We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS

BOOK CITATION INDEX

INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Automated Mapping of Hydrographic Systems from Satellite Imagery Using Self-Organizing Maps and Principal Curves

Marek B. Zaremba
Université du Québec en Outaouais
Canada

1. Introduction

A fully automatic and high-precision cartographic mapping of terrain features such as forests, rivers or roads from multispectral satellite or aerial images is a challenging problem in remote sensing, largely due to the fact that it requires an adequate representation of irregular and discontinuous objects. Being able to provide sub-meter resolution multispectral images, high-resolution satellites such as QuickBird or Ikonos broaden the application possibilities of satellite imagery and offer the possibility of making them sensors of choice for a variety of environmental applications. Their high spatial and radiometric resolution facilitates visual interpretation. Temporal resolution of image databases can be largely increased, due to the known revisit time and pointing capabilities of the satellite platform, which facilitates large-scale change-detection and monitoring of selected areas in order to keep natural environment databases up-to-date. Certain analyses involving spectral change detection and dynamically obtained maps can be performed more easily and in a more automated fashion. Finally, high-resolution satellite images offer the potential for being ever more competitive in terms of price with the aerial images.

This chapter presents methods based on Self-Organizing Maps (Kohonen et al., 1996) developed in efforts to fully automate the generation of hydrographic maps from remotely sensed imagery. The complexities of generating cartographic representations of hydrological objects, such as rivers and lakes, from satellite and aerial images consists generally of two categories of tasks: the first involves the extraction process of the required linear or a real feature while the second involves generation of a suitable representation in a form appropriate for cartographic map presentation. The presented approach applies the technology of Self-Organizing Maps (SOM) at both stages of the hydrological mapping process, i.e., the detection of water bodies from multispectral images and the subsequent tracing of hydrological systems or networks.

The first task can be approached by applying a classification technique or through a scene analysis method. A number of different methods have been reported in the literature. Conventional image processing techniques typically apply edge detection algorithms (Ma & Manjunath, 2000) in efforts to define water regions. A similar problem of road detection from satellite images was discussed in (Auclair Fortier et al., 2001). A rule-based approach to segmentation of satellite images was presented in (Ton et al., 1991). Selection of

multispectral bands for segmentation of low-resolution ocean images was discussed in (Ainsworth & Jones, 1999). Availability of three-dimensional information can enhance the mapping process. In addition to providing the raw elevation data, the Digital Elevation Models (DEMs) have been processed (Jenson & Domingue, 1988) to extract hydrographic features, such as basins or flowlines. The conventional methods generally used are sensitive to noise and usually provide only a rough shape description or present extractions that are discontinuous suggesting erroneous heterogeneity of objects. The extraction process described in this chapter makes use of the topological relationships between or among the objects - automatically extracted by SOM.

The second task requires derivation of objects suitable for representation in map form. The conventional approaches resort to classical image processing techniques, such as skeletal thinning algorithms or splines. Extraction of rivers from digital SPOT images using the approximation of the river shape by the snake method was reported in (Dillabaugh et al., 2002). A two stage, multi-resolution procedure was applied to produce a detailed outline of the channel banks and, subsequently, fit high resolution imagery to the channel banks using the dynamic contour technique. The calculation of the snake shapes is based on dynamic programming and is computationally expensive. Another disadvantage of the snake-based method is that the resulting shape can often be located outside of the area of the river. A powerful technique for obtaining skeletons of planar objects turned out to be neural network architectures. A self-organizing map was initially investigated for skeletal shape extraction in (Datta et al., 1996). An initial linear topology was able to evolve to circular and - depending on the angle between the neuron and its neighbors - forked forms. However, the skeletonization results were not invariant to the rotations of the image because of the flow-through version of SOM adopted in that work. By using the batch formulation (Ritter et al., 1996) of the SOM algorithm, this problem was solved in (Sing et al, 2000).

Many of the conventional approaches, such as skeletal thinning algorithms or splines, will fail. This is as a result of the quality of the derived objects from the input images which generally contain noise or objects that are rendered discontinuous as a result of object scarcity. One of the objectives of this study was to develop techniques able to deal with images in the form of sparse data sets. Sparse data sets refer to spatial discontinuity of a homogeneous and topologically related object within an object class. There are known several methods to cope with the sparse shape and occlusions, the most popular of them (by the compromise of efficiency and simplicity) are based on a median filtering and morphological set operations such as opening and closing (Haralick & Shapiro, 1992). Afterwards, a standard thinning algorithm is applied to extract the shape skeleton. Such methods usually introduce a large amount of distortions when the shape sparseness and occlusion is significant.

In order to effectively deal with sparse and discontinuous remote sensing data at the stage of tracing hydrographic systems, the principal curve approach was adopted. In this approach, the skeletal description of point sets is extracted directly, without a sequential thinning procedure. Principal curves are one of the nonlinear generalizations of principal components. They were first defined in (Hastie & Stuetzle, 1989) as self-consistent smooth curves defined by the property that each point of the curve is the average of all data points that project to it, i.e., for which that point is the closest point on the curve. Consequently, the curve passes through the "middle" of a multi-dimensional probability distribution or data cloud. It is shown in (Duchamp & Stuetzle, 1996) that Hastie & Stuetzle principal curves are saddle points of the distance function (MSE) which explains why cross-validation is not a

viable method for choosing the complexity of principal curve estimates. The definition of principal curves in (Kégl et al, 2000) assures that, assuming the distribution has finite second moments, principal curves always exist. This makes it possible to carry out a theoretical analysis of learning principal curves from training data. Based on their definition, the authors developed the Polygonal Line Algorithm. The algorithm starts with a straight line segment and, in each iteration, increases the number of segments by one by adding a new vertex to the previously produced polygonal curve. After adding a new vertex, the positions of all vertices are updated in an inner loop so that the resulting curve minimizes a penalized distance function. The algorithm is extended in (Kégl & Krzyzak, 2002) to find principal graphs of data sets, and is applied to handwritten character skeletonization.

In many aspects principal curves are related to several other unsupervised learning methods, including Vector Quantization and Generative Topographic Mapping (Chang & Ghosh, 2001). It was demonstrated in (Mulier & Cherkassky, 1995) that discretized principal curves are essentially equivalent to SOMs. In (Sun & Yang, 2007), a new definition of principal curves - Principal Curve with Feature Continuity (PCFC) - was proposed. PCFC focuses on both reconstruction error minimization and feature continuity. It builds a continuous mapping from samples to the extracted features so the features preserve the inner structures of the sample set. The properties of PCFC make it a powerful tool for the tasks of feature extraction for signal representation. Principal curves and surfaces were investigated in (Einbeck et al., 2010) in the context of multivariate regression modeling. The situation was analyzed where the intrinsic dimensionality of the data tends to be very small due to the high redundancy induced by complex dependency patterns between the involved variables. It is then useful to approximate the high-dimensional predictor space through a low-dimensional manifold (i.e., a curve or a surface), and use the projections onto the manifold as compressed predictors in the regression problem. The authors used the local principal curve algorithm for the compression step, and provided an algorithm extendible to manifolds of arbitrary dimension.

One of the first applications of principal curves was identification of ice flows and their outlines in satellite images (Banfield & Raftery, 1992). Initial estimates of floe outlines come from the erosion-propagation (EP) algorithm, which combines erosion from mathematical morphology with local propagation of information about floe edges. The edge pixels from the EP algorithm are grouped into floe outlines by grouping them about principal curves rather than points on lines. Local Principal Curve algorithm was used on the Beaver County, PA, digitalized floodplain data to reconstruct rivers in the valleys (Einbeck et al., 2005). The curve moves along the data cloud with along first local principal components. First local principal component is considered to be a (biased) approximation of the tangent to the crest line of the estimated density, gives locally the best fit. The principal graph algorithm proposed in (Kégl & Krzyzak, 2002) was used to obtain principal curves that could be served as skeletons of a fingerprint (Miao et al., 2007). Based on the obtained principal curves, a minutiae extraction algorithm is proposed to extract minutiae of the fingerprint. An application aiming at the understanding the potential impact of the microbicide on HIV viral transmission by assessing the kinetics of a microbicide lubricant was presented in (Caffo et al., 2007). The experiment was conducted by imaging a radiolabeled lubricant distributed in the subject's colon. The tracer imaging was conducted via single photon emission computed tomography (SPECT), a non-invasive, in-vivo functional imaging technique. A novel principal curve algorithm was developed to construct a three dimensional curve through the colon images. The algorithm was tested on several two dimensional images of familiar curves. An approach based upon the theory of principal

curves to find vessel lumen centerlines for the application to 3D CT angiography (CTA) was reported in (Wang et al., 2009).

In the approach presented in this chapter, the location of each vertex of piecewise linear generating curves that represent skeletons of the objects corresponds to the position of a particular SOM unit, treated here as a centre of the neural gas architecture (Martinetz et al., 1993). The proposed method makes it possible to extract the object skeletons and to reconstruct the planar shapes of sparse objects based on the topological constraints of generating lines and the estimation of local scale (Zaremba & Palenichka, 2002). In the final stage of the mapping process, the skeletal description of point sets (extracted directly, without a sequential thinning procedure) is used for tracing the final hydrographic shape by applying the principal curve algorithm between consecutive vertices of the skeleton. The experimental test results using satellite images demonstrate the accuracy of the proposed approach and its utility for fully automated mapping of hydrological objects.

2. Satellite data

2.1 Source image data

In the case of multi-dimensional image data, such as multi-spectral satellite imagery, several images of the same scene can be provided simultaneously. In remote sensing, they represent, for example, a view of the earth surface obtained by using electromagnetic radiation of different wavelengths. A pixel with image coordinates (i,j) of a multi-spectral image is a vector $\boldsymbol{b}_{i,j} = [b_1, b_2, \ldots, b_T]$ of T intensity values corresponding to T wavelength bands. The extraction and tracing of hydrological objects was performed on a data set consisting of medium-resolution Landsat imagery. Images obtained from the Landsat-7 satellite using the Enhanced Thematic Mapper Plus (ETM+) scanning radiometer instrument are produced with seven bands of reflected energy and one band of emitted energy. Since one of the eight ETM+ channels is a panchromatic channel, in Landsat-7 imagery $T = 7$. In many application areas, multi-dimensional image data are also obtained from a single image by extracting local intensity properties with respect to each image point (i,j). This approach is frequently applied to images with textured objects and backgrounds. In the case of Landsat-7, such extraction could be performed on the panchromatic channel data. The characteristics of the Landsat-7 ETM+ satellite sensor are given in Table 1.

Fig. 1. Landsat-7 image of the Eastern Ontario area

Spectral Band	Half-Amplitude Bandwidth (μm)	Ground Resolution (m)
1	0.450±0.005 - 0.515±0.005 (Blue)	30
2	0.525±0.005 - 0.605±0.005 (Green)	30
3	0.630±0.005 - 0.690±0.005 (Red)	30
4	0.775±0.005 - 0.900±0.005 (Near IR)	30
5	1.550±0.010 - 1.750±0.010 (Mid IR)	30
6	10.40±0.100 - 12.50±0.100 (Thermal)	60
7	2.090±0.020 - 2.350±0.020 (IR)	30
8	0.520±0.010 - 0.900±0.010 (Panchromatic)	15

Table 1. Spectral and spatial characteristics of Landsat-7

Landsat-7 images were pre-processed at level 1G: The 1G level image is a radiometrically and geometrically corrected raw level 0R image. The main scene of the study covers the area of the Ottawa region (Fig. 1), with a diversified river system: the Ottawa River with wide sections and islands of different size, the Gatineau River in the northern part of the image, and the Rideau river with some sections narrower than the image resolution.

3. Extraction of water regions

3.1 Topology of the network

The extraction is performed using pixel-by-pixel multispectral classification using unsupervised image segmentation and a set of labelled examples. The mapping from the space of spectral features to the space of terrain classes is performed by a self-organizing feature map (SOM) architecture.

The classical (SOM) is a two-layer network, where each input layer neuron x_j has a feed-forward connection w_{ij} to each output layer neuron y_i [Kohonen et al., 1996]. A feature map is created by adjusting the connection weights from m-dimensional-vector input nodes to n-dimensional interconnected output nodes ($n < m$). The topology of a three-dimensional self-organizing feature map is shown in Fig. 2. The resulting network has the ability to adapt its

behavior in such a way that the location (k, l, m) of the activated output neuron in the network becomes specific to certain characteristic features (a_k, b_l, c_m) of the input signal X.
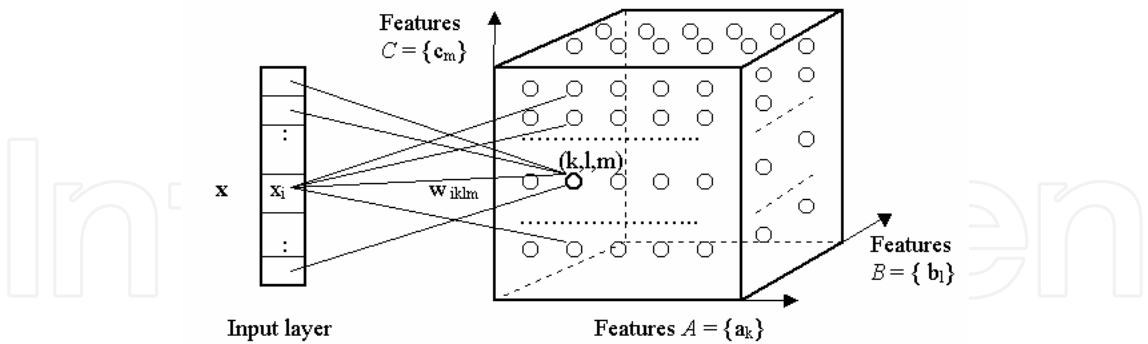

Fig. 2. Topology of a three-feature self-organizing map

In this application, the definition of the input vector consists in finding a discriminative set of spectral bands from Table 1. The selection is performed based on the distribution of pixels for the water and the non-water regions. An example of the distribution for the Near-IR band (Band 5) is shown in Fig. 3.

Fig. 3. Distribution of pixel blue-channel intensity for water and non-water regions
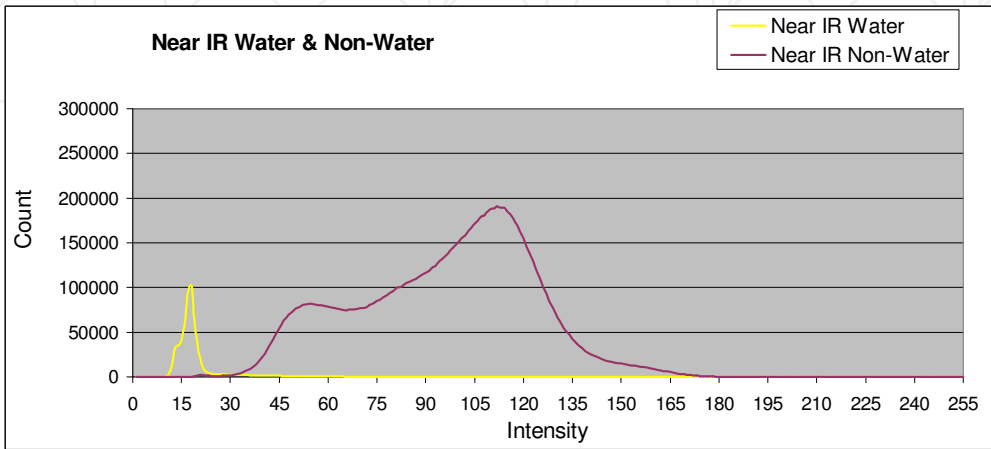

Fig. 4. Distribution of pixel Near IR-channel intensity for water and non-water regions

The distributions in Figures 3 and 4 show that the Near IR channel is more discriminative than the Blue channel. Combinations of channels offer better classification results than using a single channel. For the final architecture of the classification network, three features were selected, corresponding to the Green, Blue and Near IR bands.

3.2 Learning method

The lesrning method consists of two phases: the SOM network learning phase, enhanced by the use of fuzzy logic paradigm, and the merging and evaluation phase.

Network learning phase

The learning of the network consists in updating the connection weight vector **w**. Let, for a 3-D network, $\mathbf{x} = (x_1, x_2, x_3)^T$ represent the input vector and w_{ij} represent the weight matrix for the weights between unit *j* of the input layer (j = 1,2,3) and unit *i* of the output layer (let us assume, for simplicity, a linear order of the output neurons). The basic learning algorithm takes place in three steps.

1. Computing of matching values for each unit in the output layer. If we define the winner with index c, the neuron with the smallest network input

$$\| x - w_c \| = \min\{\| x - w_{ij} \|\} \tag{1}$$

is declared the winner neuron.

2. Updating the weights. The weights corresponding to the winner unit and its neighboring units are updated according to the formula:

$$w_j(k+1) = w_j(k) + \frac{\sum_{i=1}^{n} \alpha_{ij}(k)^{m(k)}(x_i(k) - w_j(k))}{\sum_{i=1}^{n} \alpha_{ij}(k)^{m(k)}} \quad \textit{for } i \in \textit{topological neighbourhood} \tag{2}$$

$$w_j(k+1) = w_j(k) \qquad\qquad \textit{otherwise}$$

where

$$w_i = \frac{\sum_{i=1}^{n} x_i(\alpha_{ij})^m}{\sum_{i=1}^{n}(\alpha_{ij})^m} \tag{3}$$

The topological neighborhood, may be defined in several ways. In a typical choice for the neighborhood function it falls off with distance between the output hypercube units. Thus nearby neurons receive stronger updates and end up responding more strongly to nearby input patterns. This arrangement enables the topological information to be fed into the network (Villmann et al., 1997). The presence of uncertainty and non-Gaussian, nonstationary statistics the methodology of fuzzy sets offers significant advantages. In the fuzzy set paradigm, a pixel can belong to several clusters with different certainty α mapped to a range [0,1]. Fuzzy C-means was the first algorithm to incorporate fuzzy login in the learning process. It is an iterative algorithm used to find the grades of membership α_{ij} and cluster centers w_j to minimize the criterion

$$J = \sum_{i=1}^{n} \sum_{j=1}^{c} (\alpha_{ij})^m \mid x^i - w_j \mid^2 \qquad (4)$$

where c is the number of clusters, and n represents the number of input samples. A major drawback of the algorithm is that it becomes time-consuming when the number of clusters is large. The combination of SOM and fuzzy C-means can overcome the drawbacks of both methods and improve the performance of classification. In the fuzzy SOM (FSOM), the learning rate of SOM can be replaced by the membership values calculated as

$$\alpha_{ij} = \left[\sum_{k=1}^{c} \left(\frac{\mid x^i - w_j \mid}{\mid x^i - w_k \mid} \right)^{\frac{2}{m-1}} \right]^{-1} \qquad (5)$$

The rate is related to the distance between the centers of the clusters and the samples. The clusters closest to the sample (the winner) take the largest membership value. Unlike fuzzy C-means, which calculates all membership values, for fuzzy SOM the k in (5) relates to the size of neighbourhood, so only membership values of the winner and its neighbourhood are calculated. In order to reduce the computational time, the parameter m in equation (5), which is equivalent to the neighbourhood size, changes with time according to:

$$m(k) = m_0 (1 - e^{-E(k)}) + m_{\min} \qquad (6)$$

where m_0 and m_{min} are positive constants greater than one and E(k) is the Euclidean distance between successive weight values.

3.   Calculate Euclidean distance

$$E(k) = \max \{ \| w_j(k+1) - w_j(k) \|^2 \} \qquad (7)$$

If $E(k) < \varepsilon$, stop. Else, go to Step 2.

Merging and evaluation phase

An essential question that has to be addressed in the design of a SOM network is the size of the network, i.e., the number of neurons. The network can not be too small so that patterns which are important but can be distinguished only among a large number of patterns are not missed. On the other hand, an excessive number of neurons may hinder the detection of a pattern by splitting it into several sub-patterns. The optimal size of the network can be determined by initially selecting a large network, and subsequently limiting its size through merging closely located neurons.

Assume two neurons are defined by the weights $\mathbf{w}_i = (w_{i1}, w_{i2} \dots w_{in})$ and $\mathbf{w}_j = (w_{j1}, w_{j2} \dots w_{jn})$. If $\| w_i - w_j \| < T$, where T is a pre-specified threshold, then the two corresponding clusters can be merged. The new neurons will be defined by

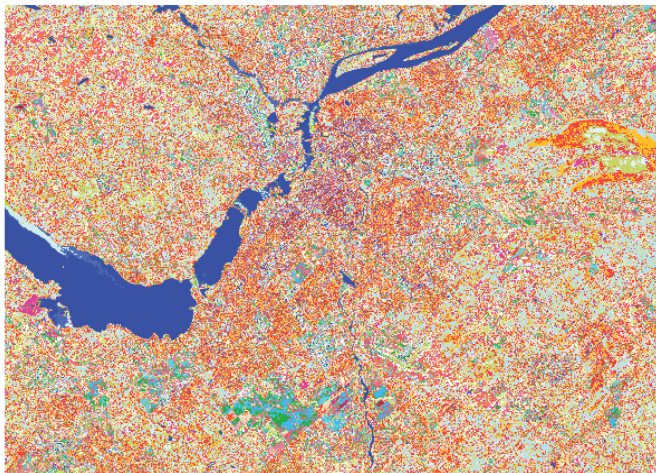$$\mathbf{w}_{new} = \frac{n\mathbf{w}_i + m\mathbf{w}_j}{n + m}, \qquad (8)$$

where n and m are the hit rates (input vectors corresponding to the neuron). The more the input data activate the neuron, the greater the value of the weight assigned.
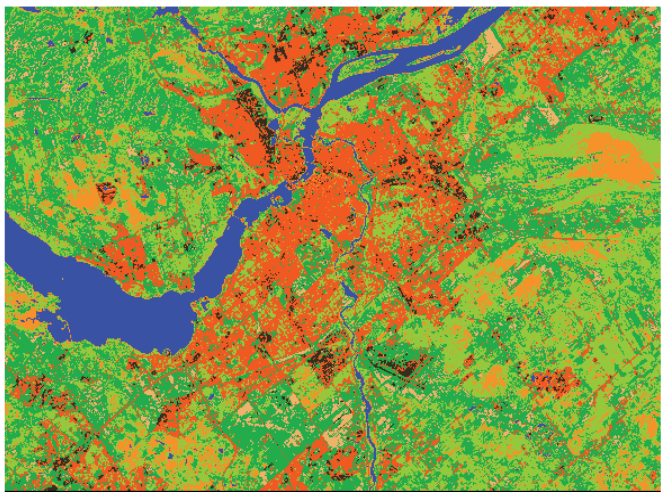
A cluster validity method is required in order to measure the quality of the merging results. We defined the following quality function:

$$V = \frac{\frac{1}{n}\sum_{i=1}^{c} \sum_{x \in X_i} ||x_k - w_i||}{\min_{i,j} ||w_i - w_j||} \tag{9}$$

where $c > 2$ is the total number of neurons, and n is the total number of samples. The optimal number of neurons is that for which V becomes minimum. The parameter V is the ratio of the average distance between the inputs and their corresponding neuron to the minimum distance between neurons. The inputs should be made as close to their corresponding neurons as possible, and the neurons distributed as far apart as possible. The procedure of merging neurons continues as long as the distance between the neurons remains smaller than a given threshold value. Different thresholds can be selected, with a subsequent use of V to evaluate the quality of merging results.

a)

b)

Fig. 5. Clustering results   a) 100 nodes, b) 7 nodes

In summary, the procedure for merging neurons consists of the following steps:
1. Select T; load the learned neurons and the hit rates of the neurons.
2. Calculate the smallest Euclidean distance between pairs of neurons. If the distance is smaller than T, merge the two neurons according to equation (8).
3. Repeat Step 2 based on the values of neurons after merging. If the smallest distance is larger than T, stop the iteration.

The number of clusters that has the smallest V, for different T, is the final optimal number of clusters.

Typical results of the merging procedure are depicted in Fig. 5. Figure 5a shows the segmentation of the satellite image into 100 classes. The results of merging the neurons into a network with 7 units are shown in Fig. 5b.

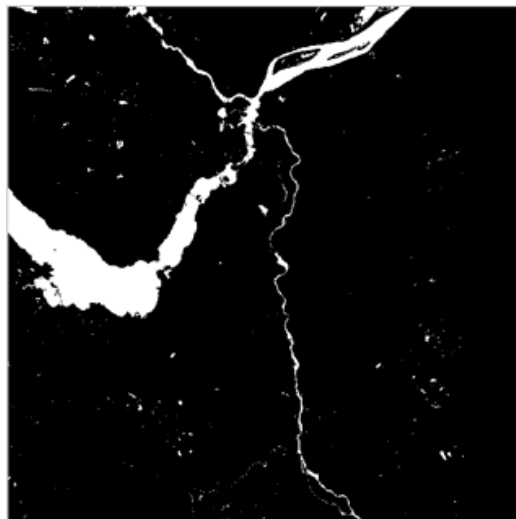The optimum results were obtained for 25 nodes. The resulting water areas are shown in Fig. 6.

Fig. 6. Binary map of the water bodies

4. Mapping and tracing of hydrographic networks

The basic idea of the hydrographic network tracing algorithm consists in generating principal curves between the vertices of the network skeleton graph. In the first stage of the process, the set of vertices of the principal curve skeleton graph is constructed by applying a SOM-based algorithm. The algorithm consists in iterative updating of vertex coordinates in the data space while progressively decreasing the span of a SOM kernel function. The resulting piece-wise linear skeleton is constructed as follows:

Step 1: Initialization. The input data vectors $v_1, …, v_N$ represent the image plane coordinates of input binary image $\phi(m,n)$ obtained by the segmentation process, such as shown in Figure 6. The SOM units $u_1, …, u_K$ are initialized in the two-dimensional data space, with weight vectors $w_1, …, w_K$ representing the grid topology. The initial number of units K should be comparable with the maximal number of vertices in the object skeletons. The scale attributes of all units are set to the maximal value, which is comparable with the image size. This means that each unit is initially connected to its q nearest neighboring SOM units, where q is the maximal number of vertices.

Step 2: Determination of vertex neighborhood connectivity. The SOM units with the same scale value are connected into a local structure by checking the connectivity between the

neighborhood units. During the first iteration, the full connectivity is established, the unit scales being all equal and maximal in size.

Step 3: Matching point coordinates with vertices. With the Euclidean distance used as a similarity measure, this step corresponds to the Voronoi tessellation of input data set. The number a SOM unit in the topological space closest to the ith data point is determined by:

$$z_i = \arg \min_r | | v_i - w_r | |, \qquad i = 1, \ldots, N \tag{10}$$

As a result, the object support regions of the input binary image $\phi(m,n)$ will be partitioned into N Voronoi regions $\{V_i\}$. The new weights of the rth unit will be determined based on pixels belonging to the rth Voronoi region and the regions neighboring to V_r.

Step 4. Evaluation of vertex scale. The scale attribute of the rth SOM unit, ρ_r, is estimated by considering the Voronoi regions for the SOM units, and is proportional to the number of points in the region V_r. The more precise scale estimate is made by finding the size of the maximal disk inscribed into the rth Voronoi region:

$$\rho_r = \arg \max_k \left\{ \max_{(u,v) \in V_k} \left\{ \left\| \frac{1}{|S_k|} \sum_{(m,n) \in S_k(u,v)} \varphi(m,n) - \frac{1}{|R_k|} \sum_{(m,n) \in R_k(u,v)} \varphi(m,n) \right\| \right\} \right\} \tag{11}$$

where $R_k(u,v)$ is kth disc and $S_k(u,v)$ is a ring around it. The ring width is equal to the minimal distance between two objects considered as isolated. Such a scale estimate is used because of the assumption of shape sparseness (the direct method of maximal inscribed disk produces poor results). The point $t_r = (u_r,v_r) \in V_r$, which corresponds to the maximal value of scale (Eq. 10), is called the *attraction point* of region V_r. The use of such a scale estimate and attraction point is useful when calculating conditional expectation of unit coordinates in the next step of this algorithm. Figure 7 depicts the detected vertices represented as circles proportional to the vertex scales.
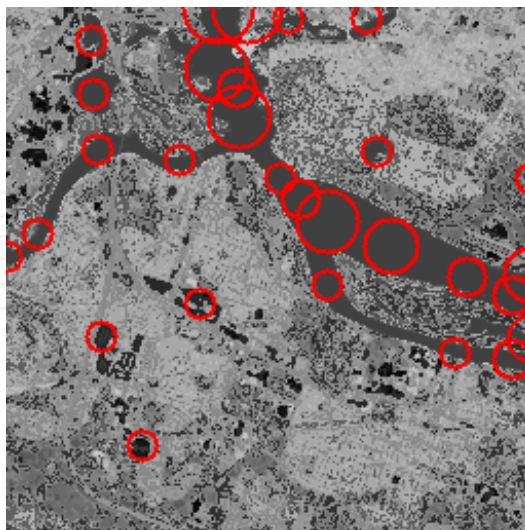

Fig. 7. Vertices represented as circles of different scale

Step 5: Updating of coordinates using conditional expectation. The weights of SOM units, i.e., the vertex coordinates, are updated at this stage. The approach of conditional expectation [Singh et al., 2000] is used here, modified in order to take into account the current unit scale:

$$w_r = \frac{\sum_{i=1}^{K} v_i \Phi(w_r, w_i, t_r)}{\sum_{i=1}^{K} \Phi(w_r, w_i, t_r)}, \quad r = 1, \ldots, K \tag{12}$$

where $\Phi(w_r, w_i, t_r)$ is a monotonically decreasing kernel function defined indirectly in terms of two intermediate variables: distance $\Delta_1(w_r, w_i)$ between the positions of the rth and the ith SOM units, and distance $\Delta_2(w_r, t_r)$ between the position of unit r and the attraction point $t_r \in V_r$. A Gaussian kernel function has been used as the smoothing kernel function $\Phi(.)$ of the total distance $\Delta_1 + \Delta_2$:

$$\Phi(\Delta_1 + \Delta_2) = \exp\left(-\frac{(\Delta_1 + \Delta_2)^2}{2\sigma^2}\right) \tag{13}$$

where σ is the span of the kernel function. The distance function $\Delta_1(w_r, w_i)$ is defined between SOM neighborhood units connected to u_i. The simultaneous use of two variables in the kernel function allows the algorithm to adapt to the segments of objects of different scales of interest. The smoothing function (Eq. 13) represents the conditional probability that the point v_i belongs to unit u_r at the current iteration step.

Step 5. Determination of SOM vertices connectivity graph. During this step, the connectivity is determined between all non-connected vertices and the local structures. Vertices and elementary straight-line segments are also determined. If a vertex is connected to a line segment, then this vertex is added to the skeleton and a new SOM unit with its local connectivity attributes is inserted into the list of units. Such a vertex insertion provides more exact approximation of skeletons with crossing lines and other higher-order connections.

In order to avoid ragged skeleton lines for elongated thick objects and false connections between skeleton vertices of disconnected object parts, which often result from applying the Minimum Spanning Tree algorithm (Singh et al., 2000), The decision about the vertex connectivity is made by using a context-dependent connectivity test. The test is based on the Markov random chain model of vertices belonging to the same elementary curve and the Bayesian principle of a decision-making process. The Markov chain model is used here to provide the probability of the position of the next vertex with respect to the positions of the two immediately preceding and connected vertices. This model of the curve skeleton macro-growth process can be defined by the conditional probabilities of the new vertex position with respect to the positions of the previous two adjacent vertices on the same curve. Let u_s, u_l and u_k be three consecutive vertices which generate segments $[u_s, u_l]$ and $[u_l, u_k]$. The Markov chain model of a skeleton piece-wise linear graph G can be defined in terms of a conditional probability of the slope $\theta_{l,k}$ of line segment $[u_l, u_k]$ with respect to the slope $\theta_{s,l}$ of its preceding straight-line segment $[u_s, u_l]$:

$$\{P(\theta_{l,k}, [u_l, u_k] \subset G) = P(\theta_{l,k} / \theta_{s,l}, \forall\{[u_s, u_l], [u_l, u_k]\} \subset G\}, \tag{14}$$

where $P(\theta_{l,k} / [u_k, u_l] \in G)$ is the unconditional probability of the slope of the straight-line segment between vertices k and l of the same line G, and $P(\theta_{k,l} / \theta_{s,l}, \forall\{[u_s, u_l], [u_l, u_k]\} \subset G$ is the conditional probability of the slope of the straight-line segment between vertices k and l

of the same generating line G provided that the slope of straight-line segment $[u_s, u_l] \subset G$ equals $\theta_{s,l}$. After adopting a limit probability, δ, in the connectivity test by a probability thresholding, i.e., $P(u_k \wedge u_l / x) \geq \delta$, the connectivity test can be written in an explicit form:

$$\frac{P(x \, / \, u_k \wedge u_l)}{P(x \, / \, u_k \, | \, | \, u_l)} \geq \frac{\delta}{(1-\delta)} \cdot \frac{P(u_k \, | \, | \, u_l)}{P(u_k \wedge u_l)} \, , \tag{15}$$

where $P(u_k | | u_l) = 1 - P(u_k \wedge u_l)$ is the prior probability of no connection between vertices u_k and u_l, and $P(x/u_k | | u_l)$ is the conditional probability of the connectivity feature x provided vertices u_k and u_l are not connected. The connectivity test by Eq. (15) guarantees that the error due to a wrong connection of vertices will be less than $(1 - \delta)$.

Figure 8 gives an example of the results of skeletonization of a network of hydrographic objects. The piece-wise linear approximation of the principal curve is shown. The skeleton demonstrates the utility of the presented approach in a situation of multiple rivers. In addition to showing the skeleton graph, the white form in Fig. 8b depicts also the results of a reconstruction of the hydrographic shape in Fig. 8a from the skeleton. The reconstruction consists in interpolating the scales of adjacent vertices along the skeleton segments. Only large values of the vertex scales (see Step 4) were used.

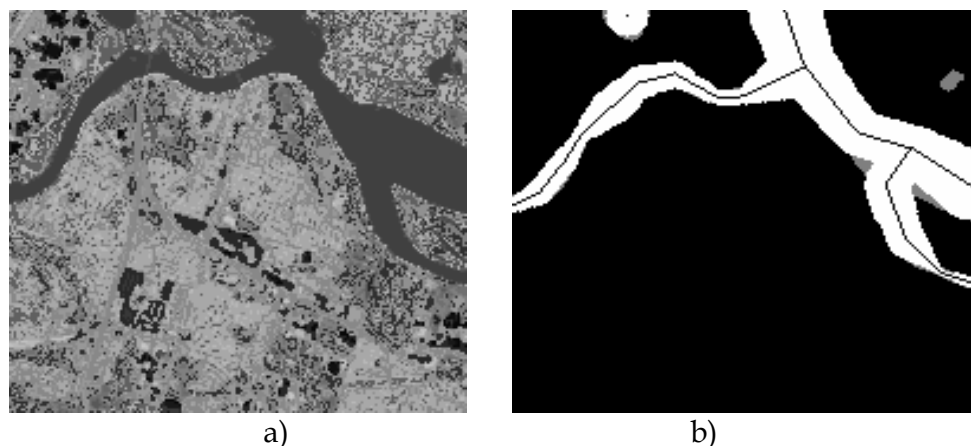a)   b)

Fig. 8. Skeleton graph (b) for the water basin in (a)

In the second stage of the hydrographic network tracing process, the vertex points are used for tracing the river by connecting them by a principal curve. A curve $\mathbf{C}$ is said to pass through the middle of a dataset if every point $\mathbf{x}$ on the curve is the average of the observations projecting onto it. The Hastie and Stuetzle definition of principal curves is based on probability densities. Let $\mathbf{X}$ denote a two-dimensional random vector distributed according to a probability density p, and let $\mathbf{C} \subset \mathbf{R}^2$ be a smoothly embedded closed interval. For each point $\mathbf{x} \in \mathbf{R}^2$, let $d(\mathbf{x}, \mathbf{C})$ denote the distance from $\mathbf{x}$ to $\mathbf{C}$. Because $\mathbf{C}$ is compact, for each $\mathbf{x} \in \mathbf{R}^2$ the distance $d(\mathbf{x}, \mathbf{C})$ is realized by at least one point of $\mathbf{C}$. The *projection map*

$$\pi_{\mathbf{C}} : \mathbf{R}^2 \to \mathbf{C} \tag{16}$$

is the map which assigns to each $\mathbf{x} \in \mathbf{R}^2$ a point $\pi_{\mathbf{C}}(\mathbf{x}) \in \mathbf{C}$ realizing the distance from $\mathbf{x}$ to $\mathbf{C}$, that is,

$$d(\mathbf{x}, \mathbf{C}) = \| x - \pi_{\mathbf{C}}(\mathbf{x}) \|. \tag{17}$$

We can now formalize a definition of a principal curve. A curve **C** is called *self-consistent* or a *principal curve* of a density p if $E(\mathbf{X} \mid \pi_C(\mathbf{X}) = \mathbf{x}) = \mathbf{x}$ for almost every $\mathbf{x} \in C$.

The notion of projection also leads to a natural definition of the distance between a random vector **X**, or its associated density, and a curve **C**:

$$d^2(\mathbf{x}, \mathbf{C}) = E(\; \|\mathbf{X} - \pi_C(\mathbf{X})\|^2 \;). \tag{18}$$

As proved in (Hastie & Stuetzle, 1989), principal curves are critical points of the distance in the variational sense. Since all principal curves are saddle points of the distance, there are no local minima. In order to better illustrate this point, let us set this property against the regression problem, where the conditional expectation $E(Y \mid x)$ minimizes the expected squared distance $E(Y - f(X))^2$ among all functions f. Although principal curves are not local minima, in general, they are local minima of the distance for "low frequency variations" (Duchamp & Stuetzle, 1996). The definition of "low-frequency" depends on the principal curve itself, and not only on the underlying density. In practical situations, and for the type of shapes required to trace hydrographic objects, this limitation of principal curves is not critical.

An example of a river traced using a polygonal principal curve is shown in Figure 9. The examples shown in Figures 8 and 9 illustrate the versatility and good performance of the SOM-based approach to solve the problem of tracing hydrographical systems. The presented method works on long stretches of rivers. It can also effectively deal with a system of connected, tributary rivers.
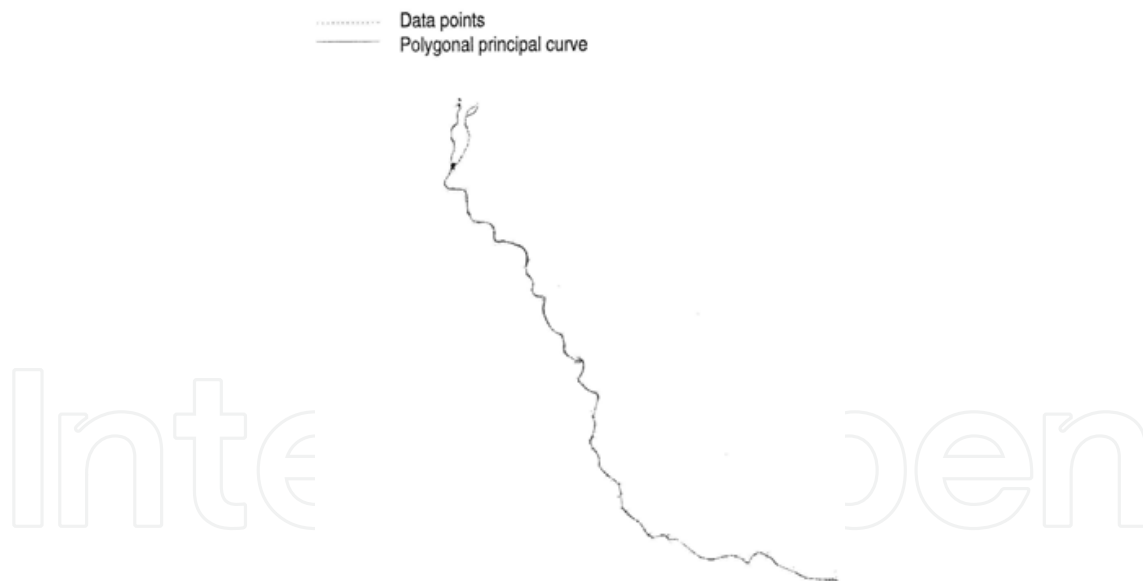

Fig. 9. A river mapping example

5. Conclusion

A semi-automated approach to the extraction and tracing of river networks was presented, where SOM networks are used in both stages of the hydrological mapping process, i.e., the detection of water bodies from multispectral images and the subsequent tracing of hydrological systems or networks. By adopting the principal curve approach, the algorithms precented in this chapter permit to effectively deal with sparse image data. The tracing

method makes it possible to extract the river skeletons based on the topological constraints of river segments and the estimation of local scale. A distinctive feature of the proposed SOM-based approach is the introduction of such attributes as local scale, used in the scale-based updating phase of SOM units, and connectivity of SOM units. A probabilistic model-based connectivity test is performed while connecting adjacent SOM units, i.e., skeleton vertices. The presented method for the extraction of the skeletal shape and the tracing of hydrographic objects by using structured self-organizing maps was tested on Landsat-7 ETM+, SPOT and QuickBird satellite images, demonstrating the utility of the presented approach.

6. Acknowledgement

7. References

Ainsworth, E.J. & Jones, I.S.F. (1999). Radiance spectra classification from the ocean color and temperature scanner on ADEOS, *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 37, No. 3, pp. 1645-1656.

Auclair Fortier, M.-F.; Ziou, D.; Armenakis C. & Wang, S. (2001). Automated Correction and Updating of Roads Databases from High-Resolution Imagery, *Canadian. J. Remote Sensing*, Vol. 27, No.1, pp. 76-89.

Banfield, J. D. & Raftery, A. E. (1992). Ice floe identification in satellite images using mathematical morphology and clustering about principal curves, *Journal of the American Statistical Association*, Vol. 87, No. 417, pp. 7-16.

Caffo, B.S.; Crainiceanu, C.M.; Deng, L. & Hendrix, C.W. (2007). A case study in pharmacologic imaging using principal curves in Single Photon Emission Computed Tomography, *Working Paper 143* (May 2007), John Hopkins University, Dept. of Biostatistics.

Chang, K.-Y and Ghosh, J. (2001). A unified model for probabilistic principal surfaces, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 23, No. 1, pp. 22-41.

Datta, A.; Parui, S.K. & Chaudhuri, B.B. (1996). Skeletal shape extraction from dot patterns by self-organization, *Proc. 13th Int. Conf. Pattern Recognition*, Vol. 4, 1996, pp. 80-84.

Dillabaugh, C.R.; Niemann, K.O. & Richardson, D. (2002). Semi-Automated Extraction of Rivers from Digital Imagery, *Geomatica*, Vol. 6, No. 3, pp. 263-284.

Duchamp, T. & Stuetzle, W. (1996). Extremal properties of principal curves in the plane. *The Annals of Statistics*, Vol. 24, No. 4, pp. 1511-1520.

Einbeck, J.; Tutz, G. & Evers, L. (2005). Local Principal Curves. *Statistics and Computing*, Vol. 15, pp. 301-313.

Einbeck, J.; Evers, L. & Powell, B. (2010). Data compression and regression through local principal curves and surfaces. *International Journal of Neural Systems*, Vol. 20. pp. 177-192.

Haralick, R.M. & Shapiro, L.G. (1992). *Computer and Robot Vision*, Addison-Wesley, Reading, MA.

Hastie, T. & Stuetzle, W. (1989). Principal curves, *Journal of the American Statistical Association*, Vol. 84, No. 406, pp. 502-516.

Jenson, S.K. & Domingue, J.O. (1988). Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric Engineering and Remote Sensing*, Vol. 54, pp. 593-600.

Kégl, B.; Krzyzak, A.; Linder, T. & Zeger, K. (2000). Learning and design of principal curves, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 3, pp. 281-297.

Kohonen, T.; Oja, E.; Simula, O.; Visa, A. & Kangas, J. (1996). Engineering applications of the self-organizing map. *Proceedings of the IEEE*, Vol. 84, pp. 1358-1384.

Ma, W.-Y. & Manjunath, B.S. (2000). EdgeFlow: A Technique for Boundary Detection and Image Segmentation, *IEEE Trans. Image Processing*, Vol. 9, pp. 1375-1388.

Martinetz, T.; Berkovich, S. & Schulten, K. (1993). Neural-gas network for vector quantization and its application to time-series prediction, *IEEE Trans. on Neural Networks*, Vol. 4, No. 4, pp. 558-569.

Miao, D.Q; Tang, Q.S. & Fu, W.J. (2007). Fingerprint minutiae extraction based on principal curves, *Pattern Recognition Letters*, Vol. 28 , Issue 16 (December 2007), pp. 2184-2189.

Mulier, F. & Cherkassky, V. (1995). Self-organization as an iterative kernel smoothing process, *Neural Computation*, Vol. 7, No. 6, pp. 1165-1177.

Ritter, H.; Martinetz, T. & Schulten, K. (1992). *Neural Computation and Self-Organizing Maps: An Introduction*. Addison-Wesley, Reading, MA, 1992.

Singh, R.; Cherkassky, V. & Papanikopoulos, N. (2000). Self-Organizing Maps for the Skeletonization of Sparse Shapes, *IEEE Trans. on Neural Networks*, Vol. 11, pp. 241-248.

Sun, M. & Yang, J. (2007). Principal curves with feature continuity. Proceedings of the 11th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining, Nanjing, China, May 22 - 25, 2007. In: *Lecture Notes In Artificial Intelligence*. Zhou, Z., Li, H. & Yang, Q. (Eds.), pp. 785-792, Springer-Verlag, Berlin, Heidelberg.

Ton, J.; Sticklen, J. & Jain, A.K. (1991). Knowledge-Based Segmentation of Landsat Images, *IEEE Trans. on Geoscience and Remote Sensing,* Vol. 29, pp. 222-231.

Villmann, T.; Der, R.; Herrmann, M. & Martinez. T.M. (1997). Topology preservation in self-organizing feature maps: Exact definition and measurement. *IEEE Trans. on Neural Networks*, Vol. 8, No. 2, pp. 256-266.

Wong, W.C. K.; So, R.W. K. & Chung, A.C. S. (2009). Principal Curves: a Technique for Preliminary Carotid Lumen Segmentation and Stenosis Grading, *The MIDAS Journal*, July 29, 2009. http://hdl.handle.net/10380/3096.

Zaremba, M.B. & Palenichka, R.M. (2002). Probabilistic morphological modelling of hydrographic networks from satellite imagery using self-organizing maps. *Control and Cybernetics*, Vol. 31, No. 2, pp. 343-369.

Self Organizing Maps - Applications and Novel Algorithm Design

Edited by Dr Josphat Igadwa Mwasiagi

Kohonen Self Organizing Maps (SOM) has found application in practical all fields, especially those which tend to handle high dimensional data. SOM can be used for the clustering of genes in the medical field, the study of multi-media and web based contents and in the transportation industry, just to name a few. Apart from the aforementioned areas this book also covers the study of complex data found in meteorological and remotely sensed images acquired using satellite sensing. Data management and envelopment analysis has also been covered. The application of SOM in mechanical and manufacturing engineering forms another important area of this book. The final section of this book, addresses the design and application of novel variants of SOM algorithms.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Marek B. Zaremba (2011). Automated Mapping of Hydrographic Systems from Satellite Imagery Using Self-Organizing Maps and Principal Curves, Self Organizing Maps - Applications and Novel Algorithm Design, Dr Josphat Igadwa Mwasiagi (Ed.), ISBN: 978-953-307-546-4, InTech, Available from: http://www.intechopen.com/books/self-organizing-maps-applications-and-novel-algorithm-design/automated-mapping-of-hydrographic-systems-from-satellite-imagery-using-self-organizing-maps-and-prin

INTECH

open science | open minds