Self-Organizing Maps for the Skeletonization of Sparse Shapes

Rahul Singh, Vladimir Cherkassky, and Nikolaos Papanikolopoulos

***Abstract*—This paper presents a method for computing the skeleton of planar shapes and objects which exhibit sparseness (lack of connectivity), within their image regions. Such sparseness in images may occur due to poor lighting conditions, incorrect thresholding or image subsampling. Furthermore, in document image analysis, sparse shapes are characteristic of texts faded due to aging and/or poor ink quality. Due to the lack of pixel level connectivity, conventional skeletonization techniques perform poorly on such (sparse) shapes. Given the pixel distribution for a shape, the proposed method involves an iterative evolution of a piecewise-linear approximation of the shape skeleton by using a minimum spanning tree-based self-organizing map (SOM). By constraining the SOM to lie on the edges of the Delaunay triangulation of the shape distribution, the adjacency relationships between regions in the shape are detected and used in the evolution of the skeleton. The SOM, on convergence, gives the final skeletal shape. The skeletonization is invariant to Euclidean transformations. The potential of the method is demonstrated on a variety of sparse shapes from different application domains.**

***Index Terms*—Image degradation, principal curves, self-organizing maps, shape skeleton, sparse shapes.**

I. Introduction

SHAPE analysis and recognition is one of the fundamental problems in pattern recognition, image processing, and computer vision. A central issue in shape analysis and recognition is the framework used for shape representation. In this context, the medial axis [2], or the shape skeleton of a two-dimensional (2-D) image has been a popular approach for shape representation. Intuitively, a skeleton can be thought of as a *stick figure* that retains the connectivity information of a shape. 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 [22], [25], industrial part inspection [17], inspection of printed circuit boards [24], analysis of medical imagery [7] and optical character recognition. Conventional skeletonization techniques implicitly assume connectivity of pixels inside image regions. In this paper we consider the problem of computing the shape skeleton for shapes that lack pixel level connectivity (henceforth, termed as *sparse shapes*). In Fig. 1, we present two instances of the handwritten numeral two. The figure on the right has an artificially introduced internal noise level of 3.01 dB (50% noise). By comparing it with its counterpart on the left, which has no internal noise, we can observe the *sparseness* within the binary image regions. Another example of images having sparseness in their regions is shown in Fig. 2. 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. Again, one may observe the sparseness introduced in the image regions due to a combination of factors including the change in illumination intensity, albedo of the surface, and the threshold value. In general, factors causing sparseness in images may be classified as extrinsic or intrinsic. Extrinsic factors are related to control of imaging conditions, thresholding, sensor noise, and image subsampling. Intrinsic factors, on the other hand, are related to the inherent nature of the available data. Examples of the latter include degraded texts due to fading or artificially generated distributions. Before presenting our approach we enumerate the following list of properties that are desirable in a solution methodology to this problem:

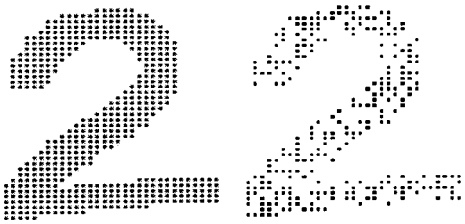

Fig. 1. A nonsparse image and its sparse counterpart.

- 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 nonsparse images, should not differ substantially from those obtained with conventional algorithms.
- The shape skeleton produced should correspond to our intuitive notions of shape.

It may be noted that the problem of computing the skeletal description of a sparse shape is a special case of the more general problem of describing the shape of a point set. The latter problem has received considerable attention in computational geometry where a variety of techniques has been developed to characterize both the external shape (convex hull, α-hull) and the internal shape (nearest neighbor graph, minimum spanning tree, relative neighborhood graph, Gabriel graph, and Delaunay triangulation) of a point set. For an overview of these descriptors

Manuscript received December 1, 1998; revised August 27, 1999. This work was supported by the NSF under Grants IRI-9410003 and IRI-9502245.

R. Singh and N. P. Papanikolopoulos are with the Artificial Intelligence, Robotics, and Vision Laboratory, Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455 USA.

V. Cherkassky is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA.

Publisher Item Identifier S 1045-9227(00)01198-X.

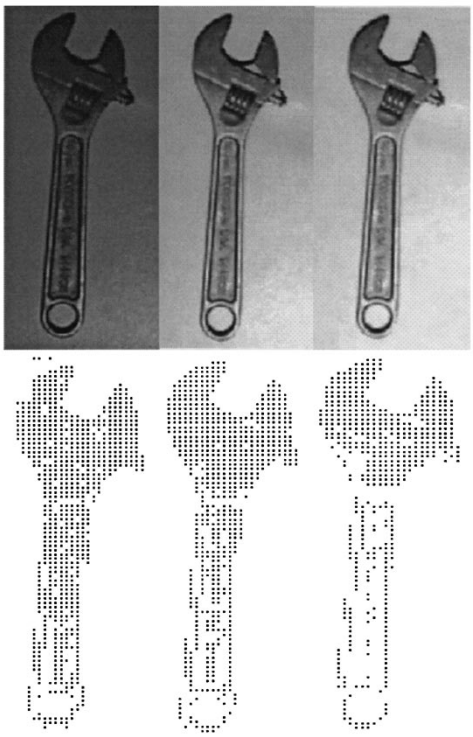

Fig. 2. (Top) three images taken under different illumination. (Bottom) corresponding thresholded (binary) images.

we refer the reader to [11] and [20] and references therein. Computing a skeletal description for such shapes is of particular interest not only because skeletons provide an intuitive shape description, but also because of the wide use of skeletons as shape descriptors and the possible applicability of such a technique in situations where the presence of sparseness causes conventional algorithms to perform poorly.

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 II). The proposed method is formulated in Section III. In Section IV some issues related to the evolution of the shape skeleton are discussed. Results of applying the method to various shapes are reported in Section V. Finally, the conclusions are given in Section VI.

II. Skeletonization of Sparse Shapes

Various skeletonization algorithms can be broadly grouped into the following four classes [18]:

1) techniques based on the grassfire model;
2) analytical techniques;
3) distance transforms;
4) topological thinning algorithms.

Sparse images of objects (as the examples in the previous section exhibit), are characterized by anisotropy and noncontiguity of the shape regions. For such shapes, a pixel may not lie on the boundary of a region and yet have zero valued pixels in its neighborhood. 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 difficult 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. Skeletonization of sparse shapes is thus difficult to handle in the framework of standard techniques.

Conventional image processing techniques that are generally used to ameliorate this problem include filtering techniques like median filtering or morphological operations such as dilation-erosion [8]. While median filtering ceases to be effective in images with large amounts of sparseness, the use of morphological operators introduce limitations that may preclude their use in certain applications. In Fig. 3, for instance, we present some results of performing a sequence of dilation-erosion operations on a word from a faded text. Note that not only there is no *a priori* way to determine the number of dilation-erosion operations that may be needed, but distortion of topological features as well as merging of adjacent letters may also occur. In the area of document image analysis, various other techniques have been proposed to deal with image degradation including sparseness. In [10], bitmap averaging was used to cancel out the noise and produce images with smooth outlines. This technique was based on the hypothesis that the noise introduced during scanning and transmission is random. Image degradations due to salt-and-pepper noise and partial omission of figures was handled in [19]. The method was based on producing gray level images from binary images by computing the density of black pixels in an $N \times N$ grid placed on each pixel. A similar approach was also used in [3]. Our work differs significantly from the above mentioned approaches in that, given a faded image, the issue of its shape representation is directly addressed by computing its shape skeleton. Unlike [10], we do not assume that the sensor noise is uniformly random. Furthermore, the skeletonization technique proposed by us is valid not only for sparse shapes, but may also be applied to conventional binary images. Recent efforts toward solving the problem of obtaining skeletons for images which may lack pixel-level connectivity include an entropy based method [4] and the use of neural models [6]. The entropy based technique in [4] 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 is then computed. This symmetry information is treated as a gray-scale image which is thinned to obtain the shape skeleton. In [6], a flow-through self-organizing map (SOM) is used to obtain the shape skeleton of a connected set of points. The SOM is initialized with a linear topology and evolves to arc patterns, forked patterns or circular patterns based

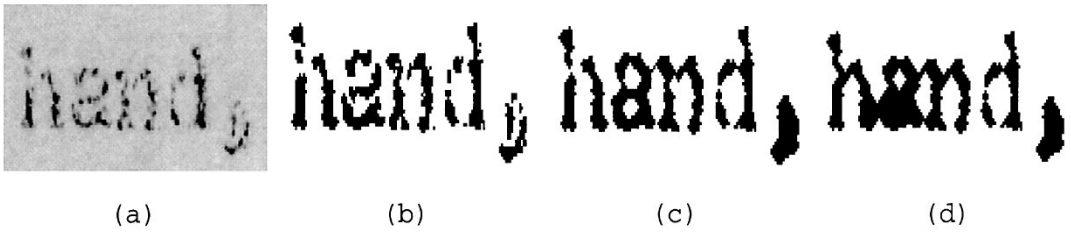

Fig. 3. Dilation-erosion on faded words: (a) gray-scale image, (b) after one dilation-erosion operation, (c) after two dilation-erosion operations, and (d) after three dilation-erosion operations.

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). While sharing the framework of self-organization with [6], our approach differs from it in the following significant respects. First, the use of a flow-through SOM, as in [6], may be sensitive to the choice of learning schedule and the order of presentation of the data [16]. This implies that the final result is not invariant to image plane rotations. By using the batch formulation of the SOM algorithm (see Section III), we circumvent these difficulties. Second, the shape skeleton, in our approach, is based on a data-driven minimum spanning tree topology. Unlike [6], the evolution of the skeleton does not require hand crafted rules. Finally, the convergence of a batch-SOM is much faster than the flow-through version.

III. The Proposed Method

In the proposed technique, computing the skeletal description of a sparse shape is based on approximating the *principal curve* of the shape distribution. The concept of principal curves was introduced in statistics [9]. Principal curves are smooth, nonlinear generalization of principal components and provide nonlinear dimensionality reduction. For a given data distribution, its principal curves are smooth self-consistent curves that pass through the *middle* of the convex distribution. This property links the principal curve to the medial axis of a shape. A direct application of the principal curve algorithm to compute a shape skeleton is, however, precluded, since by definition principal curves do not self-intersect and thus can not represent crossover of regions. Our solution to this problem is based on the following.

- For a given shape distribution, we obtain a discrete approximation of its principal curve by using the batch formulation of the SOM algorithm [5], [13], [21]. The approximation of the principal curve thus obtained is parameterized according to the topological coordinates of the SOM units. Invariance to Euclidean transformations follows, since in the batch method the final position of the map units is insensitive to the order of presentation of the data [15].
- Instead of the traditional fixed topologies used with SOM (*linear* and *grid*), a flexible minimum spanning tree (MST) on the map units is used to define the neighborhood relationships. Such a topology can provide better representation of structured (nonconvex) distributions [12]. In this topology, arcs are assigned between map units and 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 sections of the spanning tree between the branching points thus provide a piecewise approximation of the principal curve for the data points projecting onto these sections. Furthermore, the connectedness and thinness of the MST are ideal for skeletal representation. The use of a MST topology also obviates the need to define threshold based criteria as in [6] for evolution of the skeletal topology.

Let $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{\mathcal{D}}\}$ be the set of input vectors (data), and $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{\mathcal{U}}\}$ 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 $\mathbf{T}_j$. Let $\mathcal{D}$ denote the size of the data set and $\mathcal{U}$, the number of units in the map, respectively. The SOM units are randomly (and automatically) initialized in the sample space. The initialization range is determined by the bounding-box of the data. After the initialization, the batch-mode SOM algorithm consists in iterating the following steps with progressively decreasing span of the kernel function being used.

1) *Minimum Spanning Tree Computation*: Prim's algorithm is used to compute a MST on the SOM units in the sample space. The number of branches in the tree between any two units constitutes a topological distance measure.
2) *Voronoi Tessellation of the Input Data*: The data is partioned into Voronoi regions of the units by finding the closest center to each data point

$$\mathbf{T}_p = \arg\min_j (\|\mathbf{x}_p - \mathbf{W}_j\|)\ p = 1, \cdots, \mathcal{D}.$$

Conceptually, this step may be considered as a single iteration of the Lloyd vector quantization algorithm (K-means algorithm).
3) *Conditional-Expectation*: The conditional expectation is determined using the kernel regression estimate

$$\mathbf{W}_j = \frac{\sum_{p=1}^{\mathcal{D}} \mathbf{x}_p C(\mathbf{T}_j - \mathbf{T}_p)}{\sum_{p=1}^{\mathcal{D}} C(\mathbf{T}_j - \mathbf{T}_p)}, \qquad j = 1, \cdots, \mathcal{U} \tag{1}$$

where $C(.)$ is a monotonically decreasing kernel function defined in the topological space. The kernel function is defined in terms of the distance between the units in the topological space (number of branches between them in the MST). In the present work, a Gaussian was used as the kernel function.

We note that the results of the SOM do not generally depend on the initialization of the units, unlike feedforward MLP networks where the resulting model depends on weight initialization [5]. This implies that the skeleton of a shape is determined by the data itself and not by the initial coordinates of the SOM.

IV. Issues in the Evolution of the Skeleton

The approximation of the skeletal representation for a given shape distribution depends not only upon the map topology, but on the number of map units as well. 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 [6] and add a new unit in the middle between two existing units $\mathbf{u}_i$ and $\mathbf{u}_j$ if $||\mathbf{W}_i - \mathbf{W}_j|| > \delta_{\max}$. Similarly, two existing units $\mathbf{u}_{ik}$ and $\mathbf{u}_j$ are merged into a single one if the distance between their weight vectors $||\mathbf{W}_i - \mathbf{W}_j|| < \delta_{\min}$. For a given shape distribution skeletal descriptions of varying details can be obtained by suitably tuning the parameters $\delta_{\min}$ and $\delta_{\max}$.

While the spanning tree topology provides an intrinsic method to span complex shapes, it is unable to represent certain topological properties like the closure of circular regions. Circular regions can be represented by knowing the adjacency relations between the Voronoi regions of the shape as represented by their Delaunay triangulation. We define the *neighborhood* relationship between two Voronoi regions $\mathbf{V}_{\mathcal{P}}$ and $\mathbf{V}_{\mathcal{Q}}$ with centroids $\mathbf{u}_{\mathcal{P}}$ and $\mathbf{u}_{\mathcal{Q}}$ based on their second-order Voronoi polyhedra $\mathbf{V}_{\mathcal{PQ}}$, where $\mathbf{V}_{\mathcal{PQ}}$ is defined as [14]

$$\begin{aligned}\mathbf{V}_{\mathcal{PQ}} = \{\mathbf{x}_i \in \Re^d \colon & ||\mathbf{x}_i - \mathbf{W}_{\mathcal{P}}|| \le ||\mathbf{x}_i - \mathbf{W}_{\mathcal{Q}}|| \\ & \le ||\mathbf{x}_i - \mathbf{W}_{\mathcal{K}}||, \forall \mathcal{K} \ne \mathcal{P}, \mathcal{Q}\}. \end{aligned} \tag{2}$$

The second-order Voronoi polygon, thus consists of points that have $\mathbf{u}_{\mathcal{P}}$ as their closest centroid and $\mathbf{u}_{\mathcal{Q}}$ as their second closest centroid. It may be noted (see [14] for a proof), that

$$\mathbf{V}_{\mathcal{PQ}} \ne \phi \Leftrightarrow \mathbf{V}_{\mathcal{P}} \cap \mathbf{V}_{\mathcal{Q}} \ne \phi. \tag{3}$$

The computation of the neighborhood relationships can then be done by determining for each $x_i \in \mathcal{X}$, the two closest centroids $\mathbf{W}_k$ and $\mathbf{W}_l$

$$||\mathbf{x}_i - \mathbf{W}_k|| \le ||\mathbf{x}_i - \mathbf{W}_j||, \forall j \tag{4}$$

and

$$||\mathbf{x}_i - \mathbf{W}_l|| \le ||\mathbf{x}_i - \mathbf{W}_j||, \forall j, j \ne k. \tag{5}$$

The relations in (4) and (5) can be implemented by joining, for each datum, the two closest centroids during the Voronoi tessellation in the batch SOM algorithm. Since the shape distributions we are dealing with form sparse manifolds, the above relations lead to an induced Delaunay triangulation representing the neighborhood property expressed in (2). Skeletal representation of circular regions are obtained by joining two disjoint units in the SOM if there is an edge between them in the induced Delaunay triangulation. Thus, the shape skeleton is a proper subset of the Delaunay triangulation of the shape. 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.

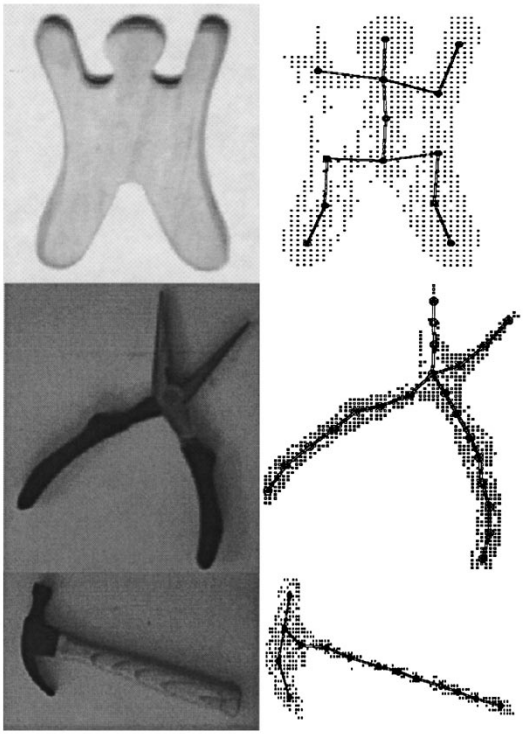

Fig. 4. Results for some objects in our database.

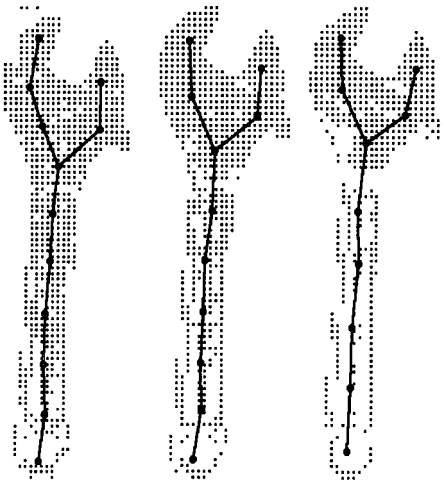

Fig. 5. Skeletons for the object from the object in Fig. 2.

V. Experimental Results

The performance of the proposed algorithm was tested on three sets of sparse images. In the first experiment, the algorithm was applied on a collection of objects. The images of these objects were captured and bilevel thresholding was performed for each object using both moment preserving thresholding and

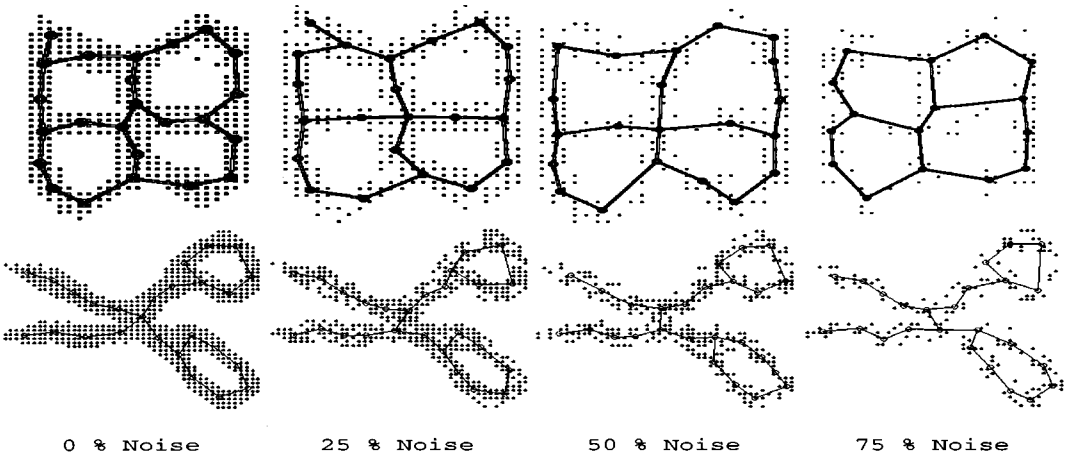

Fig. 6. Skeletal shape extraction on a pair of objects under decreasingSNR.

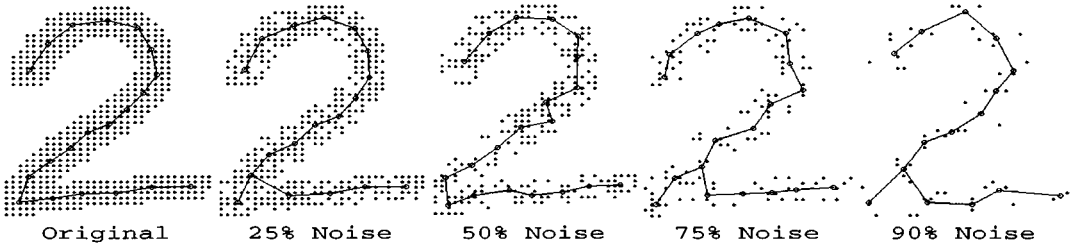

Fig. 7. Skeletonization performance under decreasing SNR. The original image is a dilated handwritten numeral. Uniformly distributed random noise is added to the shape.

the Otsu algorithm (minimizing within group variance) [8]. The sparseness in the object regions, was primarily due to poor control over the imaging conditions and/or incorrect thresholding. The skeletonization results for three objects from this experiment are presented in Fig. 4. In Fig. 5, the shape skeletons of the object shown under varying imaging conditions in Fig. 2 are shown.

In the second experiment, the 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. This procedure was similar to that used in [4] to generate the test data. We note that shapes obtained by dilating thin patterns tend to be much more uniform in thickness than the shapes considered by us in the other experiments. A predetermined amount of random noise, uniformly distributed in the bounding box containing the image, was then added to the image regions. The amount of the noise added was computed by counting the number of black pixels which were changed. Skeletons for a pair of objects at different noise levels are shown in Fig. 6. The noise is varied between 0 to 75% (1.25 dB) of the original number of pixels in the shape. The results on the dilated shape of a handwritten numeral are shown in Fig. 7. Typically, for shapes from this class, our method produced skeletal descriptions consistent with the visual form of the data beyond SNR of 1.25 dB (noise levels of 75%). In Fig. 8, a skeleton obtained with the proposed method is compared with those obtained by thinning a sparse handwritten numeral after dilation-erosion operations. The degenerate skeleton in part Fig. 8(c) occurred due to sparseness that remained after five dilation-erosion operations. The sparseness was removed after ten dilation-erosion operations [Fig. 8(d)]. However, as we have stated in Section II, there is no way to estimate the number of dilation-erosion operations and their overuse may cause changes in the shape topology. In Fig. 9, the results obtained by applying the proposed approach to three different shapes are presented. The examples were specifically chosen to highlight the performance of the algorithm for shapes with region crossovers. The images in the bottom row were obtained by adding 50% (SNR = 3.01 dB) uniformly distributed random noise to the corresponding nonsparse shapes from the top row.

In the third experiment, the algorithm was applied for skeletonization of letters in a faded text. The pertinent document pages were scanned and bilevel thresholding was performed using moment preserving thresholding. Skew estimation and correction was performed using the algorithm in [1]. After skew correction, horizontal projection profiles were used to separate lines in the text. The separation of words and individual letters was done using a method based on vertical projection profiles. For a detailed description of the method as applied to document image analysis, we refer the reader to [23]. A section from a

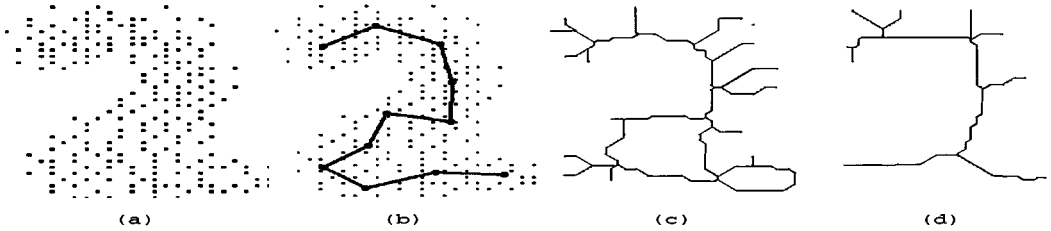

Fig. 8. Comparison of skeletons obtained with the proposed method and those after dilation-erosion followed by thinning: (a) sparse handwritten numeral, (b) skeleton using the proposed method, (c) skeleton by thinning after five dilation-erosion operations, and (d) skeleton by thinning after ten dilation-erosion operations.

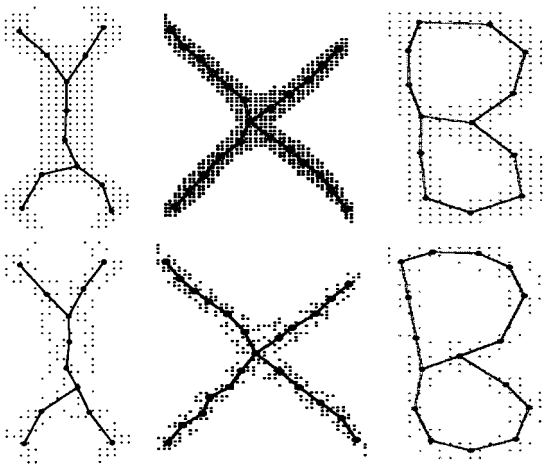

Fig. 9. Skeletons of shapes with region crossovers.

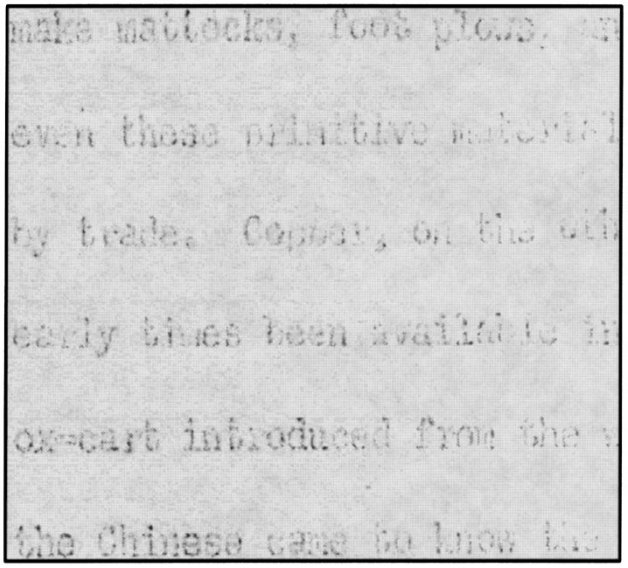

Fig. 10. Section from a text faded due to aging.

scanned page of a faded text is presented in Fig. 10. The skeletal representation of words from the third and fourth lines of the text shown in Fig. 10 are shown in Fig. 11. The reader may note the breaks in the skeletons caused due to excessively large gaps within many individual letters. In all the above experiments, the convergence of the algorithm occurred within 50 to 70 iterations of the data set. The convergence rate (in terms of the number of iterations) is much faster than that reported in [6], where a flow-through SOM was used.

A. Analysis of the Experimental Results

An objective criteria to quantify the performance of the skeletonization algorithm is difficult to formulate. This is primarily due to the lack of a formal definition for the skeleton of a sparse shape. Since increasing sparseness leads to deformations in the shape skeleton, a numerical characterization of the algorithm's performance may be obtained by quantifying the deformations that relate the skeleton of a sparse shape to the skeleton of the corresponding nonsparse shape (whenever available). For this, we use the approach suggested in [4], and compute the deformation of the skeleton as the normalized Euclidean distortion defined below.

Let $\mathcal{I}_1$ and $\mathcal{I}_2$, be two binary images. Let these two images be indexed (in terms of their rows and columns) by the variable pairs (i, j) and (k, l), respectively. Let also $||(i, j), (k, l)||$ denote the Euclidean distance between the pixels with coordinates (i, j) and (k, l). The normalized Euclidean distortion between the images $\mathcal{I}_1$ and $\mathcal{I}_2$ is then defined as

$$d_{1,2} = \frac{1}{2}\left[\frac{\sum_{\forall \mathcal{I}_1(i,j)\neq 0} \mu_{\forall \mathcal{I}_2(k,l)\neq 0}||(i, j), (k, l)||}{|\mathcal{I}_1|} + \frac{\sum_{\forall \mathcal{I}_2(k,l)\neq 0} \mu_{\forall \mathcal{I}_1(i,j)\neq 0}||(k, l), (i, j)||}{|\mathcal{I}_2|}\right]. \quad (6)$$

In the above,

μ minimization operator;

$|\mathcal{I}_n|$ number of nonzero pixels in the binary image $\mathcal{I}_n$.

In Fig. 12(a), we plot the distortion of the skeletons of the handwritten numeral shown in Fig. 7. The plots for the skeletons of the two objects shown in Fig. 6 are presented in Fig. 12(b). A rapid increase in the distortion can be observed from 90% sparseness (SNR 0.22 dB) onwards for the handwritten numeral and from 75% sparseness (SNR 1.25 dB) onwards for the two objects. In each of these cases the increase corresponded to a rapid deterioration in the quality of the skeletons.

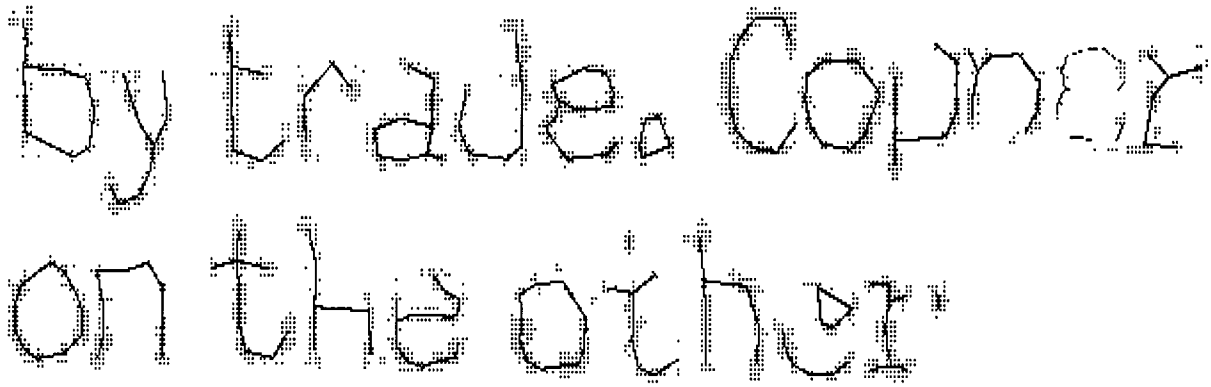

Fig. 11. Skeletonization results for a line from the faded text shown in Fig. 10.

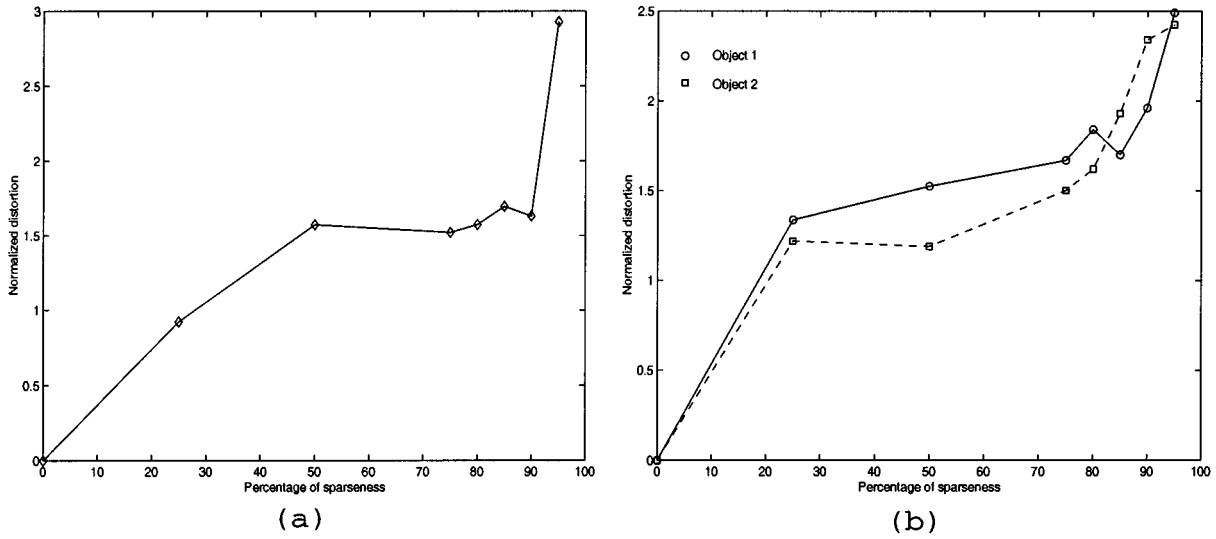

Fig. 12. Normalized Euclidean distortion of the skeletons from Figs. 7 and 6. The distortion is computed with respect to the skeleton of the corresponding nonsparse image.

VI. Conclusion

This paper presents a method for obtaining the skeletal description of sparse shapes. The proposed approach is based on obtaining the shape skeleton by using an MST-based batch mode SOM. The proposed algorithm is interesting in that it neither involves assumptions about demarcation of object regions from the background, nor does it require pixel-level connectivity inside regions. The skeletonization is invariant to Euclidean transformations and is adaptive in terms of the topology of the underlying shape distribution as well as in the number of map units. Additionally, the batch formulation leads to rapid convergence of the skeleton. Due to lack of contiguity in the image region of sparse shapes, conventional skeletonization techniques are either inapplicable or perform poorly in such problems. Since sparseness in images can occur in many real-world problems where imaging conditions can not be completely controlled and because the results of this technique on nonsparse images are similar to those obtained with standard techniques, the proposed algorithm can be used in the development of robust vision systems. Additionally, it can be applied to novel applications like direct retrieval of character shapes from faded documents. An important direction of future research would be to develop techniques for a data-driven estimate of the parameter values used in the algorithm.

Acknowledgment

The authors would like to thank M. C. Wade for his help in applying the method for description of letters in faded text. They would also like to thank the James J. Hill Reference Library, St. Paul, MN for the faded text used as an example in this paper.

References

[1] H. S. Baird, "The skew angle of printed documents," in *Document Image Analysis*, L. O'Gorman and R. Kasturi, Eds: IEEE Comput. Soc. Press, 1995, pp. 204–207.

[2] H. Blum, "A transformation for extracting new descriptors of shape," in *Models for the Perception of Speech and Visual Form*, W. Wathen-Dunn, Ed. Cambridge, MA: MIT Press, 1967, pp. 362–380.

[3] M. Bokser, "Omnidocument technologies," *Proc. IEEE*, vol. 80, no. 7, pp. 1066–1078, 1992.

[4] Y. S. Chen and Y. T. Yu, "Thinning approaches for noisy digital patterns," *Pattern Recognit.*, vol. 29, no. 11, pp. 1847–1862, 1996.

[5] V. Cherkassky and F. Mulier, *Learning From Data*, 1998.
[6] A. Datta, S. K. Parui, and B. B. Chaudhuri, "Skeletal shape extraction from dot patterns by self-organization," in *Proc. 13th Int. Conf. Pattern Recognition*, vol. 4, 1996, pp. 80–84.
[7] A. R. Dill, M. D. Levine, and P. B. Noble, "Multiple resolution skeletons," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, no. 4, pp. 495–504, 1987.
[8] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*. Reading, MA: Addison-Wesley, 1992.
[9] T. Hastie and W. Stuetzle, "Principal curves," *J. Amer. Statist. Assoc.*, vol. 84, no. 406, pp. 502–516, 1989.
[10] J. D. Hobby and T. K. Ho, "Enhancing degraded document images via bitmap clustering and averaging," in *Proc. 4th Int. Conf. Document Anal. Recognition*, vol. 1, 1997, pp. 394–400.
[11] J. W. Jaromczyk and G. T. Toussaint, "Relative neighborhood graphs and their relatives," in *Proc. IEEE*, vol. 80, 1992, pp. 1502–1517.
[12] J. A. Kangas, T. K. Kohonen, and J. T. Laaksonen, "Variants of self-organizing maps," *IEEE Trans. Neural Networks*, vol. 1, pp. 93–99, 1990.
[13] S. P. Luttrell, "Derivation of a class of training algorithms," *IEEE Trans. Neural Networks*, vol. 1, no. 2, pp. 229–232, 1990.
[14] T. Martinetz and K. Schulten, "Topology representing networks," *Neural Networks*, vol. 7, no. 3, pp. 507–522, 1994.
[15] F. Mulier and V. Cherkassky, "Self-organization as an iterative kernel smoothing process," *Neural Comput.*, vol. 7, no. 6, pp. 1165–1177, 1995.
[16] F. Mulier and V. Cherkassky, "Statistical analysis of self-organization," *Neural Networks*, vol. 8, pp. 717–727, 1995.
[17] J. L. Mundy and R. E. Joynson, "Automatic visual inspection using syntactic analysis," in *Proc. Int. Conf. Pattern Recognition Image Processing*, 1977, pp. 144–147.
[18] R. L. Ogniewicz and O. Kubler, "Hierarchic voronoi skeletons," *Pattern Recognit.*, vol. 28, no. 3, pp. 343–359, 1995.
[19] M. Oguro, T. Akiyama, and K. Ogura, "Faxed document image restoration using gray level representation," in *Proc. 4th Int. Conf. Document Anal. Recognition*, vol. 2, 1997, pp. 679–683.
[20] J. D. Radke, "On the shape of a set of points," *Computational Morphology*, pp. 105–137, 1988.
[21] H. Ritter, T. Martinetz, and K. Schulten, *Neural Computation and Self-Organizing Maps: An Introduction*. Reading, MA: Addison-Wesley, 1992.
[22] F. Roll, J. M. Chassery, and A. Montanvert *et al.*, "3-D medial surfaces and 3-D skeletons," in *Visual Form*, C. Arcelli *et al.*, Eds. New York: Plenum, 1992, pp. 443–450.
[23] R. Singh, M. C. Wade, and N. P. Papanikolopoulos, "Letter-level shape description by skeletonization in faded documents," in *Proc. 4th IEEE Workshop Applicat. Comput. Vision*, 1998, pp. 531–536.
[24] Q. Z. Ye and P. E. Danielsson, "Inspection of printed circuit boards by connectivity preserving shrinking," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, no. 5, pp. 737–742, 1988.
[25] S. C. Zhu and A. L. Yuille, "FORMS: A flexible object recognition and modeling system," *Int. J. Comput. Vision*, vol. 20, no. 3, pp. 187–212, 1996.