Object Skeletons From Sparse Shapes In Industrial Image Settings

Rahul Singh† Nikolaos P. Papanikolopoulos† Vladimir Cherkassky‡

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

Abstract

This paper presents a new method for computing the shape skeleton of planar objects in presence of noise occurring inside the image regions. Such noise may be due to poor control of lighting conditions, incorrect thresholding or image subsampling. Binary images of objects with such noise exhibit sparseness (lack of connectivity), within their image regions. Such non-contiguity may also be observed in thresholded images of objects which consist of regions having varying albedo. The problem of obtaining the skeletal description of sparse shapes is ill posed in the sense of conventional skeletonization techniques. We propose a skeletonization method which is based on obtaining the shape skeleton by evolving an approximation of the principal curve of the shape distribution. Our method is implemented as a batch mode Kohonen self-organizing map algorithm and involves iterating the following two steps: (1) *Voronoi tessellation of the data,* (2) *kernel smoothing on the Voronoi centroids. Adjacency relationships between the Voronoi regions are obtained by computing a Delaunay triangulation of the centroids. The Voronoi centroids are connected by a minimum spanning tree after each iteration. The final shape skeleton is obtained by joining centroids which are disjoint in the spanning tree, but have adjacent Voronoi regions. The skeletal descriptions obtained with the method are invariant to translation, rotation, and scale changes of the shape. The potential of the method is demonstrated on industrial objects having varying shape complexity under different imaging conditions.*

1 Introduction

Shape recognition is a fundamental task of computer vision with applications in a variety of contexts, including manufacturing, inspection, modeling, target tracking and robotic-grasping. A central issue in any shape recognition formulation is the framework used for shape representation. In this context, the medial axis or the shape skeleton has been a popular technique. In general, skeletal descriptions incorporate the advantages of both region based and boundary based descriptors and provide a compact and often highly intuitive shape description. Shape skeletons have been applied to object recognition and representation [12, 14], industrial part inspection [9], inspection of printed circuit boards [13], analysis of medical imagery [3] and optical character recognition. A major difficulty in creating robust systems for object recognition and representation has been the sensitivity of most vision techniques to changes in imaging conditions. For instance, small changes in illumination or viewpoint often cause disproportionately large changes in the appearance of the object. In Figure 1, we show a variation of this problem which is of interest to us in the present work. The upper row of the figure consists of three images of a tool taken under varying lighting conditions. The lower row comprises of the corresponding binary images obtained after moment preserving thresholding. One can observe the *sparseness* introduced in the image regions due to the the change in lighting. In this paper, we look at the problem of computing skeletal description of such shapes. It may be brought to the notice of the reader that in addition to improper illumination, factors such as incorrect thresholding, camera noise, as well as image subsampling may lead to such sparseness. While comparatively small amounts of sparseness in images can often be dealt with, by using filtering techniques or a sequence of dilation-erosion operations, such methods, however, become ineffective when the level of sparseness in the image is large. In this paper we present a new approach for obtaining skeletal shapes from sparse images. The method is based on obtaining a skeletal description by approximating the *principal*

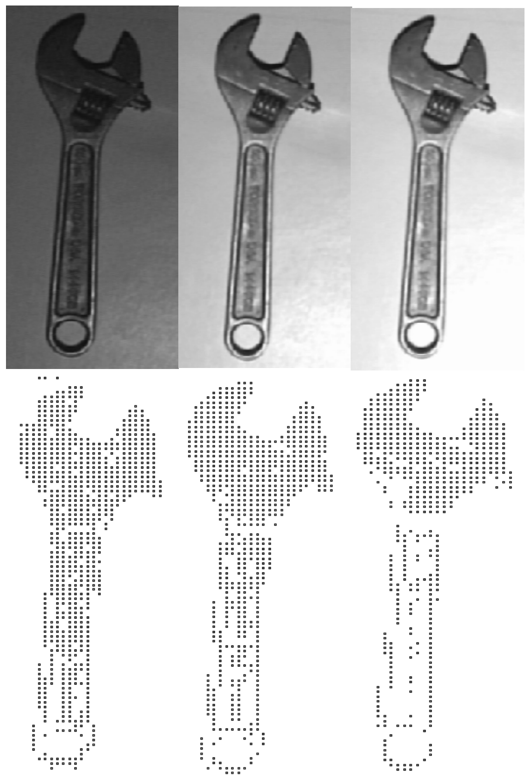

Figure 1: The top row of the figure shows three images of taken under different illumination. The bottom row consists of the corresponding thresholded images.

curve of the shape distribution. The concept of principal curves was introduced in Statistics [5], to summarize patterns exhibited by points in a scatter-plot. The principal curve is a non-linear generalization of the first principal component and is defined for both sparse and non-sparse distributions.

We begin this paper by analyzing the difficulties encountered in skeletonizing sparse shapes using conventional techniques and present an overview of the prior research in skeletonization of sparse shapes (Section 2). The proposed method is formulated in Section 3. Results of applying the method to various shapes are reported in Section 4. The conclusions of the present research and possible directions of future work are reported in Section 5.

2 Skeletonization of Sparse Shapes

Various skeletonization algorithms can be broadly grouped into the following four classes [10]:

1. Techniques based on the grassfire model.
2. Analytical techniques.
3. Distance transforms.
4. Topological thinning algorithms.

Sparse images of objects are characterized by anisotropy and non-contiguity of the shape regions. Furthermore, it is difficult to formulate rigorous criteria, to distinguish the foreground and background in such shapes. Due to these reasons, a direct application of conventional skeletonization techniques is precluded for sparse shapes. For instance, in the grassfire model, the skeleton is defined as the location where the propagating wavefront initiated on the shape boundary intersects itself. For sparse shapes it is not only difficult to define the shape boundary, but the very rate of wave propagation may be influenced by the anisotropy within the regions. Analytical techniques, on the other hand, are built around the concepts of continuity and connectivity of the shape regions. These properties often do not hold for sparse shapes. The application of distance transforms as well as topological thinning require formal criteria for separating the foreground from the background. Since a sparse shape might conceivably, be a distribution where the object pixels have neither 4-connectivity nor 8-connectivity, both thinning as well as distance transforms are inapplicable.

Recent efforts towards solving this problem include an entropy based method [1] and the use of neural models [2]. The entropy based technique of Chen and Yu [1] is based on computing the circular range containing the maximal information for each pixel. The symmetry score of the pixel distribution in the circular range in then computed. The symmetry information is treated as a grey-scale image which is then thinned to obtain the shape skeleton. The method proposed by us in this paper is motivated by the work of Datta et. al. [2], on using a flow-through self-organizing map (SOM) to obtain a shape skeleton. In the method of Datta et. al., the SOM is initialized with a linear topology and evolves to arc patterns, forked patterns or circular patterns based on thresholds on the angle formed at each map unit by its neighbors (for changing from linear to forked patterns) and the distance between two map units (for evolving from linear to circular patterns). Our approach differs from [2] in the following fundamental aspects:

- We seek to obtain the skeleton as an approximation of the principal curve of the shape distribution. The batch mode SOM algorithm used by us is more closely related to the concept of principal curve than the traditional flow-through version. Furthermore, in contrast to the batch formulation, the results of using a flow-through SOM are susceptible to the order in which data is presented [8]. Skeletal descriptions obtained with a flow-through SOM are therefore not invariant to rotations in the image plane.

- The evolution of the skeleton in our method is guided by a minimum spanning tree on the Voronoi centroids (map units). The use of a spanning tree topology provides us with a natural way to span the input shape distributions, without resorting to threshold based criteria.

- Skeletons reflect topological properties of shapes like adjacency and closure of circular regions. The Kohonen SOM does not necessarily form topology preserving representations [7]. We obtain the topological properties of the shape distribution by computing its Delaunay triangulation. The minimum spanning tree based skeleton is a subset of this triangulation [11]. In our approach, after convergence, the spanning tree based skeleton is augmented with proximity information from the Delaunay triangulation to give the final skeletal shape.

3 The Proposed Method

3.1 Principal Curves, Shape Skeletons and the Batch-mode SOM

Principal curves are smooth, non-linear generalization of principal components [5]. For a given data distribution, its principal curves are smooth, self-consistent curves that pass through the *middle* of the distribution. The self-consistency of the principal curve implies that for any point on the principal curve, the average of all data points projecting onto it, coincides with this point on the curve. Figure 2, shows a principal curve fit to a data set. The principal curve algorithm consists of the following steps:

1. *Projection step*: For each data point, find its projection on the curve. The projection is determined by finding the closest point on the curve (in the orthogonal sense).

2. *Conditional-Expectation step*: Scatterplot smooth the projected values along the curve length.

3. Decrease the scatterplot smoothing span and iterate steps 1 and 2 till convergence.

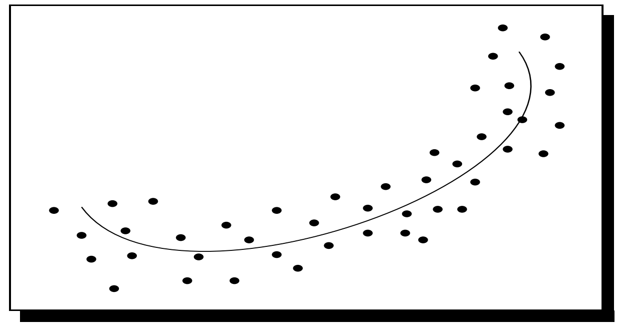

Figure 2: Principal curve fit on a data set.

The relation of the principal curve, to the medial axis of a shape, follows from its property of self-consistency. Furthermore, the non-linear nature of the principal curve is important in obtaining smooth, unidimensional representations of (possibly) high dimensional, complex distributions. Since principal curves are well defined for both dense (4 or 8 connected) and sparse data, skeletal descriptions based on them can be obtained for objects both in the setting being considered in this paper (sparse shapes), as well as for the more conventional non-sparse images.

For a given distribution, a discrete approximation of its principal curve can be obtained by using the batch formulation of the SOM algorithm. The representation of the curve thus obtained is parametrized according to the topological coordinates of the SOM units.

Let $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{\mathcal{D}}\}$ be the set of input vectors (data), and $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{\mathcal{U}}\}$ be the units of the SOM respectively. Also let the location of unit $\mathbf{u}_j$ ($j \in [1,\ \mathcal{U}]$), 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 $\mathcal{D}$ denote the size of the data set and $\mathcal{U}$, the number of units in the map respectively. The batch mode SOM algorithm consists in iterating the following two steps:

1. *Voronoi Tessellation 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.

$$\arg\min_{j} \left(\parallel \mathbf{x}_k - \mathbf{W}_j \parallel \right) \quad k = 1, \cdots, \mathcal{D}$$

Conceptually this step may be considered as a single iteration of the Lloyd vector quantization 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:

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

Where $\mathcal{M}_p$ is the centroid of the Voronoi region defined by map unit $\mathbf{u}_p$ and $\mathcal{S}_p$ is the number of samples in 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. This step is equivalent to a kernel smoothing operation on the centroids in the topological space.

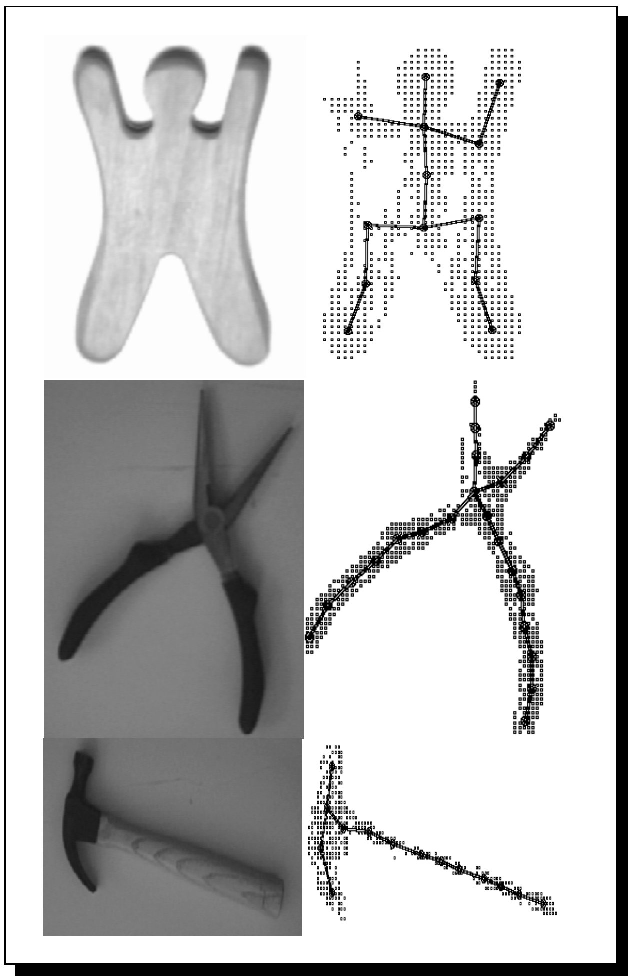

Figure 3: Results for some representative objects from the first experiment.

3.2 Evolution of the Skeletal Topology

The traditional topologies used with SOM (*linear* and *grid*) are not suitable to approximate the variation in natural shape distributions. An alternative is to use a minimum spanning tree on the map units, to define the neighborhood relationships in case of highly non-linear distributions [6]. A minimum spanning tree topology, assigns arcs between map units where the length of the arcs is defined as the distance between the nodes in the input space. By definition, the total length of all the arcs in the tree is minimal. For the purpose of extracting shape skeletons, the connectedness and thinness of the minimum spanning tree is highly useful. Additionally, the spanning tree topology provides an intrinsic mechanism to represent various shape topologies (with the exception of closed circular shapes).

The approximation of the skeletal representation for a given shape distribution depends not only on the map topology, but also on the number of map units. Since it is not possible to determine *a priori*, the number of map units for an unknown shape distribution, an adaptive procedure to change the map size is desirable. A smooth evolution of the skeleton can occur if the perturbation to the skeletal shape, during the addition and/or deletion of map units is minimal. To this end, we follow the approach suggested in [2] and add a new unit in the middle between two existing units $\mathbf{u}_i$ and $\mathbf{u}_j$ if $\| \mathbf{W}_i - \mathbf{W}_j \| > \delta_{\mathbf{max}}$. Similarly, two existing units $\mathbf{u}_k$ and $\mathbf{u}_l$ are merged into a single one if the distance between their weight vectors $\| \mathbf{W}_i - \mathbf{W}_j \| < \delta_{\mathbf{min}}$. For a given shape distribution skeletal descriptions of varying details can be obtained by suitably tuning the parameters $\delta_{\mathbf{min}}$ and $\delta_{\mathbf{max}}$. The average length of the minimum spanning tree (L) on the data set can be used to estimate values for these parameters. In our experiments we used the values $\delta_{\mathbf{min}} \approx 3\sqrt{L}$ and $2\delta_{\mathbf{min}} \leq \delta_{\mathbf{max}} \leq 3\delta_{\mathbf{min}}$.

While the spanning tree topology provides an intrinsic method to span complex shapes, it is unable to faithfully represent certain topological properties like the closure of circular regions. Closure of circular regions can be represented, by knowing the adjacency relations between the Voronoi regions of the shape. Since the shape manifolds we are dealing with are sparse, we define two Voronoi regions $\mathcal{V}_p$ and $\mathcal{V}_q$ with centroids $\mathcal{M}_p$ and $\mathcal{M}_q$, to be adjacent, if and only if, for any input $\mathbf{x}_j$, the two closest centroids are $\mathcal{M}_p$ and $\mathcal{M}_q$. We note that the above relation can be easily implemented by joining, for each datum, the two closest centroids during the Voronoi tessellation in the batch SOM algorithm. It can be shown, that the above procedure is a special case of the topology representing network algorithm [7] and forms a Delaunay triangulation of the shape manifold.

Skeletal representation of circular regions can 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.

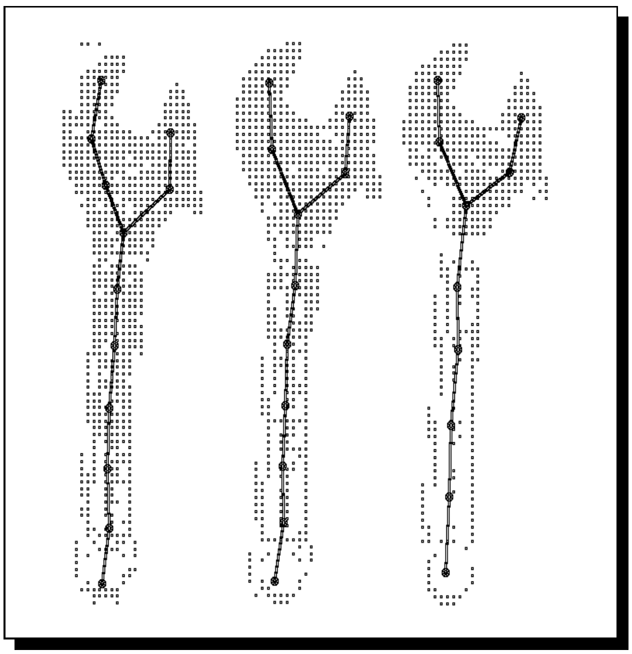

Figure 4: Skeletons for the object from Figure 1.

4 Experimental Results

Three sets of experiments were designed to test the performance of the system under varying conditions. In the first set of experiments the performance of the method was tested on a collection of objects under fairly well controlled illumination conditions. However, the objects were selectively placed on backgrounds with intensity spectra, similar to those of the object. Bi-level thresholding was performed for each object using both moment preserving thresholding and the Otsu algorithm (minimizing within group variance) [4]. The sparseness in the object regions, observed in this experiment was primarily due to poor thresholding. The skeletonization results for three objects are presented in Figure 3. The second experimental setting was designed to test the proposed approach on images of the same object, captured under varying illumination conditions. The skeletons for the object shown under varying lighting in Figure 1 are presented in Figure 4. In the final experiment, we tested the performance of the method under decreasing signal to noise ratio. A predetermined amount of random noise was generated in the bounding box containing the image. The amount of noise added to each image was computed by counting the number of black pixels (foreground pixels) which were changed. For each shape, the noise was varied from 0% to 75% (1.25 dB) of the original number of pixels in the shape. The results on two shapes from this experiment are presented in Figure 5.

5 Conclusions and Future Work

This paper presents a new method for skeletonizing sparse shapes. Due to the lack of contiguity in the image regions, conventional skeletonization techniques cannot be applied to such patterns. Our approach is based on obtaining the shape skeleton by approximating the principal curve of the sparse images. The approximations are computed by using a batch mode SOM. Both the topology and the size of the map are adaptively determined, based on the input shape distribution. Experimental results indicate that the method performs robustly under different illumination conditions as well as varying signal to noise ratio. In the present formulation, the algorithm depends on the $\delta_{\mathbf{min}}$ and $\delta_{\mathbf{max}}$ parameters. Their values are currently determined by using a global estimation technique. If the noise distribution has a local character, then such estimations are often poor. Our current research is aimed towards estimating these parameter values using local techniques.

Acknowledgements

The authors would like to acknowledge the help of Mike Wade in conducting some of the experiments. This research was supported by the NSF through Grants #IRI-9410003 and #IRI-9502245, the McKnight Land-Grant Professorship Program of the University of Minnesota, and the Department of Energy through Contracts #AC-3752D and #AL-3021.

References

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

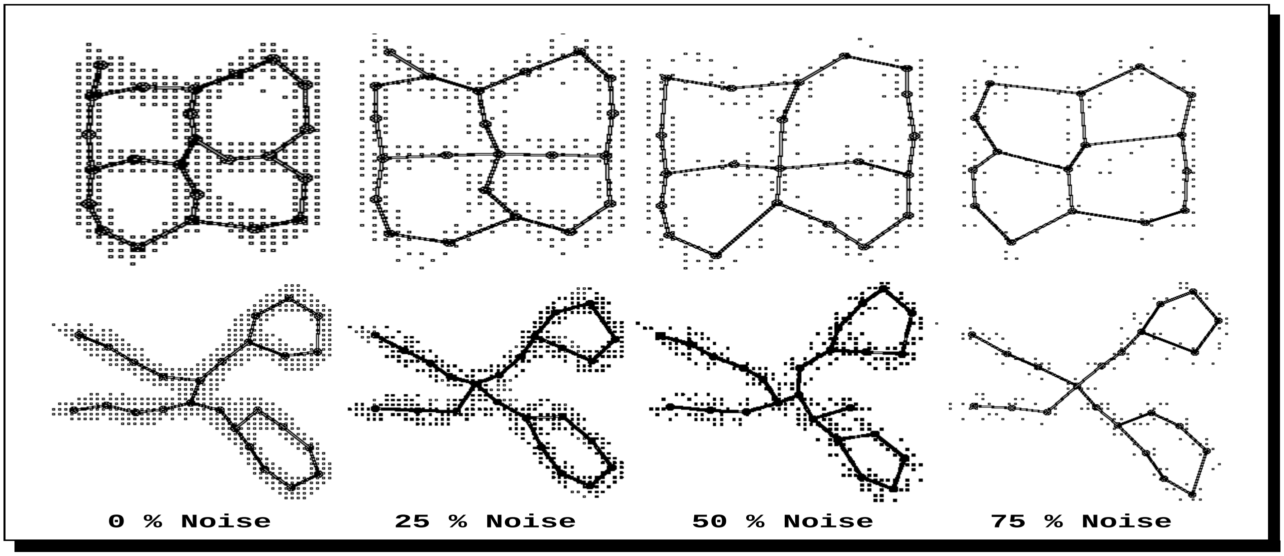

Figure 5: Skeletal shape extraction on a pair of objects under decreasing signal to noise ratio.

[2] 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.

[3] A. R. Dill, M. D. Levine, and P. B. Noble. "Multiple Resolution Skeletons". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):495–504, 1987.

[4] R. M. Haralick and L. G. Shapiro. *"Computer and Robot Vision "*. Addison -Wesley, Reading, Massachusetts, 1992.

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

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

[9] J. L. Mundy and R. E. Joynson. "Automatic Visual Inspection Using Syntactic Analysis". In *Proceedings of the International Conference on Pattern Recognition and Image Processing*, pages 144–147, 1977.

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

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

[12] F. Rolland, J. M. Chassery, and A. Montanvert. 3d medial surfaces and 3d skeletons. In C. Arcelli et. al., editor, *Visual Form*, pages 443–450. Plenum Press, New York, 1992.

[13] Q Z Ye and P. E. Danielsson. "Inspection of printed circuit boards by connectivity preserving shrinking". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5):737–742, 1988.

[14] S. C. Zhu and A. L. Yuille. "FORMS: A Flexible Object Recognition and Modelling System". *International Journal of Computer Vision*, 20(3):187–212, 1996.