Determining The Skeletal Description Of Sparse Shapes

Rahul Singh[†]   Vladimir Cherkassky[‡]   Nikolaos P. Papanikolopoulos[†]

†Artificial Intelligence, Robotics, and Vision Laboratory
Department of Computer Science
‡ Department of Electrical Engineering
University of Minnesota, Minneapolis, MN 55455

Abstract

A variety of techniques in machine vision involve representation of objects by using their shape skeleton. Many algorithms have been proposed to date for obtaining the skeletal shape of digital images. The noise models predominantly used in these techniques are restricted to boundary noise. In particular, instances of noise occurring inside object regions and causing their non-contiguity are precluded. In this paper we present a method to obtain the skeletal shape of binary images in the presence of both boundary noise and noise occurring inside object regions. We propose to obtain the skeletal shape of such images by a modified version of the Kohonen self-organizing map, implemented in a batch processing mode. The modifications allow the map to adapt to the input shape distribution. At each iteration, a competitive Hebbian rule is used to progressively compute the Delaunay triangulation of the shape. Information from the triangulation augments the map topology to yield the final skeletal shape. The batch mode implementation of the self-organizing process, allows our approach to compare very favorably, in terms of computational time, with the traditional flowthrough implementations. Encouraging experimental performance has been obtained on a variety of shapes under varying signal to noise ratios.

1   Introduction

Shape analysis and recognition constitutes a fundamental problem of computer vision. One of the commonly used technique for representing shapes in vision systems involves the computation of the shape skeleton. Intuitively, a skeleton can be thought of as a *stick figure* that retains the connectivity information of a shape. The shape skeleton can often capture the essence of the visual form of an object and sig-

nificantly reduce the amount of data required for object representation. Applications of shape skeletons include among others, industrial part inspection, optical character recognition, and object representation. There are numerous algorithms for obtaining digital skeletons based on different definitions of the skeletal shape. A classification of these methods can be found in [11] where the following four classes in which these algorithms can be broadly grouped are identified: techniques based on the grassfire model, analytical techniques, distance transformations and topological thinning algorithms. It may be noted, that the noise models (if any) considered by these algorithms, are restricted to noise distributions on object boundaries. In Figure 1, we present two instances of the handwritten numeral two. The figure on the right has an internal noise level of 3.01dB (50 % noise). By comparing it with its counterpart on the left, which has no internal noise, we can observe the *sparseness* caused by internal noise within binary image regions. Such noise can occur due to pixel subsampling or poor quantization. Even though we find no difficulty in forming a perceptual idea of the skeletal shape for the noisy image in Figure 1, skeletonization of such images in the sense of the aforementioned four classes of techniques is an ill posed problem: in the grassfire model, the anisotropy introduced by sparseness inside regions would hinder the wavefront propagation. Analytical techniques on the other hand assume connectivity properties of the image regions which may not remain valid as the image becomes sparse. In the case of algorithms based on distance transforms, computation of the minimum distance from every object pixel to the background is required. However, a rigorous definition of *foreground* and *background* is difficult to formulate for sparse regions. Similar problems may occur during the application of topological thinning techniques since a pixel may not lie on the boundary of a sparse region and yet have zero valued pixels in

its neighborhood. Before presenting our approach we point out several properties that may be desirable in a prospective solution methodology to this problem.

- The method should be invariant to translation, rotation and scale changes of the shape.

- The results of the method should have the fundamental properties of a shape skeleton: connectedness and thinness.

- The method should have a reasonable computational performance.

- The output of the method should degrade gracefully with increase in the sparseness of the image.

- The results produced by the method, when applied to non sparse images, should not differ substantially from those obtained with conventional algorithms.

- The shape skeleton produced should correspond to our intuitive notions of shape.

We begin this paper by outlining an overview of our methodology and the motivations leading up to it (Section 2). We then present some previous work in the area of skeletonizing sparse images (Section 3). In Section 4 the proposed method is formulated. The experimental performance of the approach is studied on a variety of shapes in Section 5. In Section 6 we present the conclusions obtained from this research and outline possible future directions of work.

2 Overview of the Approach

The problem of obtaining a skeletal representation for a given image (sparse or otherwise), can be interpreted as a particular case of the more general problem of compactly representing a given data set. The conventional approach to this problem consists in either maintaining the dimensionality of the data set but reducing its size (i.e., clustering) or maintaining its size but representing it in a transformed space having reduced dimensionality (e.g., by principal component analysis). Neither of these methods can be directly applied to our problem. Clustering, for instance, cannot be used by itself to obtain topological properties like connectedness of the input shape, while principal components usually fail to provide a sufficiently accurate description of the highly non-linear shape distributions that usually occur in nature. An interesting non-linear extension of the principal components,
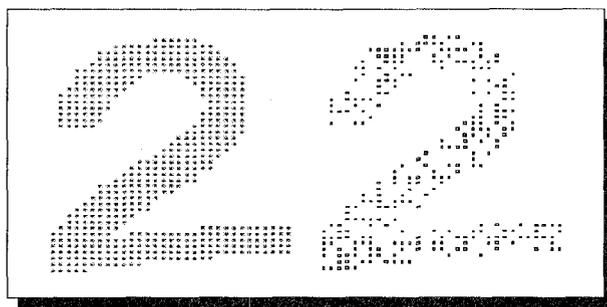

Figure 1: A nonsparse image and its sparse counterpart.

called the principal curves, was pioneered in statistics [5] to obtain smooth, curve based representation of patterns exhibited by scatter plots. For a given d-dimensional data distribution, its principal curve is a smooth uni-dimensional curve that passes through the middle of this distribution. The principal curve for a given distribution is self consistent in the sense that for any point p_i on the curve, the average of all data points projecting on p_i coincides with p_i. In the context of the two fundamental approaches enumerated above, the principal curve can be used for non linear dimensionality reduction. In [13, 10], the authors have observed that the Kohonen self-organizing procedure can be considered to provide a discrete approximation of the principal curve for a given data distribution. The approach we propose in this paper is based on this conclusion and consists of obtaining the skeletal shape of a given shape distribution by using the Kohonen self-organizing map (SOM). The skeletal description thus obtained is parameterized by the topological coordinates of the SOM units. The basic map topology employed in our method is the minimum spanning tree on the map units. The minimum spanning tree provides a flexible topology to represent the input space and its choice is based on empirical results [6] which show that the minimum spanning tree topology is well suited to approximate structured shape distributions. Since a spanning tree is insufficient to describe topological properties like closure of circular regions, we augment the tree structure of the map after convergence with proximity information obtained from the Delaunay triangulation of the shape. The batch processing SOM algorithm used in our work conceptually relates the shape structure obtained from self-organization to the mathematically well founded concept of principal curves [10], provides computational speedup, and allows easy determination of the Delaunay triangulation at each iteration.

369

3 Previous Work

The literature on methods to define a skeletal shape for sparse images is not very rich. There has been some work in identifying skeletonizing points for sparse images by density based clustering [1] and neural models [2]. However, the results from these methods constitute a disconnected set of points and the methods provide no information as to how these points may be connected. Recently two new methods [3, 4] have been proposed for obtaining skeletal shapes from sparse patterns. Chen and Yu [3] propose an entropy based approach in which for each point in the image a circular range of maximum effective information is defined. The symmetry score of the point distribution in this circular range is then computed. Following this, the symmetry score is treated as a gray-scale image and gray-scale thinning is performed on it to obtain the skeletal lines. The work of Datta et. al. [4] constitutes the motivation behind our approach. In [4], the authors obtain a piecewise linear approximation of the skeletal shape of a pattern by using a flowthrough version of the Kohonen self-organizing map. The map is initialized with a linear topology and evolves to approximate one of the three shape classes: arc patterns, forked patterns, and circular patterns. The transformation of the linear pattern to one of the above classes is done by threshold based criteria which include distance between two map units (for evolving from linear to circular patterns) and angle between three map units (for evolving from linear to forked patterns).

Our work essentially constitutes an alternate approach to [4] within the common framework of self-organization. Since [4] uses a flowthrough SOM, the final results may be influenced by the order of presentation of the data [10]. This renders the results obtained with the method susceptible to rotations in the image plane. This is avoided in our approach by using the batch SOM algorithm. Since batch processing involves the entire data set, problems related to the order of presentation are not involved. Furthermore, the use of a minimum spanning tree topology provides us with a natural way to span the whole input shape distribution, thus obviating the need to formulate criteria for transforming the map topology between different shape classes. Finally, the batch processing SOM algorithm makes the approximation of the skeletal shape representation much faster (in terms of the number of iterations on the data set).

We would like to briefly point to the reader that the general problem of defining the shape of a sparse pattern (also called a point set) has attracted a lot of attention both in the areas of Machine Vision and Com-

putational Geometry. Describing the shape of such patterns by their Convex Hull being an example of a well known method. However most of the techniques in these areas [12] do not attempt to define the inner shape of patterns.

4 The Proposed Method

4.1 The Map Update Procedure

The Kohonen self-organizing map consists of a set of units and an interconnection topology which connects these units. Each unit in the map is assigned a weight vector having the same dimensionality as that of the input space. A neighborhood function is used to define a region of influence around each unit based on the map topology.

Let $\mathbf{S} = \{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_\mathcal{D}\}$ be the set of input vectors and $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_\mathcal{U}\}$ be the the units of the SOM respectively. Let also $\mathbf{W}_i$ denote the weight vector corresponding to the unit $\mathbf{u}_i$ of the map. Let the input vector $\mathbf{d}_j$ be presented to the map at the time instant t. Furthermore let $\mathbf{u}_k$ with the weight vector $\mathbf{W}_k$ be the closest unit (in L_2 norm), to the given input. The weight vectors of $\mathbf{u}_k$ and all units in its neighborhood are then updated as

$$\mathbf{W}_k(t+1) = \mathbf{W}_k(t) + \beta(t)\mathbf{C}_j(i(k), k)[\mathbf{d}_j - \mathbf{W}_k(t)] \quad (1)$$

where t denotes the time. $\mathbf{C}_j(i(k), k)$ is the neighborhood defined on the map topology around the unit $i(k)$. $\beta(t)$ is the learning rate which decreases with time. The updates are repeated till the change in the weight vectors becomes smaller than a predetermined threshold. This often requires many presentations of the entire data set. The effect of the update process is to *pull* the closest weight vector and those in its neighborhood towards the input in a topologically constrained manner, ultimately leading to an orderly approximation of the unknown distribution of the input vectors. The above description of the SOM update procedure constitutes the flowthrough version of the algorithm. The application of the flowthrough version to the problem of obtaining skeletal representation of sparse images is limited by its slow training time and dependence on the order of presentation of the input data. Since most vision systems follow a standard image entry procedure (e.g., raster scanning), any rotation in the input image would cause a change in the data presentation order leading to possibly different skeletal descriptions for the same object. The solution to both these problems can be obtained by using the
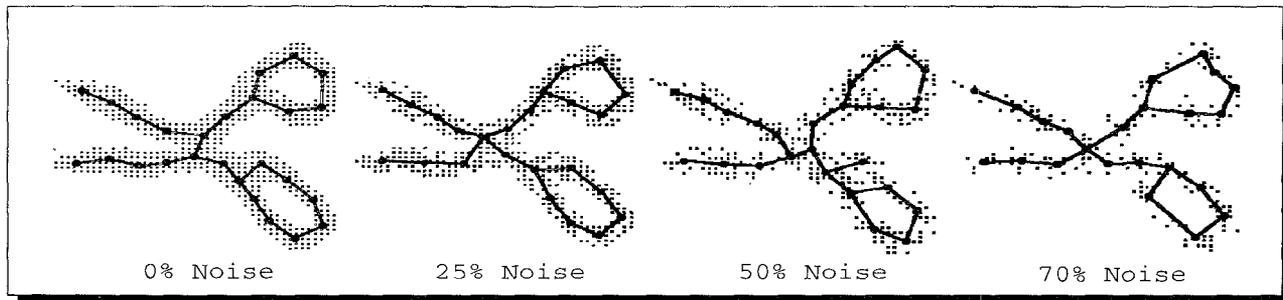
Figure 2: Skeletal shape extraction on a pair of scissors under decreasing signal to noise ratio.

batch formulation of the SOM algorithm [7, 8]. Empirical tests show that batch processing provides faster training time [7]. Furthermore since batch processing works with the entire data set at the same time, order of presentation of the data has no influence on the final result [10].

Let the location of unit j in the sample space be $\mathbf{W}_j$ and let its coordinate location in the topological space of the map be $\tilde{\mathbf{j}}$. Let $\mathcal{D}$ and $\mathcal{U}$ be the size of the data set and the number of units in the map respectively. The SOM update algorithm in the batch mode consists of the iteration of the following two steps with progressively decreasing neighborhood size:

1. *Voronoi Partition of the input data*: The data is partioned into Voronoi regions of the units. The centroid of each Voronoi region is computed along with its size. The partitioning and centroid computation is done by the Lloyd algorithm (K-means algorithm).

2. *Weighted centroid update in the topological space*: The units are updated by a weighted centroid of the data. The weights of each datum are determined by the voronoi region it belongs to and correspond to the neighborhood function of the flowthrough SOM update formulation. The formula used for centroid update is [10]:

$$\mathbf{W_j} = \frac{\sum_{p=1}^{\mathcal{U}} C(\tilde{\mathbf{p}} - \tilde{\mathbf{j}})\mathcal{M}_p \mathcal{S}_p}{\sum_{p=1}^{\mathcal{U}} C(\tilde{\mathbf{p}} - \tilde{\mathbf{j}})\mathcal{S}_p} \qquad (2)$$

where $\mathcal{M}_p$ is the centroid of the Voronoi region defined by map unit p and $\mathcal{S}_p$ is the size (in terms of the number of samples) of this region. $C(.)$ is a monotonically decreasing neighborhood function defined in the topological space. In the present work, a Gaussian was used as the neighborhood function.

Conceptually, a single iteration of the batch mode SOM update procedure can be thought to consist of a single iteration of the Lloyd algorithm followed by a kernel smoothing operation on the centroids in the topological space.

4.2 The Map Topology

In the original SOM formulation proposed by Kohonen, the number of processing units and their interconnection topology is fixed. The traditional topologies used with SOM (*linear* and *grid*) however, may not suffice to approximate the variation in natural shape distributions. The use of a minimum spanning tree to define the neighborhood relationships in cases where the input vector distribution is structured, has been suggested in [6]. This paper suggests assigning arcs between nodes, where the length of the arcs is defined as the distance between the nodes in the input space. By definition, the sum of the length of all the arcs in the tree is minimal. The neighborhood of a unit in this topology is defined in terms of the arcs emanating from the unit. The use of a spanning tree as the topology for the map during the batch update procedure provides an intrinsic mechanism to represent various planar shape topologies (closed circular shapes being the exception). The minimum spanning tree is recalculated after each iteration of the self-organization. The ability of a SOM to approximate a given shape distribution however depends not only on the basic map topology, but also on the number of processing units in the map. Since it is not possible to determine *a priori* the number of units required to approximate an input shape distribution, a procedure for automatically adding or deleting units in the map needs to be adopted. A principal constraint governing the addition and/or deletion of map units in our problem domain is that the existing skeletal structure should not be excessively perturbed by the change in the number of units. We follow the approach suggested in [4] and add a new unit in the middle between two existing units if the distance between their weight vectors exceeds a threshold $\delta_{\mathbf{max}}$. Similarly two existing
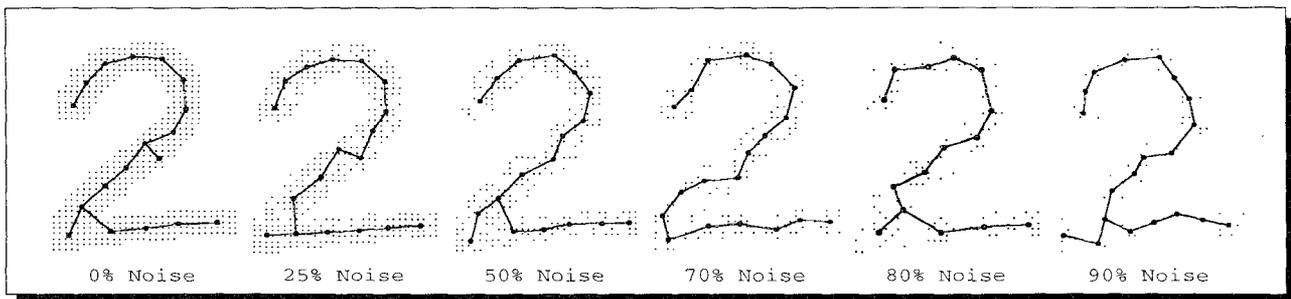
371

Figure 3: Skeletal shape extraction for a handwritten numeral. The shape distribution, prior to noise addition, was obtained by dilating the original thin handwritten shape pattern.

units are merged into a single one if the distance between their weight vectors falls under another threshold $\delta_{\min}$. The unit obtained after merging is placed at the midpoint of the line joining the two older units. Since the addition or removal of units is restricted to the existing skeletal structure, sudden changes in the shape skeleton are avoided. For a given shape distribution skeletal descriptions of varying details can be obtained by suitably tuning the parameters $\delta_{\min}$ and $\delta_{\max}$. The average length of the minimum spanning tree (L) on the data set can be used to obtain values for these parameters. In our experiments visually pleasing skeletons were obtained by using $\delta_{\min} \approx 3\sqrt{L}$ and $2\delta_{\min} \leq \delta_{\max} \leq 3\delta_{\min}$. The closure of circular regions is an important property that cannot be described due to the inherent limitations of the spanning tree based SOM topology. In order to augment the SOM, so as to produce faithful skeletal representation for circular regions, the proximity relationship between the Voronoi regions of the shape distribution is needed. Such a relationship can be obtained by computing the Delaunay triangulation of the Voronoi regions. We use a Hebbian rule proposed in [9] to compute the Delaunay triangulation after each iteration of the batch SOM algorithm. Essentially the method proposed in [9] joins Voronoi centroids (map units) $\mathbf{V}_i$ and $\mathbf{V}_j$, if for an input datum $\mathbf{d}_k$, the two closest centroids are $\mathbf{V}_i$ and $\mathbf{V}_j$. The authors in [9] prove that the triangulation thus obtained is topology preserving (in the sense that adjacency of centroids in the Delaunay triangulation implies adjacency of their respective Voronoi regions in the input space and vice-versa), if for every datum $\mathbf{d}_k$ from the shape manifold $\mathcal{S}$, the triangle $\triangle(\mathbf{d}_k, \mathbf{V}_i, \mathbf{V}_j) \subseteq \mathcal{S}$. Since in our approach, the minimum spanning tree is computed on the map units, which progressively tend to lie on the input shape distribution (manifold), it follows that the skeletal shape we obtain is a subset of the Delaunay triangulation obtained by the Hebbian rule of [9]. Skeletal representation of circular regions can therefore be obtained

by joining two disjoint units in the SOM, if there is an edge between them in the Delaunay triangulation. The occurrence of small loops in the skeleton can be prevented by allowing two map units to join only if the resultant cycle has more than K edges. We use the value of $K = 3$ in all our experiments.

5 Experimental Results

The proposed method was tested with two sets of data. The first set consisted of binary images of planar industrial parts and the second set consisted of handwritten letters and numerals captured by using a graphics tablet. Each thin pattern in the second set was dilated to obtain a thickened shape. The data in the second set was similar to that used in [3] (obtained by dilating thin patterns). We note that such data tends to be much more uniform in thickness than the shapes considered in the first set. A predetermined amount of random noise, uniformly distributed in the bounding box containing the image, was added to the image regions. The amount of the noise added was computed by counting the number of black pixels which were changed. Figure 2 exhibits the performance of the method on a pair of scissors. The noise is varied between 0% to 70% (1.55dB) of the original number of points (pixels) in the shape. The performance of our method on the dilated shape of a handwritten numeral is shown in Figure 3. Typically, for shapes from this class, our method produced skeletal descriptions consistent with the visual form of the data beyond SNR of 1.55dB (noise levels of 70%). The performance of our method on the dilated shape of a handwritten numeral for noise level upto 90% (0.46dB) is shown in Figure 3. The algorithm suggested in [3], was reported to maintain good performance in experiments till a signal to noise ratio (SNR) of 2.22dB (60%noise). Beyond this, the performance of the algorithm deteriorated rapidly. In Figure 4, the results obtained by applying the proposed approach
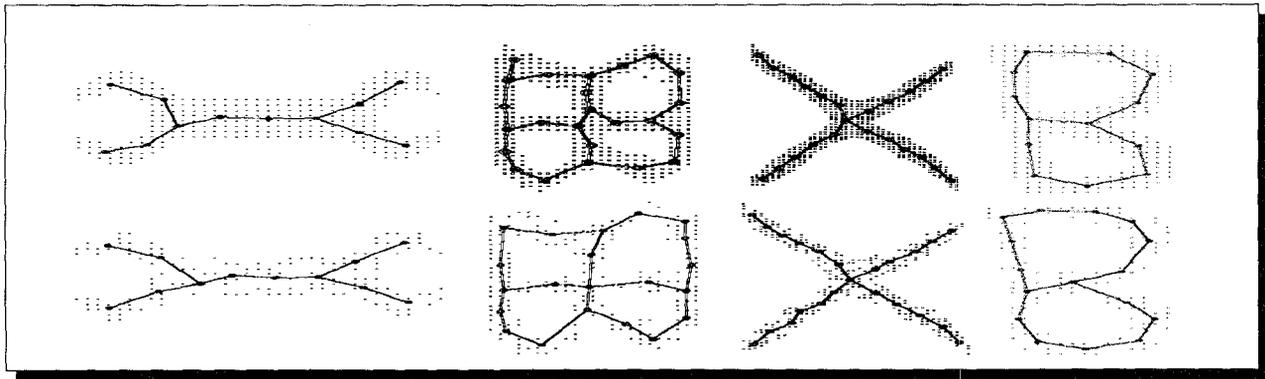
372

Figure 4: Results for some representative shapes.

to four different shapes are presented. The top row is comprised of the original planar shapes with no added noise (0% noise). The images in the bottom row were obtained by adding 50% (SNR = 3.01dB) uniformly distributed random noise to the corresponding shapes from the top row.

6 Conclusions and Future Work

In this paper, we have proposed a method to obtain skeletal shape descriptions for sparse planar shapes. Owing to the sparseness of the shape regions, conventional skeletonization techniques cannot be applied to such patterns. The approach suggested by us is based on a batch mode SOM. The SOM uses a minimum spanning tree topology to approximate the input shape distribution. In conjunction with the iterations of the SOM a Hebbian rule is used to compute the Delaunay triangulation of the Voronoi centroids of the input. The topology of the SOM on convergence is modified, if necessary, by adding edges between the map units based on information obtained from the Delaunay triangulation. The method is invariant to translation, rotation and scale changes in the input. Empirical results suggest that it performs robustly under varying signal to noise ratio. In the present formulation, the performance of the method depends on certain parameters $(\delta_{min}, \delta_{max})$, which essentially have to be provided manually. Our further research is directed towards resolving these parameters using shape statistics.

Acknowledgements

References

[1] D. Chaudhuri, B. B. Chaudhuri, and C. A. Murthy. "A Data Driven Procedure for Density Estimation with some Applications". *Pattern Recognition*, 29(10):1719–1736, 1996.

[2] Y. S. Chen, C. T. Wu, and W. H. Hsu. "A Neural-like Model Capably Detects the Structural Points from Line Patterns Having Noises". In *Proceedings of the 3rd International Conference on Visual Search*, 1992.

[3] Y. S. Chen and Yu. T. Yu. "Thinning Approaches for Noisy Digital Patterns". *Pattern Recognition*, 29(11):1847–1862, 1996.

[4] A. Datta, S. K. Parui, and B. B. Chaudhuri. "Skeletal Shape Extraction from Dot Patterns by Self-Organization". In *Proceedings of the 13th International Conference on Pattern Recognition*, volume 4, pages 80–84, 1996.

[5] T. Hastie and W. Stuetzle. "Principal Curves". *Journal of the American Statistical Association*, 84(406):502–516, 1989.

[6] J. A. Kangas, T. K. Kohonen, and J. T. Laaksonen. "Variants of Self-Organizing Maps". *IEEE Transactions on Neural Networks*, 1:93–99, 1990.

[7] T. K. Kohonen. "Things You Haven't Heard About the Self-Organizing Map". In *Proceedings of the IEEE International Conference on Neural Networks*, volume 3, pages 1147–1156, 1993.

[8] S. P. Luttrell. "Derivation of a Class of Training Algorithms". *IEEE Transactions on Neural Networks*, 1(2):229–232, 1990.

[9] T. Martinetz and K. Schulten. "Topology Representing Networks". *Neural Networks*, 7(3):507–522, 1994.

[10] F. Mulier and V. Cherkassky. "Self-Organization as an Iterative Kernel Smoothing Process". *Neural Computation*, 7(6):1165–1177, 1995.

[11] R. L. Ogniewicz and O. Kubler. "Hierarchic Voronoi Skeletons". *Pattern Recognition*, 28(3):343–359, 1995.

[12] F. P. Preparata and M. I. Shamos. *Computational Geometry*. Addison–Wesley, Reading, Massachusetts, 1992.

[13] H. Ritter, T. Martinetz, and K. Schulten. *"Neural Computation and Self-Organizing Maps: An Introduction"*. Addison-Wesley, Reading, Massachusetts, 1992.