SEGMENTATION OF SHAPES*

I. Introduction

What is meant by shape segmentation in this paper is decomposition of a 2D shape into ribbons.The basic idea is to delineate the main body of the shape by segmenting out protrusions. The inclusion relations among protrusions induce the structure of a directed graph on the segmentation, analogous to the graph structure associated with shape skeletons. A finer segmentation may be obtained by segmenting the shape across narrow necks.

The approach is based on the concept of local symmetry axes which was developed by Tari, Shah and Pien in [3] and further developed in [4]. Axes of local symmetry are analogous to the more commonly used medial axes. If a 2D shape is viewed as a collection of ribbons glued together, then the medial axis of each ribbon may be thought of as a local symmetry axis. In contrast to the usual use of medial axes to obtain shape skeletons, here they are used to segment the shape. The set of local symmetry axes is found by analyzing the level curves of a function, v, which is the solution of an elliptic PDE. The function v smooths the characteristic function of the shape boundary. A point on a level curve of v is a point of local symmetry if the level curve is locally symmetric about the gradient vector of v at that point upto second order. The local symmetry axes may also be described as the ridges and the valleys of the graph of v. The rationale underlying this approach is that if a shape has certain symmetries, then the solution of the PDE ought to reflect these symmetries.

One of the motivations behind this approach was to carry out noise suppression and extraction of shape properties simultaneously. In this spirit, the level curves of v may be thought of as successive smoothings of the shape boundary and thus the approach described above has close similarities with that based on curve evolution. However, the advantage here is that the necessary properties of the level curves are calculated from the differential properties of the function v itself, without having to locate the level curves of v. This is not true in the case of curve evolution because the time of arrival of the evolving curve at various points in the domain almost never defines a function over the domain. (The evolving curve may cross a given point several times.) Moreover, differential derivatives needed in the case of curve evolution are one order higher than in the approach proposed here, making the numerical calculations more sensitive to noise. Like curve evolution, there is a smoothing parameter in the elliptic PDE and as it tends to zero, function v tends to the (rescaled) distance function. Of course, shape analysis based on distance function has a long history. (For some recent developments, see the [2].) Local symmetry axes may be used to extract shape skeletons, but as in the case of curve evolution, smoothing built into the process disconnects the skeleton. It is more useful to employ local symmetry axes to segment the shape instead.

The work described here is also related to that of Zhu [5] who has formulated a segmentation functional to draw optimal chords. The optimal chords determine the basic ribbon structure of the shape. The medial axes of these ribbons are joined together in an optimal way to determine the shape skeleton. The advantage of this approach is that the segmentation functional provides a goodness criterion for evaluating optimality of shape skeletons and also permits a statisti-

*This work was supported by NIH Grant I-R01-NS34189-04.

cal framework. However, the technique requires calculation of the shape normals and thus shape has to be presmoothed.

The version of the algorithm described here segments shapes at a very basic level, namely, it extracts the longest possible ribbon from the shape by segmenting out protrusions. The shapes need not be presmoothed and the algorithm may be applied to a complex scene consisting of many objects. Continuity properties of the solution of elliptic PDEs with respect to deformation of the domain boundary suggest a possible approach for establishing similar properties for shape features.There are several obvious refinements that can be made. It might be necessary to further segment the ribbon segments found by the algorithm. The algorithm does provide the option of further segmenting each of the ribbon segments by creating cuts originating at the saddlepoints of v. However, this may or may not be desirable depending on the application. For example, in shapes with a long neck, a saddlepoint will be formed at its narrowest point and the neck will be segmented there. However, it might be preferrable to isolate the whole neck as one object. Another example is that of a bottle-shape with a narrow uniformly tapering neck. If the proportions are properly arranged, the algorithm treats this as one long ribbon instead of isolating the neck. The algorithm as formulated deals with generic shapes in the sense that it is not sensitive to special symmetries such as those exhibited by a square or a rectangle. It has to be modified if these special coincidences have to be taken into account. For instance, depending on the numerical choices made, it will segment out two of the corners of a rectangle as protrusions, indentifying the rest as the longest ribbon that can be extracted. Such a segmentation would appear reasonable in the case of a parallelogram in which case the two obtuse angles will be segmented out as protrusions and a long ribbon around the long diagonal will be extracted. Since the algorithm does not recognize special cases, it treats the rectangle as a generic parallelogram.

II. Smoothing of the Shape Boundary

A shape is described by specifying its boundary in the form of a collection of curves, Γ, inside a bounded domain D of the plane. All that is necessary is that Γ be sufficiently regular for the solution of the differential equation given below to exist. We consider the usual L^2 functional for smoothing the characteristic function of Γ :

$$E_\rho(v) = \int_D \left[\rho \left\| \nabla v \right\|^2 + \frac{(v - \chi_\Gamma)^2}{\rho} \right] \quad (1)$$

$$= \int_D \left[\rho \left\| \nabla v \right\|^2 + \frac{v^2}{\rho} \right] \quad (2)$$

with the boundary condition $v = 1$ along Γ where

$$\text{the characteristic function } \chi_\Gamma = \begin{cases} 1 \text{ along } \Gamma \\ 0 \text{ elsewhere} \end{cases} \quad (3)$$

Alternative smoothing strategies are possible. The advantage of this particular functional is that it behaves correctly in the limit as $\rho \to 0$:

$$\lim_{\rho \to 0} \inf E_\rho(v) = length(\Gamma) \quad (4)$$

The minimizer of E_ρ satisfies the elliptic differential equation

$$\nabla^2 v = \frac{v}{\rho^2} \quad (5)$$

with boundary conditions $v = 1$ along Γ and $\frac{\partial v}{\partial n} = 0$ along the boundary of D. The parameter ρ plays the role of the smoothing radius.

Although what is relevent here is the global behavior of v which determines the axes of local symmetries, it is interesting to note that when ρ is small compared to the local width of the shape and the local radius of curvature of Γ, the level curves of v locally capture the smoothing of Γ by curve evolution. As shown in Appendix (3) of [1], when ρ is small,

$$v(x, y) = -\rho \left(1 + \frac{\rho \kappa(x, y)}{2} \right) \frac{\partial v}{\partial n}(x, y) + O(\rho^3) \quad (6)$$

where $\kappa(x, y)$ is the curvature of the level curve passing through the point (x, y) and n is the direction of the gradient. If we imagine moving

from a level curve to a level curve along the normals, then a small change of δv in the level requires movement

$$\delta r \approx -\frac{\rho}{v}(1+\frac{\rho\kappa}{2})\delta v \tag{7}$$

where r denotes the arc length along the gradient lines of v. Define time t such that $\frac{dt}{\rho^2} = \frac{dv}{2v}$. Then

$$\frac{dr}{dt} \approx \frac{2}{\rho} + \kappa \tag{8}$$

As pointed out above, the global behavior of the level curves is radically different from that of curve evolution since the value of v at a point cannot be determined by its values in a local neighborhood. For example, consider a closed level curve with large width and small curvature everywhere so that its evolution mimics the curve evolution. As the level curve shrinks and narrows, interaction between its opposite sides becomes significant and the gradient of v will be less than what it would be without the interaction. For instance, the level curve speeds up as it nears a saddlepoint.

III. Local Symmetries, Medial Axes and Skeletons

Loci of local symmetries are now defined by analyzing the local symmetries of the level curves of v. These loci consist of one-dimensional branches and their terminal points. Figure 1 shows the level curves of v inside a starlike shape. Notice that along the apparent medial axes of protrusions, the level curves are further separated than they are in the neighborhood of indentations. The tips of protrusions are in some sense furthest away from the apparent center of the shape. Now the distance between two adjacent level curves is given by $\frac{dv}{\|\nabla v\|}$. If we define the semimetric

$$\frac{dv/dl}{\|\nabla v\|}dl \tag{9}$$

where dl is the infinitesimal Euclidean distance, then the the geodesics satisfy the equation

$$\frac{d\,\|\nabla v\|}{ds} = 0 \tag{10}$$

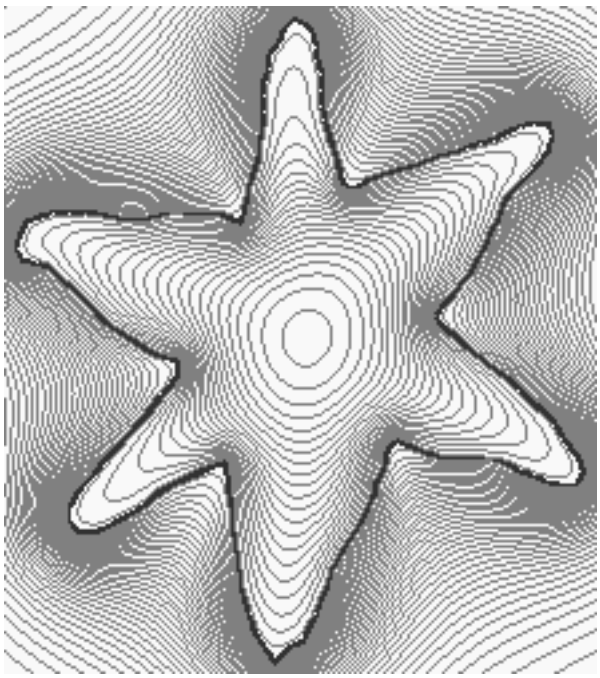

Figure 1 – Level curves of function v

where s is the arc-length along the level curves of v. The symmetry of the level curve at a point P where $\frac{d\|\nabla v\|}{ds} = 0$ is revealed by the missing $\eta\xi$-term in the Taylor expansion of v in terms of the local coordinates η and ξ where η is in the direction of ∇v and ξ is tangent to the level curve:

$$v = a_{00} + a_{10}\eta + a_{01}\xi + a_{20}\eta^2 + a_{02}\xi^2 + \cdots \tag{11}$$

Thus locally at P, the level curve $v = a_{00}$ is approximately a conic section whose one of the principal axes coincides with the gradient vector. An equivalent description of the symmetry at P is that the Hessian of v at P is diagonalized when expressed in terms of the local coordinates η and ξ. This means that the gradient vector ∇v is an eigenvector of the Hessian at P. The last description may be generalized to define partial symmetries of shapes in dimensions > 2 [4].

As explained above, along the middle of protrusions, the distance between adjacent level curves is the greatest, that is, $\|\nabla v\|$ is minimum along the level curve. So let S_1^+ denote the closure of the set of zero-crossings of $\frac{d\|\nabla v\|}{ds}$

where $\frac{d^2\|\nabla v\|}{ds^2}$ is positive and let S_1^- denote the closure of the set of zero-crossings of $\frac{d\|\nabla v\|}{ds}$ where $\frac{d^2\|\nabla v\|}{ds^2}$ is negative. The connected components of $S_1 \backslash (S_1^+ \cap S_1^-)$ are called the branches of S_1. The direction of each branch is in the direction of increasing v. Figure 2 illustrates the loci S_1^+ and S_1^- in the case of the star-like figure.

The set $S_1^+ \cap S_1^-$ consists of the terminal points of the branches of S_1 and it is the union of two sets, S_0 and J. The set S_0 is defined by the equation $\|\nabla v\| = 0$ and the set J is defined by the equations $\frac{d\|\nabla v\|}{ds} = \frac{d^2\|\nabla v\|}{ds^2} = 0$. The set S_0 may be further subdivided into the set S_0^+ of elliptic points where the determinant of the Hessian of v is positive, the set S_0^- of hyperbolic points where it is negative and the set S_0^0 of parabolic points where it is zero. At an elliptic point, v has a local minimum and has the Taylor expansion of the form $a_{00} + a_{20}x^2 + a_{02}y^2 +$ *higher order terms*. By applying the definition of S_1^+ and S_1^- to this local expression, it is easy to see that at an elliptic point, there are two branches of S_1^+ directed away from the point in the direction of the maximum second derivative and two branches of S_1^- directed away in the direction of the minimum second derivative. At a hyperbolic point, v has the Taylor expansion of the form $a_{00} + a_{11}xy +$ *higher order terms* and calculations show that at a hyperbolic point, there are four branches of S_1 all of which belong to S_1^+. Hyperbolic points are of course saddlepoints in that two of these branches are directed away from the saddlepoint and and two are directed towards it. In theory, the set S_0^0 of parabolic points may be one-dimensional, but numerically it is impossible to identify such points without setting some kind of a numerical threshold. What we find is that a parabolic line is seen numerically as a series of elliptic and hyperbolic points, making analysis of parabolic points difficult. However, from the point of view of segmentation, all that is needed is determination of the kind of branches of S_1 that are present in a tubular neighborhood.

Generically, a point in J is a junction of a branch from S_1^+ and a branch from S_1^-. As in the case of parabolic lines, in the absence of a numerical threshold, all points in J are numerically regarded as belonging to this category. A point in J belongs to the subset J^+ if the two branches of S_1 are directed away from it and it belongs to the subset J^- if they are directed towards it. At points of J, v has a local maximum or a local minimum when restricted to S_1. It is minimum if the point belongs to J^+ and maximum if it belongs to J^-. Junctions of type J^- arise when a parabolic line breaks up into a series of elliptic and hyperbolic points or when there is protrusion present near a neck. The latter case can be seen in Figure 2. Along the boundary of the rectangular domain, four saddlepoints are created by the artificially created necks outside the shape. The S_1^+-branch from such a saddlepoint links up with the S_1^+-branch from the nearby protrusion with an S_1^--branch interposing between the two. (Note that an indentation in the shape behaves like a protrusion when seen from outside the shape.) In principle, there should be exactly one point of J^+ and one point of J^- in this linkage. However, there is numerical degeneracy due to the fact that the level curves $\frac{d\|\nabla v\|}{ds} = 0$ and $\frac{d^2\|\nabla v\|}{ds^2} = 0$ are nearly coincident over some distance creating a series of points numerically identified as points of J, (see the saddlepoint on the right).

Since there are only four branches at an elliptic point, most branches of S_1 end up at points in J. The smaller the protrusion, the shorter the branch. Notice the extremely short branches near the shape boundary, created by the noise in the boundary. The two branches of S_1^+ meeting at the elliptic point inside the star define the medial axis of the longest ribbon that can be extracted from the shape.

The construction described above depends on the choice of the smoothing parameter ρ. Axes shown in Figure 2 were obtained with $\rho = 32$ pixels. (The size of the frame D is 400×400 pixels.) Figures 3 and 4 depict the axes determined using $\rho = 8$ and $\rho = 128$. The features that are sensitive to the choice of ρ are the points in J and the number of saddlepoints. The larger the value of ρ, the shorter the protrusion axes. Since the

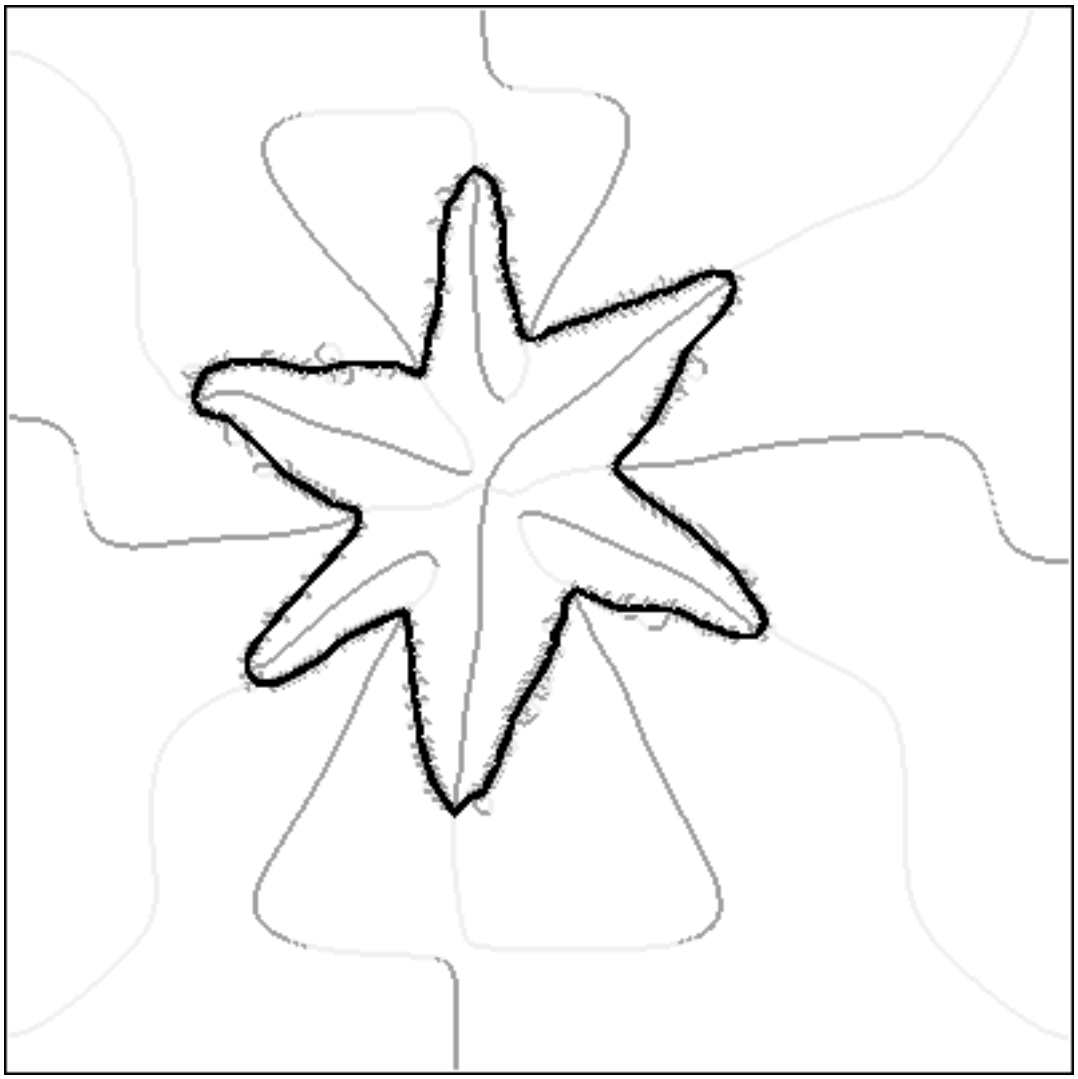

Figure 2 – Local symmetry axes. S_1^+ dark grey, S_1^- light grey

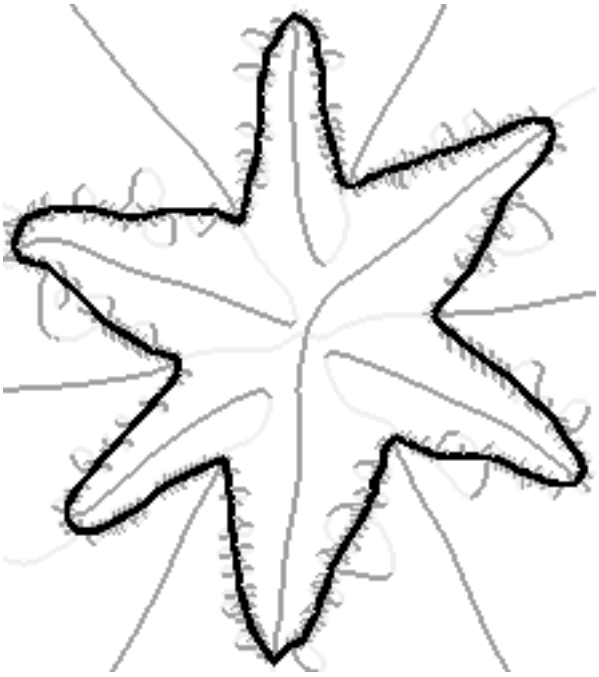

Figure 3 – Local symmetry axes, $\rho = 8$ pixels.

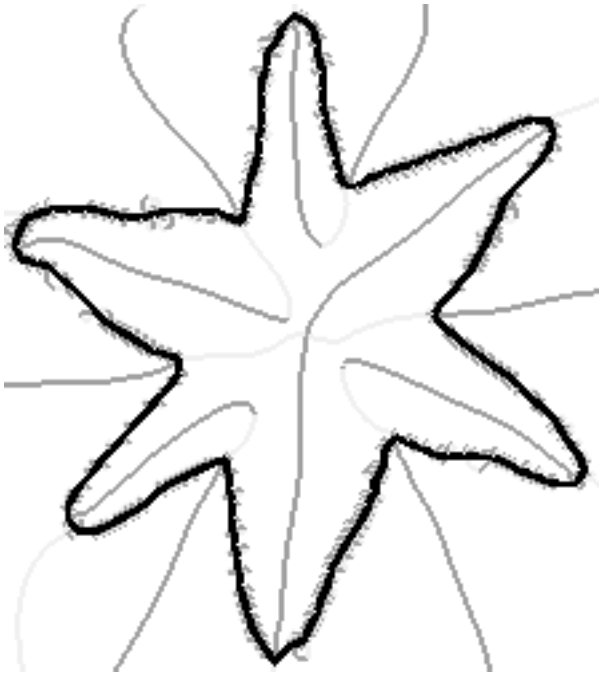

Figure 4 – Local symmetry axes, $\rho = 128$ pixels.

function v emulates the distance function more and more closely as ρ tends to zero, it begins to detect even very wide necks by creating more saddlepoints. Main segments of the medial axes remains fairly stable since they approximate the center line of the protrusions.

We conclude this section with the definition of the shape skeleton:

Definitions: A *medial axis* is a branch of S_1^+ which starts at an elliptic point or at a point in J^+ and ends either at the shape boundary or at a saddlepoint. A medial axis starting at an elliptic point will be called a *main axis* while a medial axis starting at a point in J^+ will be called a *protrusion axis*. The *skeleton* of the shape is the union of its medial axes.

As noted before, the shape skeleton is not connected.

IV. Segmentation

The medial axes detect the "corners" of the shape. The saliency or the extent of each corner may be guaged by the length of the associated medial axis. However, the main objective of this paper is to segment protrusions and indentations by means of their medial axes. The basic idea is to find the two nearest points on the shape boundary from the terminal point of the protrusion axis, one on each side of the axis and connect these two points to segment the protrusion. The important point is to restrict the search to a suitable neighborhood of the protrusion axis. To solve this problem, we use the fact that S_1 segments D and inside each connected component of $D \backslash S_1$, $\frac{d\|\nabla v\|}{ds}$ is either positive or negative. Segmentation of D in the case of the star-figure is shown in Figure 5. Each protrusion axis neighbors exactly two of these components. Therefore the search for the nearest boundary points is restricted to the interior of these two components adjoining the axis. Admittedly, the boundary points found in this way depend on where the terminal point of the protrusion axis is which in turn depends on the choice of ρ. However, if the ends of protrusion are marked by a sharp change in the local width of the shape, the

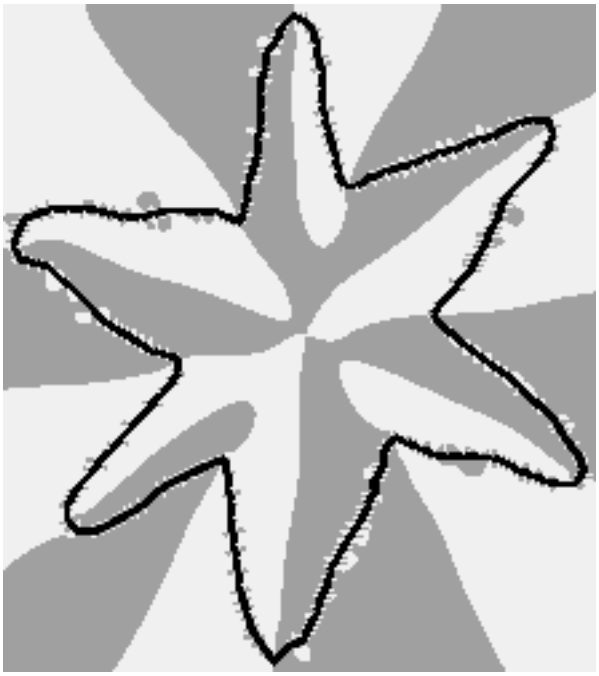

Figure 5 – Dark grey where $\frac{d\|\nabla v\|}{ds}$ is positive, light grey where negative.

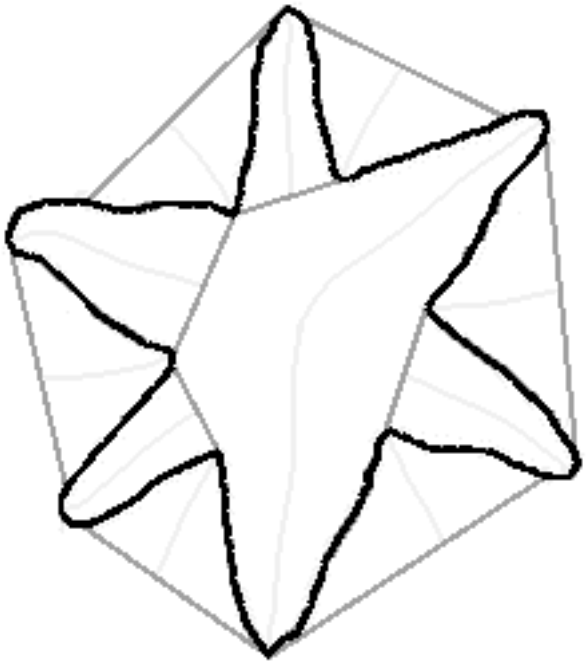

Figure 6 – Segmented shape (from inside and outside).

boundary points are insensitive to ρ. In the special cases when this is not true as in the case of a parallelogram, the larger the value of ρ, the further away the boundary points from the obtuse corners interpreted as protrusions.

It is possible to segment shapes across necks by means of the associated saddlepoints. This is a more delicate construction. Here the problem is to avoid spurious saddlepoints arising from the numerical break-up of parabolic lines. The difficulty is that there may be irrelevant small segments of $D\backslash S_1$ adjoining the saddlepoint. Therefore, a hyperbolic point is called a *true saddlepoint* if it adjoins at least three segments of $D\backslash S_1$ which touch the shape boundary and if the saddlepoint is not a point on the boundary of D. (The last condition avoids saddlepoints artificially introduced by the frame D.) Going through each saddlepoint is a medial axis and the problem is to find the two nearest boundary points, one each side of the medial axis. Restrict the search for the nearest boundary points to only these adjoining segments of $D\backslash S_1$. Once the two boundary points are found, one on each side of the medial axis, connect the saddlepoint to each of them. This construction may still produce double segmentation lines. This happens if there are two branches of S_1 leaving the theoretical parabolic line from two different points to meet the shape boundary or another true saddlepoint. The solution in this case is to search in an appropriately chosen tubular neighhood of the medial axis through the saddlepoint and treat the two terminal points on the parabolic line as a single unit.

Figure 5 shows the segmentation of the star figure. The shape is segmented from inside as well as from outside. As noted before, in its attempt to extract the longest possible ribbon, the algorithm disregards the approximate symmetry of the star and includes in the main ribbon two of what would normally be perceived as protrusions.

The segmentation has the structure of a directed graph. Its set of vertices consists of the true saddlepoints and one vertex for each of the

shape segments. If a segment X is a protrusion with a medial axis originating at a point in J^+ which is contained in another segment Y, then YX is an edge in the graph. (The direction of an edge is always in the direction of increasing v.) Each segmentation line through a saddlepoint is the common boundary between two shape segments. The vertices correponding to these two segments are connected to the vertex corresponding to the saddlepoint by edges directed towards the saddlepoint.

Additional examples of shape segmentation are shown in Figures 7 and 8. (The surrounding frame D has been cropped.) Each shape was scaled to make sure that it was nowhere less than 4 pixels wide. The approximate sizes of the shapes are: the star 208×240 pixels, the cat 208×300 pixels, the hand 280×290 pixels and the brain image 500×500 pixels. All the quantities needed in the algorithm are computed using 3×3 neighborhoods except the sign of $d^2 \left\| \nabla v \right\| / ds^2$ which required 4×3 neighborhooods. In Figure 7, the top row shows the local symmetry axes while the bottom row shows the segmented shape. Figure 8 illustrates the case of a complex of shapes involving non-simply connected shapes and triple junctions.

References

[1] D. Mumford and J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, *Comm. Pure Appl. Math,* **XLII**, 1989, 577-684.

[2] K. Siddiqi, A. Shokoufandeh, S. Dickinson and S. Zucker, Shock graphs and shape matching, *Intl. J. Computer Vision,* **35 (1)**, 1999, 13-32.

[3] S. Tari, J. Shah and H. Pien, Extraction of shape skeletons from grayscale images, *Computer Vision and Image Understanding,* **66**, 1997, 133-146.

[4] S. Tari and J. Shah, Local symmetries of shapes in arbitrary dimension, *Sixth Intl. Conf. Computer Vision,* 1998.

[5] S.C. Zhu, Stochastic jump-diffusion process for computing medial axes in markov random fields, *IEEE Trans. Pattern Analysis and Machine Intelligence,* **21 (11)**, 1999.

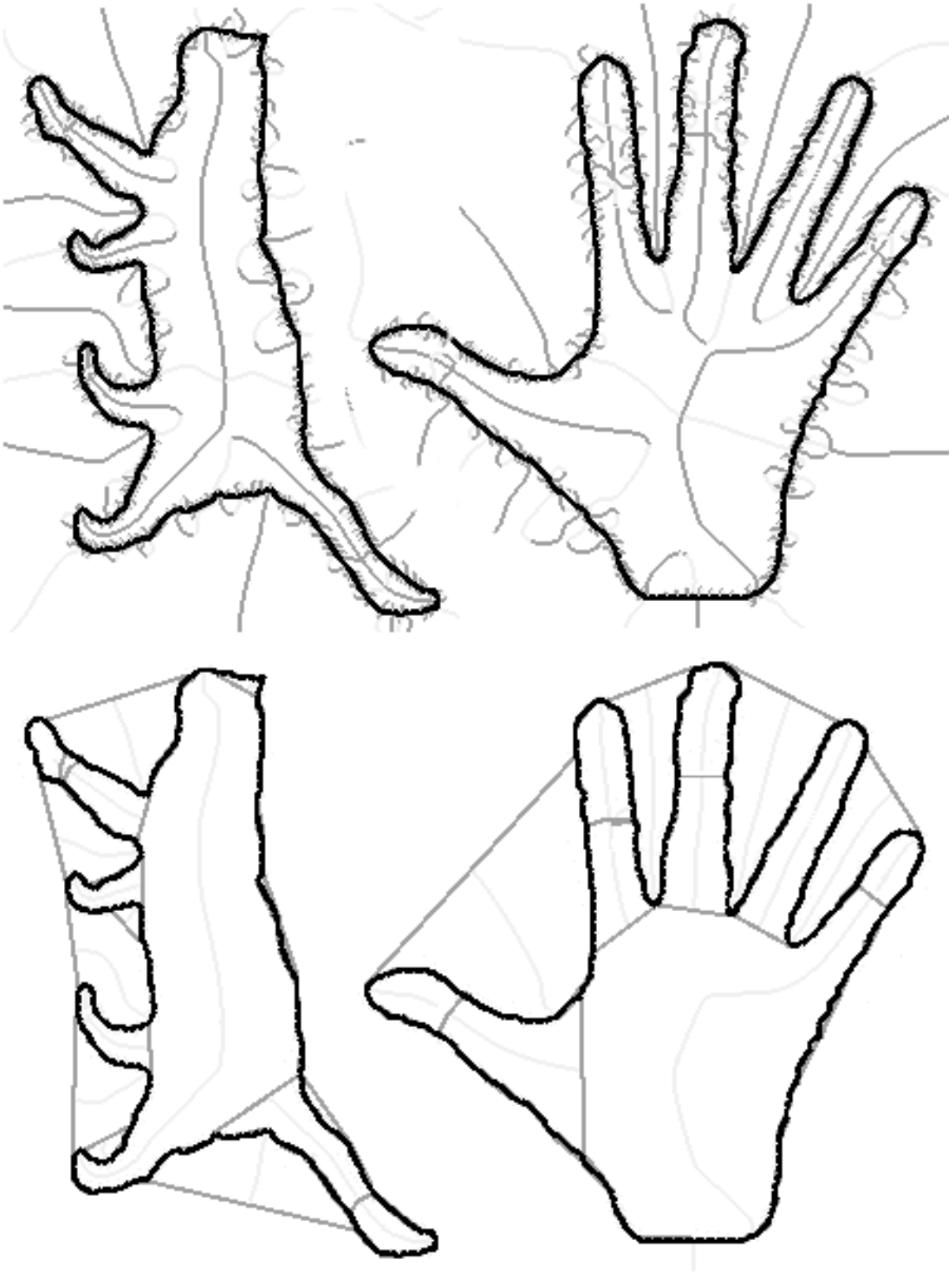

Figure 7 – Top row: local symmetry axes. Bottom row: segmentation

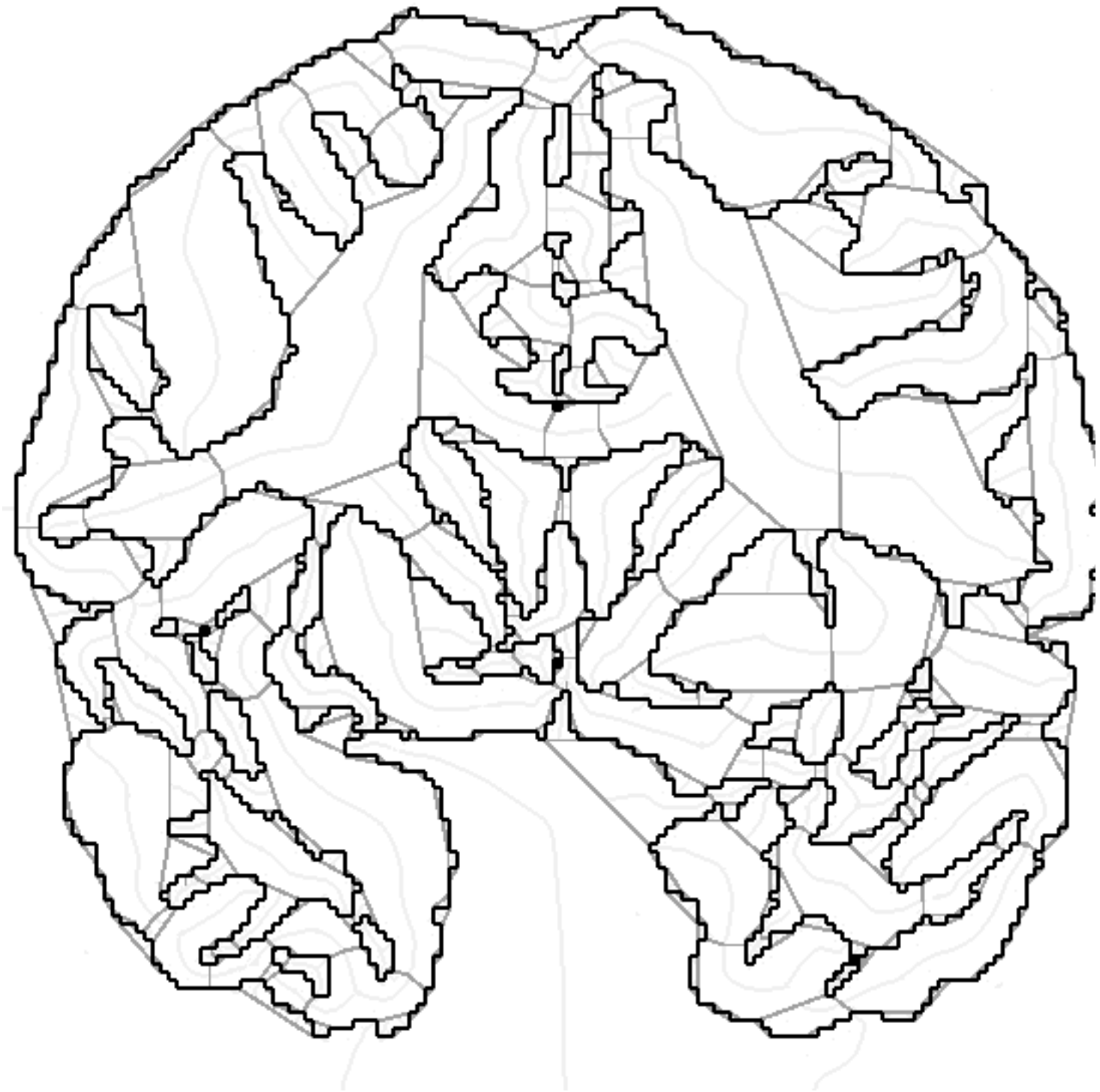

Figure 8 – Segmentation of a complex of shapes