continuous function of the shape descriptors and is constructed by taking into account the following considerations. The presence of an elongated region is taken as strong evidence for the existence of a line parallel to the principal axis of the region. Thus, the contribution to the accumulator array should increase with  $\varepsilon$ . Moreover, for a given  $\varepsilon$ , this contribution should be maximum for  $\phi = \theta$  and drop with increasing  $|\phi - \theta|$ . The rate of change with increasing  $|\phi - \theta|$  should clearly depend on  $\varepsilon$ . For very elongated regions, only the direction  $\phi = \theta$  should be incremented. On the other hand, nearly circular regions ( $\varepsilon \approx 1$ ) should be allowed to vote equally for all directions. Finally, a dependence of the kernel on the area  $A$  should be introduced in order to minimize interference effects and avoid spurious accumulator array maxima. Contributions from regions of large  $A$  and relatively small  $\varepsilon$  should be suppressed. These regions are a major source of correlated noise, because spurious lines can be formed from points lying in their interior. At the other extreme, regions of very small  $A$  should also be discarded as random noise. In-between these extreme cases, regions of small  $\varepsilon$  whose area is comparable to a characteristic intermediate area scale  $A_0$  may be part of a chain of regions contributing to a disrupted linear structure and should be taken into account.

A function used for modeling the above desired properties is the following:

$$f = \varepsilon \exp[-(\varepsilon - 1)(\phi - \theta)^2] A \exp[-(A - A_0)/(A_0 \varepsilon^2)]$$

where  $A_0$  can be chosen as the average area of regions whose elongation exceeds a predefined threshold. However, our simulations have shown that the method is robust, exhibiting stable performance for a wide range of values for  $A_0$ .

## 4 Experimental Results and Discussion

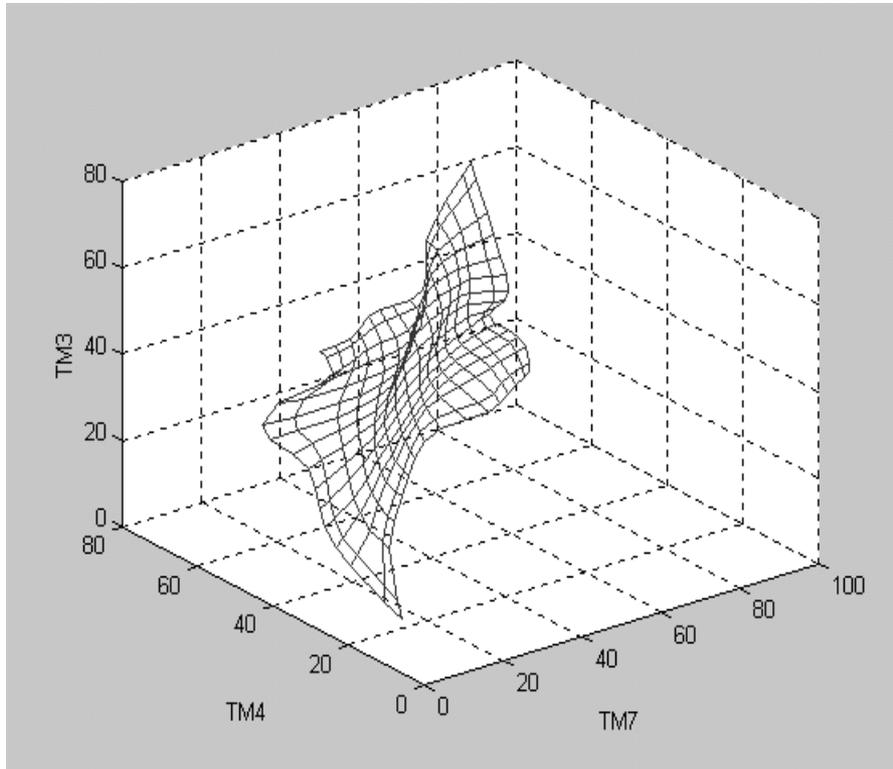
The automatic lineament detection method has been applied with success using bands 7, 4 and 3 of a multispectral LandSat TM satellite image of Vermion area in Greece. A visual lineament identification was carried out by an expert photo-interpreter on the satellite image and the output was superimposed and digitized as shown in Fig. 7.

Automatic classification in 10 categories was performed by first quantizing the space using a 16x16 self-organizing map (i.e., with 256 prototypes) and then clustering the neurons of the map using the Fuzzy Isodata algorithm [9]. Fig. 8 shows the map in 3D space.

Following unsupervised classification, the categories containing most information regarding lineaments were selected through visual inspection of the classification result. Binarization was then performed by setting all pixels to background except for those of the selected categories.



**Fig. 7.** LandSat image with photo-interpreted lineaments superimposed.



**Fig. 8.** Asymptotic weights of the self-organizing map of 16x16 neurons.

The 8-neighbors connected regions along with the *area*, *principal axis angle* and *elongation* shape descriptors of each region were then computed according to Sec. 3. Finally, the lineaments were found using the proposed modified Hough transform and are shown superimposed on the classification result in Fig. 9.

As it is displayed in Fig. 9, several of the interpreted lineaments (see Fig. 7) have been also detected using the automatic detection method. This is quite obvious for the lineaments of NE direction. On the other hand lineaments of NW direction have not been traced using the automatic lineament detection technique. This is partly due to the high relief of the area that enhances features that are on a vertical direction to sun illumination. The degree of “sinuosity” of lineaments seems also to play an important role for not detecting some of the features. However, the process has proved effective on detecting “straight” lineaments.

In short, we believe that our proposed method for automatic lineament detection can prove a useful tool to geologists and geophysicists who would like to have a first glance on a digital lineament map without significant time investment. Such a map can then serve as the starting map of a series of improved lineament maps produced within a GIS, whereby, existing lineaments can be modified, new ones added and wrong ones



**Fig. 9.** Automatic lineament detection super-imposed on the classification result.

removed, by the use of pointing devices, incorporating additional geological and/or geophysical information.

## References

1. Vassilas, N., Perantonis, S., Charou, E., Seretis, K., Tsenoglou, Th.: Automatic Lineament Detection from Geophysical Grid Data Using Efficient Clustering and Weighted Hough Transform Algorithms. In Proc.: Workshop on Intelligent Techniques for Spatio-Temporal Data Analysis in Environmental Applications. ACAI'99. Chania, Greece, July (1999) 16-25
2. Vassilas, N.: Efficient Neural Network-based Methodology for the Design of Multiple Classifiers. In: Jain, L.C., Fanelli, A.-M. (eds.): Recent Advances in Artificial Neural Networks: Design and Applications. CRC Press, Boca Raton, Florida (2000) 95-125
3. Vassilas, N., Charou, E.: A New Methodology for Efficient Classification of Multispectral Satellite Images Using Neural Network Techniques. Neural Processing Letters **9** (1999) 35-43
4. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis and Machine Vision. Chapman and Hall, London (1993)
5. Leavers, V.F.: Which Hough Transform? Computer Vision and Image Understanding **58:2** (1993) 250-264
6. Wang, J., Howarth, P.J.: Use of the Hough Transform in Automated Lineament Detection. IEEE Tran. Geoscience and Remote Sensing **28:4** (1990) 561-566
7. Karnieli, A., Meisels, A., Fisher, L., Arkin, Y.: Automatic Extraction and Evaluation of Geological Linear Features from Digital Remote Sensing Data Using a Hough Transform. Photogrammetric Engineering and Remote Sensing **62:5** (1996) 525-531
8. Perantonis, S., Vassilas, N., Tsenoglou, Th., Seretis, K.: Determination of Linear Formations Using a Weighted Region Based Hough Transform. Electronics Letters **34:7** (1998) 648-650
9. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York (1981)