COMPUTER VISION AND IMAGE UNDERSTANDING
Vol. 66, No. 2, May, pp. 133–146, 1997
ARTICLE NO. IV970612

Extraction of Shape Skeletons from Grayscale Images[1]

Z. Sibel Gőktepe Tari

Mechanical and Industrial Engineering Department, Northeastern University, Boston, Massachusetts 02115

Jayant Shah

Mathematics Department, Northeastern University, Boston, Massachusetts 02115

and

Homer Pien

Draper Laboratory, Cambridge, Massachusetts 02139

Received September 1, 1996; accepted January 2, 1997

Shape skeletons have been used in computer vision to represent shapes and discover their salient features. Earlier attempts were based on morphological approach in which a shape is eroded successively and uniformly until it is reduced to its skeleton. The main difficulty with this approach is its sensitivity to noise and several approaches have been proposed for dealing with this problem. In this paper, we propose a new method based on diffusion to smooth out the noise and extract shape skeletons in a robust way. In the process, we also obtain segmentation of the shape into parts. The main tool for shape analysis is a function called the "edge-strength" function. Its level curves are smoothed analogs of the successive shape outlines obtained during the morphological erosion. The new method is closely related to the popular method of curve evolution, but has several advantages over it. Since the governing equation is linear, the implementation is simpler and faster. The same equation applies to problems in higher dimension. Unlike most other methods, the new method is applicable to shapes which may have junctions such as triple points. Another advantage is that the method is robust with respect to gaps in the shape outline. Since it is seldom possible to extract complete shape outlines from a noisy grayscale image, this is obviously a very important feature. The key point is that the edge strength may be calculated from grayscale images without first extracting the shape outline. Thus the method can be directly applied to grayscale images.

1. INTRODUCTION

It has been long recognized that the classical descriptions of curves such as parametrization or Fourier expansion are not adequate for representing shapes [2, 3, 13]. For example, they fail to capture two-dimensional features such as necks and protrusions or blobs versus ribbons. Alternative approaches have been suggested within both computer science and psychology to take into account the two-dimensional nature of the shape. One such approach is the well-known Blum transform [2, 3]. It is commonly visualized by the grassfire analogy in which one imagines the interior of the shape filled with dry grass and fire is started simultaneously at all points on the shape boundary. The advancing front propagates with constant speed and at any time, its points are equidistant from the shape boundary. The shape skeleton is defined as the locus of singularities of the advancing front. By keeping track of the time at which the front arrives at skeleton points, it is possible to recover the original shape. Thus, no information is lost. A consequence of this fact is that if the shape boundary is noisy, the noise is also preserved in the skeleton and thus the transform is not robust with respect to noise. Many strategies have been devised to prune noisy skeletons to arrive at the essential skeleton of the shape [17]. An alternate systematic approach following the philosophy behind segmentation functionals is to introduce smoothing or regularization in the grassfire itself, thus combining skeletonization and regularization in a single formulation. A straightforward approach to regularization is the method of curve evolution in which the front velocity now depends on two components: a constant component as before and a smoothing component proportional to curvature [9, 11, 20]. In this paper, a simpler and faster method of regularization is proposed so that the resulting skeleton is robust with respect to noise.

[1] This work was partially supported under NIH Grant I-R01-NS34189-01 and NSF Grant No. DMS-9531293.

1077-3142/97 $25.00

We first illustrate the concept by means of a simple example. Consider the grassfire analogy for a rectangle. Think of time t as a function over the plane by setting $t(x, y)$ = the time when the advancing front passes through the point (x, y). The surface $t(x, y)$ is then a pyramid ending in a horizontal ridge at the top. The level curves of the surface $t(x, y)$ describe the propagation of the front and the projection of the singular locus of the surface constitutes the skeleton of the rectangle. In place of the surface $t(x, y)$, we propose to use a function v which we call the edge-strength function. Function v is smooth and equals 1 along the shape boundary and decays rapidly away from the boundary. Thus $1 - v$ may be thought of as an approximately monotonic function of distance from the boundary and hence, an approximately monotonic function of t. The level curves of v look like smoothed version of the level curves of t. The surface $1 - v$ looks like a smoothed (monotonically scaled) version of the surface $t(x, y)$. We extract information regarding the shape skeleton and the decomposition of the shape into parts from the geometry of the level curves of v.

The edge-strength function v is computed by means of a linear diffusion equation and thus easy to implement. The diffusion equation may be applied to a collection of curves which may intersect. In particular, the method permits analysis of shapes involving Y-junctions such as the line drawing of a solid cube. It is also applicable to shape outlines which are incomplete due to gaps, that is, to shape outlines consisting of a collection of disconnected pieces.

The most important advantage of the method proposed here is that it can be applied to grayscale images because one may calculate the edge-strength function corresponding to shape boundaries directly from the grayscale image without first segmenting it. This is in contrast to most methods of shape analysis in which the central assumption is that the shape outline has already been extracted in the form of a closed curve from the grayscale image. Indeed, in the approach based on curve evolution, curve evolution is first applied (for example, an intrinsic version of Terzopoulos' snakes) to recover the shape boundary. Curve evolution is then applied again to the recovered shape boundary to extract the shape skeleton. Shape recovery is not an easy task if the image is noisy. A number of recent papers [5, 6, 12, 14, 21, 22, 27] have been devoted to shape recovery from raw images by the method of curve evolution. (See [23] for a discussion and a generalization of these methods.)

The edge-strength function we use is motivated by the Ambrosio–Tortorelli approximation of Mumford–Shah segmentation functional. Indeed, we use Euler–Langrange equations of Ambrosio–Tortorelli approximation to compute the edge-strength function of grayscale images. What we demonstrate is that the Ambrosio–Tortorelli edge-strength function is much more then a technical device for applying gradient descent to the Mumford–Shah functional. It is a much better measure of ''boundaryness'' at every point in the image because the nonlinear smoothing by Ambrosio–Tortorelli approximation reduces or prevents smoothing of the image in the neighborhood of actual object boundaries. We go further and show that it has a close relationship to curve evolution and provides a much easier method for regularizing the ''grassfire'' method of Blum and thus for computing shape skeletons. The crucial point we make is that this edge-strength function is important in its own right and is much better suited for all the current applications (like ''geodesic snakes'') which need a continuous function as a measure of boundaryness across the entire image.

Although our paper is about extraction of shape skeletons from grayscale images, we decided to devote a large part of the paper to binary images for the two reasons: (A) Most of the work by other researchers on shape skeletons is done using binary images. So we think it is important that we give results by our method using the same or similar binary images so that our method can be compared with the others. (B) When our method is used for grayscale images, it integrates within a single framework three steps which are normally carried out as three separate hierachical steps, namely, smoothing of the grayscale image, shape recovery, and determination of shape skeleton. Each step involves different issues such as grayscale noise, noisy boundary, and level of detail in the skeleton. Since our focus is on determining shape skeletons by a novel use of edge-strength function, we think it would be better if the method is explained first in the context of binary images so as to demonstrate clearly the relationship between the edge-strength function and shape skeletons before other issues associated with grayscale images are brought in.

To our knowledge, neither our method nor any other method for determining skeletons can deal with occlusions. A separate ingredient such as 2.1D sketch of Nitzberg and Mumford or some type of multilayered representation is required. Note that there is some confusion in the literature between incomplete figures and occluded figures. If an occluded figure is extracted from its surroundings as an incomplete figure, our method can be applied as shown in Section 4.

This paper is organized as follows. In Section 2, we introduce the edge-strength function and analyze the behavior of its level curves. In Section 3, shape skeleton and shape segmentation are defined and illustrated by several test cases. In Section 4, we deal with incomplete figures, and Section 5, we discuss ways in which one may assign a level of significance to different branches of a shape skeleton. Finally, in Section 6, we extend our construction to grayscale images and apply it to analyze an MRI Image.

A preliminary version of this paper appeared in [26].

2. THE EDGE-STRENGTH FUNCTION OF A SHAPE

Let Γ be a curve in the plane, not necessarily a simple closed curve. We consider the following functional introduced by Ambrosio and Tortorelli [1]:

$$\Lambda_\rho(v) = \frac{1}{2}\iint \left\{ \rho\|\nabla v\|^2 + \frac{v^2}{\rho} \right\} dx\, dy \tag{1}$$

subject to the boundary condition $v = 1$ along Γ. Let v denote the unique minimizer of the functional. Then, v varies between 0 and 1 and for sufficiently small values of ρ,

$$v \approx e^{-\frac{d}{\rho}}, \tag{2}$$

where d is the (unsigned) distance from Γ. As $\rho \to 0$, $v \to 0$ everywhere except along Γ. Thus, v may be thought as a blurred version of the characteristic function of Γ and ρ as the nominal blurring radius. The key point is that as $\rho \to 0$, min $\Lambda_\rho(v)$ tends to the length of Γ. For reasons which will become clear in Section 6, we call v the edge-strength function. We compute the minimizer of functional (1) by numerically computing the steady state of the following linear diffusion equation obtained by applying gradient descent functional (1):

$$\frac{\partial v}{\partial \tau} = \nabla^2 v - \frac{v}{\rho^2}. \tag{3}$$

The equation was implemented by means of central finite differences. Alternatively, one can directly solve the steady-state equation $\nabla^2 v = v/\rho^2$, by the finite element method for example.

In order to understand the behavior of the level curves of v, assume that Γ is a simple, closed curve. As in Appendix (3) of [16], inside Γ,

$$v(x, y) = -\rho\left(1 + \frac{\rho\kappa(x, y)}{2}\right)\frac{\partial v}{\partial n}(x, y) + O(\rho^3), \tag{4}$$

where $\kappa(x, y)$ is the curvature of the level curve of v passing through the point (x, y) and n is the direction of the inward normal. Therefore, if we imagine moving from a level curve to a level curve along the normals, then for small values of ρ, a change of Δv in level requires movement

$$\Delta r \approx -\frac{\rho}{v}\left(1 + \frac{\rho\kappa}{2}\right)(\Delta v), \tag{5}$$

where r denotes the arc length along the gradient lines of v (the positive direction being the direction of the inward normals). Let $\Delta t = (\rho^2\, \Delta v/2v)$. (Notice that v is a decreasing function in the direction of the inward normal so that Δv is negative. Functions v and t have the same set of level curves because t is a monotonic function of v.) Passing to the infinitesimals, we get the velocity

$$\frac{dr}{dt} \approx \frac{2}{\rho} + \kappa. \tag{6}$$

A similar argument shows that for moving outward from Γ, the velocity is given by the formula

$$\frac{dr}{dt} \approx -\frac{2}{\rho} + \kappa. \tag{7}$$

Thus for sufficiently small values of ρ, the level curves of v mimic the usual curve evolution which is obtained by letting points on the curve move with velocity consisting of two components: a constant component corresponding to morphology (that is, Blum's grassfire) and a smoothing component proportional to curvature. Thus ρ is a scale parameter analogous to the scale parameter in the Laplacian pyramid or the entropy scale space of Kimia, Tannenbaum, and Zucker.

The above analogy of level curves of the edge-strength function with curve evolution holds only at points where ρ is small compared to the nearby radii of curvature of the level curve through the point and also small compared to the radii of the largest circles passing through the point and contained wholly inside the level curve or wholly outside of it. The differences between the behavior of the level curves of v and the advancing front under pure curve evolution emerge as ρ becomes large. The main difference is that while the velocity of the curve at a point in curve evolution depends only on the curvature at that point, the velocity of the level curves of v also depends on the interaction between nearby points. As a simple example, consider the case of a circle containing a smaller eccentric circle inside it. Thus, the shape is an annulus of varying width. The eccentricity is large so that the minimum width of the annulus is very small compared to its maximum thickness. Under curve evolution, circles remain circles. If ρ is sufficiently large, both the circles shrink toward their centers and eventually disappear without intersecting. Therefore, we assume that ρ is small enough so that eventually, the two circles touch. Note that where they touch each other, the circles are tangent to each other. Thereafter, the annulus breaks at the singular point and continues to evolve as a simple closed curve until it shrinks to a point and disappears. Consider now the level curves of v. Assume that ρ is small compared to the minimum width of the annulus and also compared to the radius of the smaller circle. Then the level curves near the boundary of the

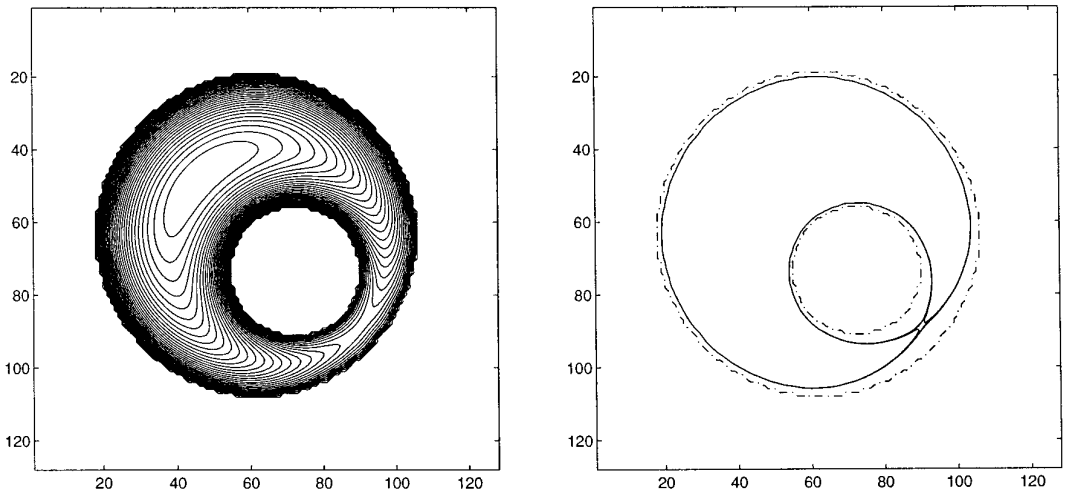

FIG. 1. Annulus: Level curves of v. Notice that at the break, the level curve forms a cross.

annulus closely approximate the initial set of fronts during curve evolution. However, for the level curves further inside the annulus, as the width of the annulus enclosed by a level curve becomes comparable to ρ, the interaction between the opposite boundaries of the annulus becomes significant and the value of v is larger than what it would be without the interaction. As a result, the gradient of v is reduced. In other words, the speed of the evolving curve is increased. Therefore, in the thinnest part of the annulus, the inner circle begins to develop a bulge and the outer circle begins to develop a dent, leading to the formation of a neck and eventual break (see Fig. 1). At the point of the break, the outer and the inner boundaries of the annulus form a cross instead of remaining tangent to each other as in the case of curve evolution. This is definitely a computational advantage as argued in the next section.

Another difference between the approach proposed here and the method of curve evolution is that the edge-strength function formulation does not extend to the limiting cases, namely, the case of pure morphological evolution obtained by omitting from (6) and (7) the curvature term (that is, the limit as $\rho \to 0$) and the case of pure smoothing obtained by setting $\rho = \infty$. This is not a serious disadvantage. On the one hand, the curvature term is essential in order to filter noise; on the other hand, pure smoothing shrinks every shape to a single "round" point [7, 8] and it seems perceptually unnatural to reduce shapes which deviate a great deal from being a circle—shapes such as dumbbells and spirals—to a single point.

The advantages of the approach proposed in this paper over the method of curve evolution are considerable. First of all, v may be calculated by solving a *linear* diffusion equation which is easy to implement by standard finite difference methods. In contrast, for implementing curve evolution, first the initial curve is embedded as zero-crossing in a surface and then all the level curves of the surface are allowed to evolve simultaneously according to the same law of motion. The evolution equation of the surface is nonlinear and because the evolving surface develops shocks, standard finite difference methods (for example, central differences) cannot be used. One must use a shock-capturing scheme such as the one proposed by Osher and Sethian [18] in which the direction of the finite difference depends on the direction in which the shock is developing. Second, as the surface evolves, it traces out a three-dimensional volume and the surface $t(x, y)$ traced out by the evolving zero-crossings is embedded in this volume. The fact that the surface $t(x, y)$ is realized as the locus of zero-crossings of a three-dimensional manifold adds another layer of computational complexity to the task of locating the critical points of its level curves [25].

We now present three examples. Figures 2–4 display selected level curves of v in a sequence of decreasing v. Note that we think of decreasing v as increasing time. The doll figure is on a 128×128 lattice while the cat and the pliers are on a 256×256 lattice. The value of ρ was set equal to 4 for all three examples. As time progresses, protrusions such as the ears of the cat disappear, shape splits into parts, parts shrink and evolve toward more circular shapes, and eventually disappear. Smaller parts disappear before the larger parts. Shapes split into parts along necks as seen in the example of the cat where the tail separates from the body. The head connected to the body by a much thicker neck disconnects much later than the tail. Clearly, v contains information regarding parts, protrusions, and necks. How to represent this information in a meaningful way is explained in the next section.

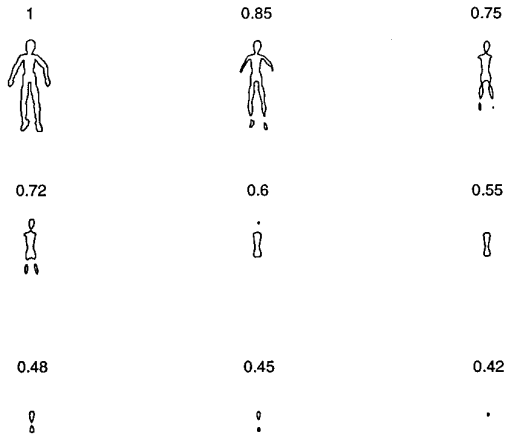

FIG. 2. Level curves for the doll shape; parts break off and disappear as we move toward the inner level curves ($\rho = 4$).

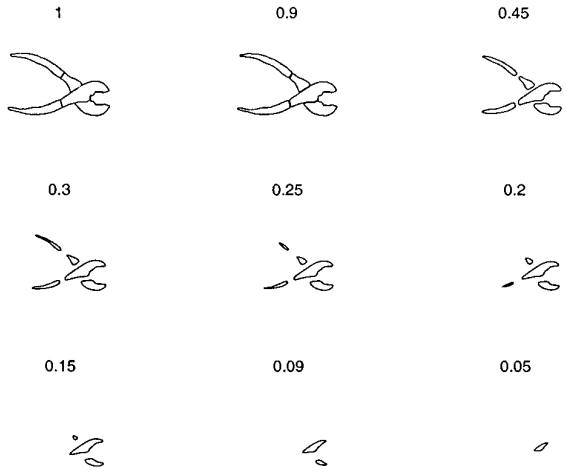

FIG. 4. Level curves for pliers; method can handle self-intersections and triple points ($\rho = 4$).

3. SHAPE SKELETON AND SHAPE DECOMPOSITION

In the purely morphological evolution (Blum's grassfire), singularities develop as corners and self-intersections form. The locus of these singularities is the skeleton of the shape. When smoothing is introduced, self-intersections still may develop (due to thinning of narrow necks), but the corners are rounded out. Therefore, when smoothing is present, points of maximum curvature serve as a substitute for corners. Now we note from Eq. (4) that for small values of ρ,

$$\left(\frac{2}{\rho} + \kappa\right) \cdot \left|\frac{\partial v}{\partial n}\right| \approx \frac{2v}{\rho^2} \tag{8}$$

so that along a level curve, the points of maximum curvature correspond approximately to the points where $|\nabla v|$ is minimum. This observation suggests an alternative way of defining shape skeleton, namely, by determining points where $|\nabla v|$ is minimum along the level curves of v. We prefer this alternative because computation of curvature involves second derivatives of v and hence more sensitive to noise than $|\nabla v|$.

Let K^+ denote the closure of the set of zero-crossings of $d|\nabla v|/ds$ where $d^2|\nabla v|/ds^2$ is positive. Here, s denotes the arc-length along the level curves.

$$\begin{aligned} \frac{d|\nabla v|}{ds} &= v_{\eta\xi} \\ \frac{d^2|\nabla v|}{ds^2} &= v_{\eta\xi\xi} + \frac{v_{\xi\xi}(v_{\xi\xi} - v_{\eta\eta})}{|\nabla v|}, \end{aligned} \tag{9}$$

where

$$\begin{aligned} v_{\eta\xi} &= \frac{\{(v_y^2 - v_x^2)v_{xy} - v_x v_y (v_{yy} - v_{xx})\}}{|\nabla v|^2} \\ v_{\xi\xi} &= \frac{\{v_y^2 v_{xx} - 2v_x v_y v_{xy} + v_x^2 v_{yy}\}}{|\nabla v|^2} \\ v_{\eta\eta} &= \frac{\{v_x^2 v_{xx} + 2v_x v_y v_{xy} + v_y^2 v_{yy}\}}{|\nabla v|^2} \\ v_{\eta\xi\xi} &= \frac{1}{|\nabla v|^3}\{v_x v_y^2 v_{xxx} + v_y(v_y^2 - 2v_x^2)v_{xxy} \\ &\quad + v_x(v_x^2 - 2v_y^2)v_{xyy} + v_x^2 v_y v_{yyy}\}. \end{aligned} \tag{10}$$

Note that $v_{\xi\xi}/|\nabla v|$ is the curvature of the level curves and $v_{\eta\xi}/|\nabla v|$ is the curvature of the gradient lines.

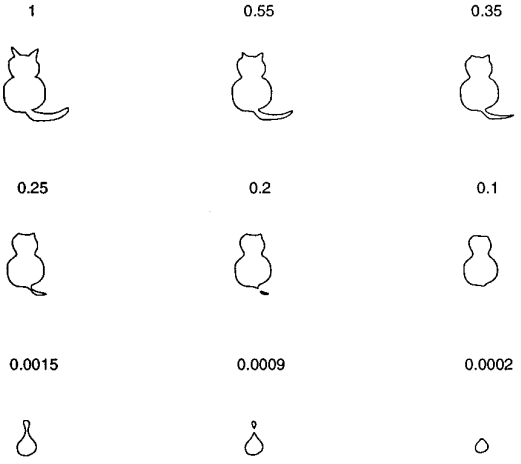

FIG. 3. Level curves for the cat shape; parts break off and disappear as we move toward the inner level curves ($\rho = 4$).

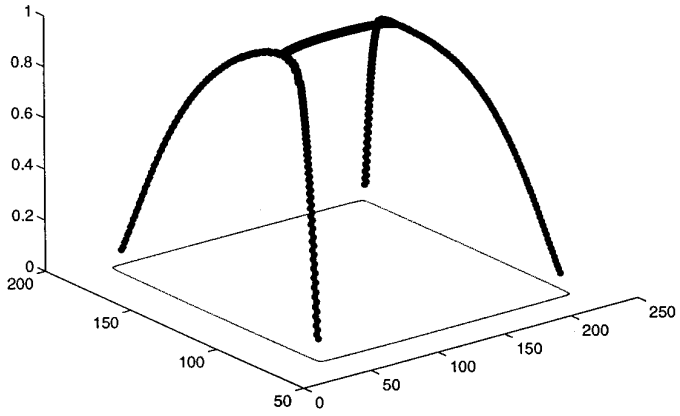

FIG. 5. $1 - v$ plotted along the skeleton; "CORE" of a rectangle.

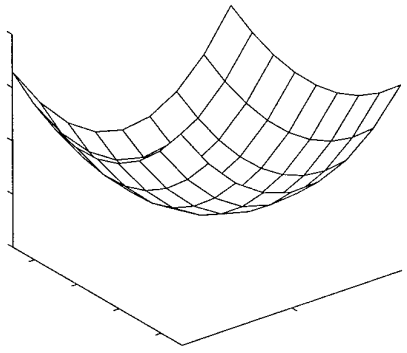

FIG. 7. Surface v near an elliptic point.

Dually, let K^- denote the closure of the set of zero-crossings of $d|\nabla v|/ds$ where $d^2|\nabla v|/ds^2$ is negative.

Let $K = K^+ \cup K^-$.

Let S denote the set of points where $|\nabla v| = 0$.

The direction of evolution at each point of K is the direction of decreasing v. The branches of K should be thought of as local axes of symmetry. For a perfect circle, K^+ and K^- are empty and S consists of a single point where v has its unique minimum. A simple closed curve may be thought of as a deformation of a circle by means of protrusions and indentations. As it evolves toward a more circular shape, K^+ tracks evolution of its protrusions, while K^- tracks evolution of its indentations. During the evolution, a protrusion might merge with an indentation, $(d|\nabla v|/ds = d^2|\nabla v|/ds^2 = 0)$, joining a branch of K^+ with a branch K^- and terminating both the branches. Of course, more complicated merges between the branches of K^+ and the branches of K^- are also possible and in such a case, a new branch might start from the junction. It is also possible that a branch might bifurcate. This typically happens during the thinning of a neck. As two indentations on the opposite sides of a neck begin to evolve toward each other, each one gives rise to a branch of K^-. However, as the indentations approach a break point, they interact and slow down the rate of decay of v. Each branch of K^- splits into three branches, the middle branch belongs to K^+ and continues toward the break point, while the other two belong to K^-. If a branch is not terminated at a junction with another branch, then it will terminate at a point in S. If the point is a minimum point of v, then the evolution has come to rest at that point and it is appropriate to call such a point the center of a part. If the point is not a minimum, then it may signify a change in the topology of the evolving curve, that is, break-up of the shape due to a thinning neck. If the point signifies a change in the topology of the shape, we will call it a saddle point. There are at least two branches of K^+ leaving such a point.

In differential geometric terms, a point in S is elliptic, hyperbolic, and parabolic depending on whether the determinant of the hessian of v, $v_{xx}v_{yy} - v_{xy}^2$, is positive, negative, or zero, respectively. An elliptic point is always a center and a hyperbolic point is always a saddle point. Parabolic

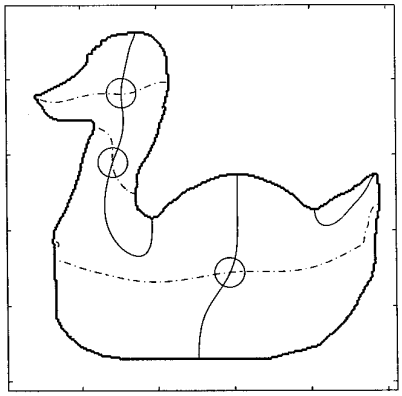

FIG. 6. Zero-crossings of v_x and v_y. There are three points in S corresponding to the neck and the centers of the head and body ($\rho = 16$).

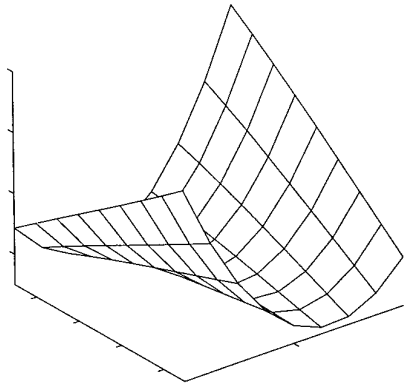

FIG. 8. Surface v near a hyperbolic point.

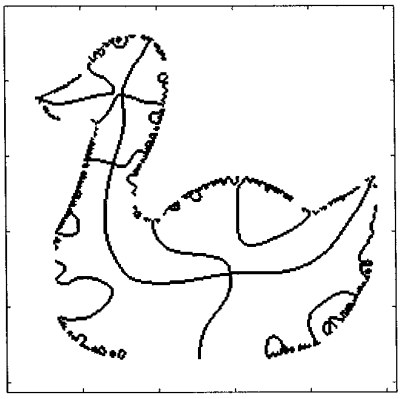

FIG. 9. K for the duck shape ($\rho = 16$).

points are more troublesome to classify because of their global nature. We must analyze the branches of K meeting the parabolic point (or the connected component of S containing the parabolic point in case it is not an isolated parabolic point). If we can find at least two branches of K^+ leaving from the connected component of S containing a given parabolic point, then it is a saddle point. If all the branches of K meeting the connected component containing the parabolic point lead into it, it is a center.

Now we can define the skeleton of a shape. A strict definition of the skeleton is that it is the subset of $K^+ \cup S$ which excludes those branches of K^+ which flow into a connected component of S containing a saddle point. The definition is designed so as to exclude the branches of K^+ along which necks of the shape evolve towards a break. However, strict skeletons are hard to calculate because one has to trace the branches of K. In order to simplify the calculation, we note that as pointed out before, at a break point, due to interaction between the opposite boundaries on the two sides of a neck, the boundaries tend to form a cross rather than be tangent to each other. In other words, at break points, v tends to have hyperbolic points rather than parabolic points. The advantage of having a hyperbolic point is that the curvature of the level curve along the branch K^+ flowing into the saddle point is always negative while it is positive along the branch flowing out. Therefore, we adopt a more managable definition of the skeleton as follows:

DEFINITION. The *skeleton* of a shape is the union of S and those branches of K^+ along which the curvature of the level curves is positive.

We define the segmentation of a shape corresponding to its break-up into parts during evolution due to the presence of narrow necks as follows:

DEFINITION. The *segmentation* of a shape is the union of branches of K^+ along which the curvature of the level curves is negative.

(A strict segmentation is the union of branches of K which flow into a saddle point.) It is interesting to note that the method of curve evolution will fail to find the segmentation curve in the example of the annulus described in the previous section, even if we use the definition of strict segmentation, because the level curves do not have critical points before the annulus breaks.)

Note that since a branch of K^+ may terminate at a junction with a branch of K^-, the skeleton need not be connected. In our description, the skeleton always extends to the boundary while in the purely morphological evolution, a branch starts only after a corner has formed. Note also that the definition of segmentation does not deal with protrusions. Significant protrusions such as the fingers of a hand must be recovered as parts of the shape from the branches of the skeleton.

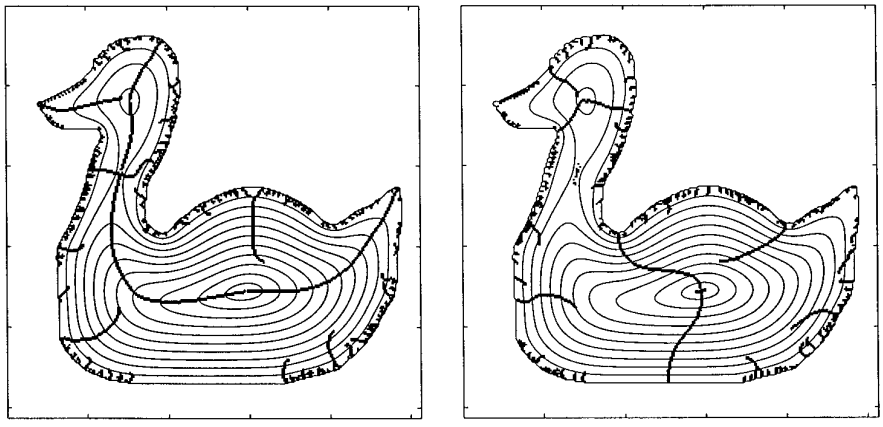

FIG. 10. K^+ (left) and K^- (right) with level curves superimposed.

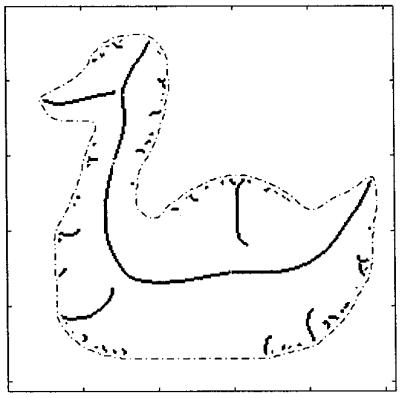

FIG. 11. Skeleton of the duck shape ($\rho = 16$).

The skeleton and the segmentation depend on the choice of ρ and should be computed for many values of ρ to provide a scale space representation, just as it is done in other contexts such as the Laplacian pyramid or the entropy scale space of Kimia, Tannenbaum, and Zucker. The branches of segmentation usually will not extend to the boundary of the shape because they may start out as branches of K^- or level curves along the branch may initially have positive curvature. The length of the segmentation branch depends on ρ. The larger the value of ρ, the longer the branch. The definition of strict segmentation has to be used if the segmentation must be extended to the boundary.

This description can be readily translated into the language of the shock grammar of Siddiqi and Kimia [10, 25]. The first-order shocks are the branches of the skeleton not belonging to S. The second-order shocks are the hyperbolic points. The third-order shocks correspond to a line of parabolic points. The elliptic points are the fourth-order shocks. The rules of the grammar and properties follow easily from the fact that v is monotonically decreasing along the branches of the skeleton and the fact the v is smooth.

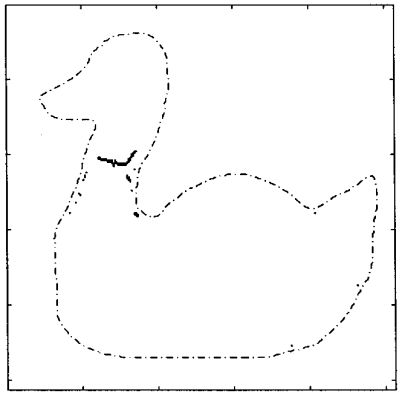

FIG. 12. Segmentation of the duck shape ($\rho = 16$).

Our work is also related to GRADMAT of Rosenfeld [28] and the related more recent development of cores by Burbeck and Pizer [4]. In our context, the value of $1 - v$ measures the medialness of a point on the skeleton. Our version of the core may be visualized by plotting the values of $1 - v$ along the skeleton. Figure 5 shows such a plot for a rectangle.

We now illustrate our construction in detail through the duck example. This simple shape covers all the phenomena that may arise. Figure 6 shows computation of S, namely the points where the gradient of v vanishes. We detect these points as the simultaneous zero-crossings of the derivatives v_x and v_y. In the case of the duck figure, S consists of three points: two elliptic points corresponding to the centers of the head and the body and one hyperbolic point corresponding to the neck. Figures 7 and 8 depict the surface v in the vicinity of one of the centers and in the vicinity of the saddle point in the neck.

Figure 9 shows all of K with $\rho = 16$. Figure 10 shows K^+ and K^- separately with level curves of v superimposed, illustrating that K^+ tracks protrusions, while K^- tracks indentations. As we move toward the inner level curves in the direction of decreasing v, the level curves deform toward a more circular form and some branches of K terminate. Figures 11 and 12 show the skeleton and the segmentation which are what one would expect.

Figure 13 shows skeletons of various shapes. (None of them happens to involve segmentation.) Examples of a pair of pliers and the outline of a cube illustrate the effectiveness of the method for complex shapes involving junctions.

In these examples, we did not do any pruning.

4. MISSING INFORMATION AND FILLING THE GAPS

Although the construction described above were motivated by the example of simple closed curves, they work equally well for more general curves. For example, if the letter C is drawn as an open curve, we can compute its skeleton. If the ends of the letter are sufficiently close, a segmentation line joins the two ends. If the same letter is drawn as a thick shape with a simple closed curve as its boundary, we recover the same information as above and also the medial axis. The point is that v should be computed over the whole plane and the skeleton and the segmentation really describe the complement of Γ. Whether the complement of Γ is connected or not is not relevant to the computation of the skeleton and the segmentation. This property is very useful in particular when the shape boundary is not fully specified. As long as important features of the shape boundary are specified, evolution fills in the gaps and the essential skeleton can be recovered. The missing

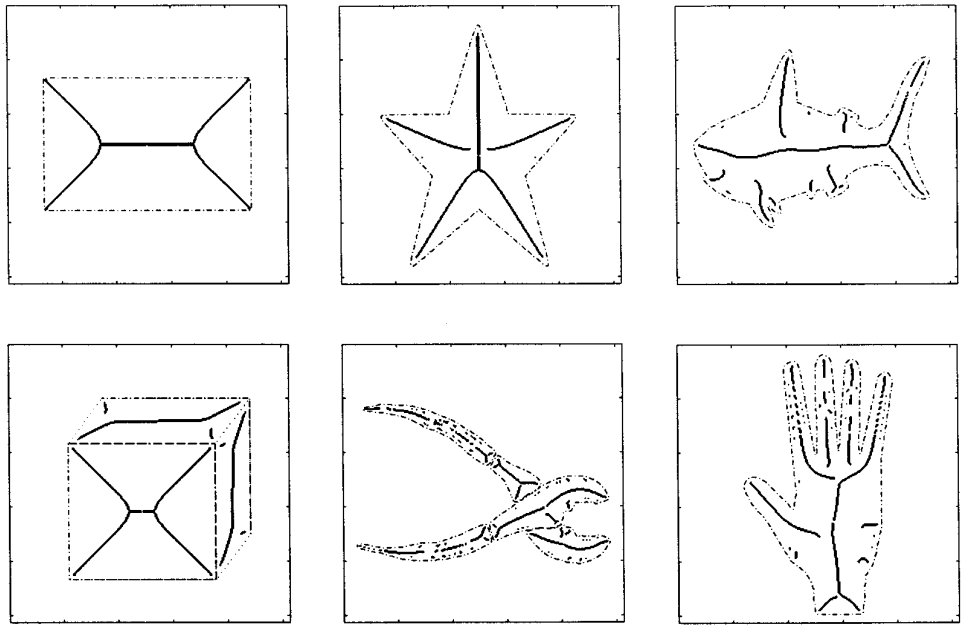

FIG. 13. Skeletons of various shapes ($\rho = 16$).

portions of the boundary are recovered as branches of the segmentation.

We illustrate our construction by means of an incomplete rectangle, obtained by removing a 56-pixel-long piece from the middle of each of the two long sides of a complete 175×140 rectangle. Figure 14 shows the skeleton and the segmentation obtained using $\rho = 16$. The skeleton consists of the skeleton of the complete rectangle and two new branches which may be interpreted as the "medial axes" of the gaps in the sides. The larger the value of ρ, the shorter these additional branches. The gaps themselves are filled in by the segmentation lines.

5. SIGNIFICANCE OF A SKELETON BRANCH

There are several ways in which one may assign a level of significance to a point on the skeleton. Branches tracking less important details terminate earlier than the branches tracking globally more prominent details. Therefore the value of $1 - v$ at the termination point is one measure of significance of a branch. Since $1 - v$ may be viewed as corresponding to time in curve evolution, our first measure is the "time of extinction" of a branch.

Another measure of significance is the curvature of the level curve at points on the skeleton. Thus, along a level curve, a skeleton point with large curvature indicates a more significant protrusion than the one with smaller curvature.

Finally, one can compare robustness of the skeleton branches with respect to increasing values of ρ. With increasing ρ, less significant branches become shorter, while the most prominent branches remain essentially unchanged. This effect is clearly seen in Fig. 15 which shows skeletons for different values of ρ.

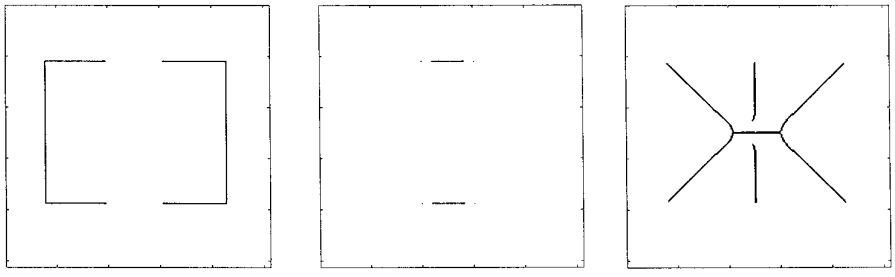

FIG. 14. Example of an incomplete shape ($\rho = 16$). Left, an incomplete rectangle; middle, segmentation; right, skeleton. Note that the segmentation line fills the gaps while skeleton acquires extra branches corresponding to the two gaps.

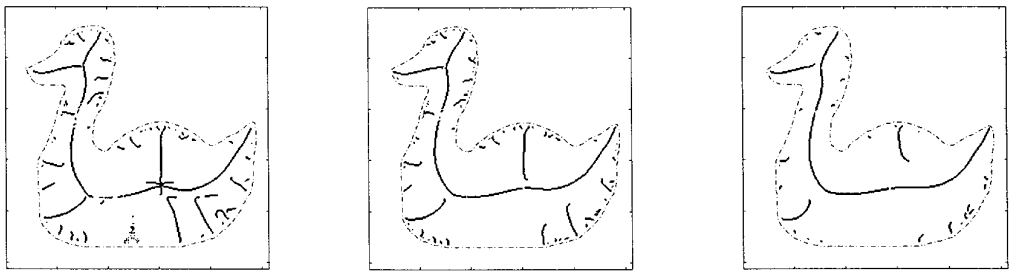

FIG. 15. Effect of ρ on the skeleton: ρ is set to 4 (left), 8 (middle), and 32 (right). Notice how less important branches shrink as ρ gets larger.

6. GRAYSCALE IMAGES

The preceding sections dealt with the case in which the shape outline is already extracted in the form of a line drawing although it may be incomplete. However, segmentation of a grayscale image for obtaining such a line drawing is not an easy task. In this section, we extend the previous constructions to grayscale images in a natural way, without requiring the intermediate step of extracting a line drawing. We start with the segmentation functional introduced in [15]. Even finding a segmentation which only approximately minimizes this functional is an extremely difficult task. Instead of looking for approximate minimizers, we employ its elliptic approximation introduced by Ambrosio and Tortorelli [1]. They achieve this by replacing the segmentation curve by a continuous function which we call the edge-strength function. Skeletons and segmentation are then extracted from the level curves of this edge-strength function.

The segmentation functional is

$$E_{MS}(u,B) = \alpha \int\int_{R\backslash B} \|\nabla u\|^2 \, dx\, dy + \beta \int\int_R (u-g)^2 \, dx\, dy + |B|, \tag{11}$$

where R is a connected, bounded, open subset of $\mathbf{R}^2$ (usually a rectangle), g is the feature intensity, B is a curve segmenting R, u is the smoothed image $\subset R\backslash B$, $|B|$ is the length of B, and α, β are the weights. Let $\sigma = \sqrt{\alpha/\beta}$. Then, σ may be interpreted as the smoothing radius in $R\backslash B$. With σ fixed, the higher the value of α, the lower the penalty for B and hence, the more detailed the segmentation.

Ambrosio and Tortorelli [1] replace

$$|B| \quad \text{by} \quad \frac{1}{2}\int\int_R \left\{\rho\|\nabla v\|^2 + \frac{v^2}{\rho}\right\} dx\, dy \tag{12}$$

and

$$\int\int_{R\backslash B} \|\nabla u\|^2 \, dx\, dy \quad \text{by} \quad \int\int_R (1-v)^2 \|\nabla u\|^2 \, dx\, dy. \tag{13}$$

The result is the functional

$$E_{AT}(u,v) = \int\int_R \{\alpha(1-v)^2\|\nabla u\|^2 + \beta(u-g)^2 + \frac{\rho}{2}\|\nabla v\|^2 + \frac{v^2}{2\rho}\Big\} dx\, dy. \tag{14}$$

As $\rho \to 0$, v converges to 1 at points on B and to zero elsewhere. The corresponding gradient descent equations are

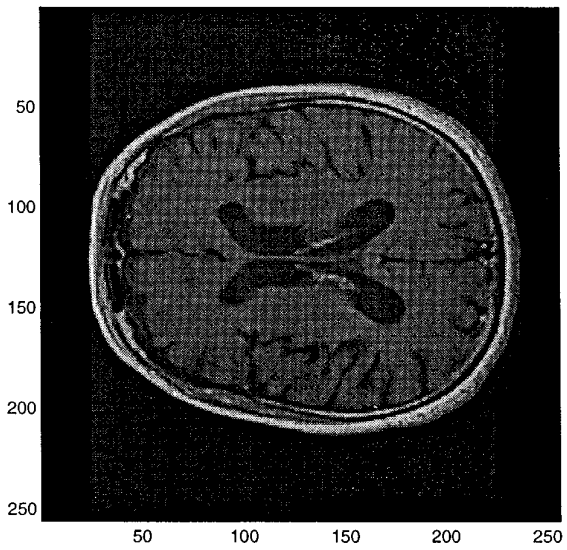

FIG. 16. MRI image.

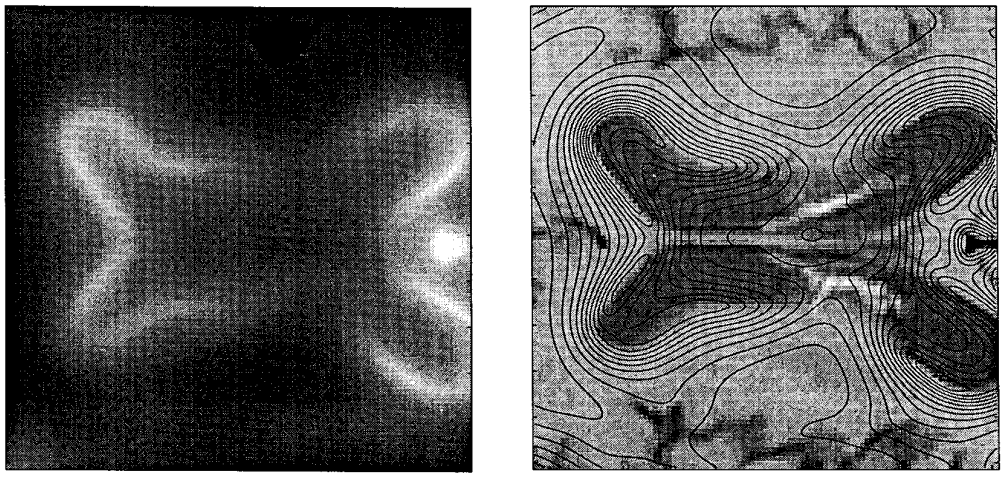

FIG. 17. Ventricle area: left, edge-strength function v_a; right, its level curves ($\rho = 8$).

$$\frac{\partial u}{\partial t} = \nabla \cdot (1 - v)^2 \nabla u - \frac{\beta}{\alpha}(u - g)$$

$$\frac{\partial v}{\partial t} = \nabla^2 v - \frac{v}{\rho^2} + \frac{2\alpha}{\rho}(1 - v)\|\nabla u\|^2 \qquad (15)$$

$$\frac{\partial u}{\partial n}|_{\partial R} = 0 \qquad \frac{\partial v}{\partial n}|_{\partial R} = 0,$$

where ∂R denotes the boundary of R and n denotes the direction normal to ∂R. The solution of these equations gives us the edge-strength function v corresponding to the segmentation locus B without determining B itself.

Notice that equation for each variable is a diffusion equation which minimizes a convex quadratic functional in which the other variable is kept fixed:

Keeping v fixed, the first equation minimizes

$$\int\int_R \{\alpha(1 - v)^2\|\nabla u\|^2 + \beta(u - g)^2\}\, dx\, dy \qquad (16)$$

Keeping u fixed, the second equation minimizes

$$\int\int_R \left\{\|\nabla u\|^2 + \frac{1 + 2\alpha\rho\|\nabla v\|^2}{\rho^2}\left(v - \frac{2\alpha\rho\|\nabla u\|^2}{1 + 2\alpha\rho\|\nabla u\|^2}\right)^2\right\} dx\, dy.$$

Thus the edge strength function v is nothing but a nonlinear smoothing of

$$\frac{2\alpha\rho\|\nabla u\|^2}{1 + 2\alpha\rho\|\nabla u\|^2}, \qquad (17)$$

where u is a simultaneous nonlinear smoothing of g.

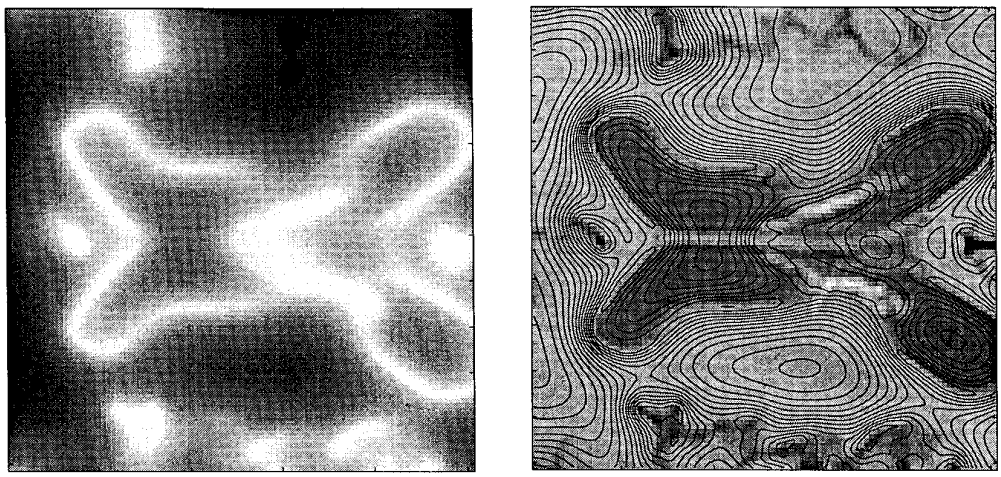

FIG. 18. Ventricle area: left, edge-strength function v_b; right, its level curves ($\rho = 8$).

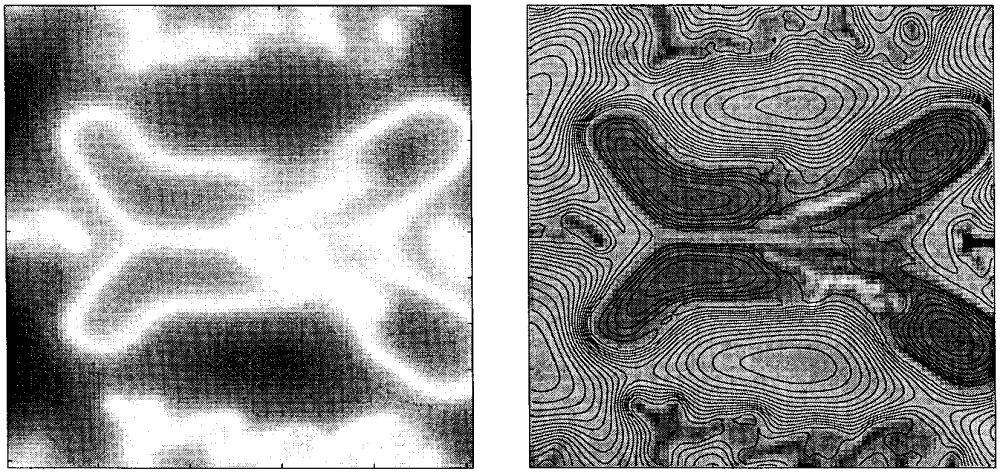

FIG. 19. Ventricle area: left, edge-strength function v_c; right, its level curves ($\rho = 8$).

In our integrated framework, there are three separate scales represented by parameters α, β, and ρ. Thus we have an integrated 3-dimensional scale space. All of these scales are also present in all other approaches as well, but separately. To take a particular case, consider the method of shape recovery and construction of skeletons by curve evolution. First, the grayscale image is presmoothed for determining the gradient from which the edge-strength function (boundariness function) is computed. Next, curve evolution is applied (for example, an intrinsic version of Terzopoulos' snakes) to detect the shape boundary. This also involves a scale in the form of ratio between the curvature component and the morphology component of the curve velocity. This determines the extent of smoothing of the boundary. Finally, the skeleton is extracted by applying curve evolution again to the extracted boundary, introducing a third scale determining the amount of detail in the skeleton. The issue of automatic scale selection is a research topic in its own right. Even when there is only one scale involved such as in the case of image smoothing by Gaussians, automatic scale selection is difficult as the current research by various investigators such as Koenderink, ter Romney, and Lindeberg show. Our purpose is to demonstrate a general method. For practical purposes, we suggest that scales be chosen empirically in the context of the particular application at hand.

Ideally, the edge-strength function v computed from a raw image by Eq. (15) should be constant along the object boundaries so that its level curves will correspond to the ideal cases discussed in Section 2. However, this almost never happens due to noise, differing levels of contrast along the object boundaries, and the interaction between nearby edges. Therefore, the object boundaries are no longer defined by level curves of v. Typically, we should expect a level curve corresponding to a value of v near its maximum to consist of several connected components, each surrounding a high contrast portion of B. The situation is analogous to the earlier example of the incomplete rectangle in Section 4 where the shape boundary consists of several disconnected pieces. Note that even in the case of the functional (11), if α is not high enough, B may not include portions of the object boundary where contrast is

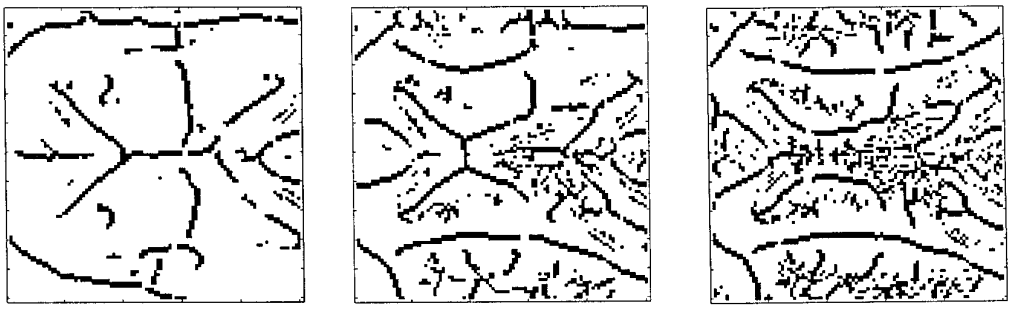

FIG. 20. Skeletons for the ventricle area corresponding to v_a (left), v_b (middle), and v_c (right).

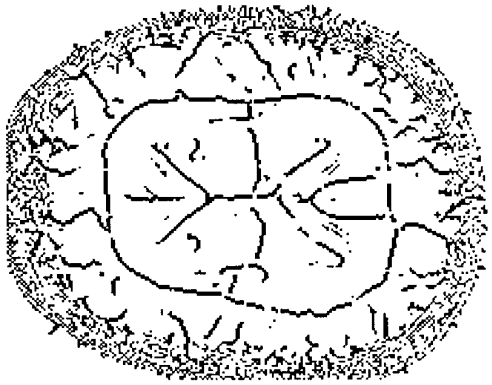

FIG. 21. Skeleton of all of MRI image corresponding to v_a.

too low. The important point is that it is not essential to have the complete shape boundary to compute its skeleton. As the evolution progresses, the gaps between the pieces of the boundary are filled in. Thus we still recover an essentially correct skeleton. As pointed out before, as the value of α is increased, B becomes more and more detailed, adding to the skeletal details. This effect on the skeleton is distinct from the effect of varying ρ discussed in Section 5 where the details of the initial shape are fixed and the skeleton is smoothed as ρ is increased.

Before computing the skeleton, the definition of S must be modified by excluding the points where v is maximum. The definition of the skeleton applies unchanged to raw images because along the branches of K^+ emanating from the maxima of v, the level curves of v have negative curvature and hence are automatically excluded from the skeleton. However, they do become part of the segmentation. Note that now we have segmentation of the grayscale image given by B and the segmentation of the shape into parts as defined in Section 3. The branches of K^+ along which the level curves have negative curvature included both of these segmentations.

In order to illustrate application of the method to grayscale images, we present an example shown in Fig. 16 and analyze the dark shape (ventricles) in the center of the figure. To compute the edge-strength function by Eq. (15), we set $\sigma = \rho = 8$ and picked three different values of α, obtaining three different edge-strength functions, v_a, v_b, and v_c depicted in Figs. 17–19. The value of α corresponding to v_a is sufficiently low so that only the corner portions of the central dark area are prominent in the image of v_a. The inner details such as the inner boundaries of the four disjoint corner lobes are smoothed over to a large degree. The value of α was doubled to obtain v_b and doubled again for v_c. At each stage, the edge-strength function becomes more detailed. Figure 20 shows the skeletons for the ventricle corresponding to the three values of α. The effect of increasing α is clearly seen in the bottom row of Fig. 20. The skeleton in the case of v_a is essentially that of an incomplete rectangle. As α is increased, it becomes more detailed so that the skeleton from v_c depicts the axes of the four lobes more accurately and finds a center for each lobe. Figure 21 shows the skeleton for the whole MRI image corresponding to v_a.

REFERENCES

1. L. Ambrosio and V. M. Tortorelli, On the approximation of functionals depending on jumps by quadratic, elliptic functionals, *Boll. Un. Mat. Ital.,* 1992.
2. H. Blum, A transformation for extracting new descriptions of shape, in *Models for Perception of Speech and Visual Form* (W. Wathen-Dunn, Ed.), MIT Press, Cambridge, MA, 1967.
3. H. Blum, Biological shape and visual science, *J. Theo. Biol.* **38,** 1973, 205–287.
4. C. Burbeck and S. Pizer, *Object Representation by Cores: Identifying and Representing Primitive Spatial Regions,* TR94-160, Department of Computer Science. University of North Carolina at Chapel Hill, 1994.
5. V. Caselles, F. Catte, T. Coll, and F. Dibos, A geometric model for active contours, *Numer. Math.* 1993, 66.
6. V. Caselles, R. Kimmel, and G. Sapiro, Geodesic active contours, in *Fifth International Conference on Computer Vision, 1995.*
7. M. Gage, On an area-preserving evolution equation for plane curves, *Contemp. Math.* **51,** 1986, 51–62.
8. M. Gage and R. S. Hamilton, The heat equation shrinking convex plane curves, *J. Diff. Geom.* **23,** 1986, 69–96.
9. B. B. Kimia, A. Tannenbaum, and S. W. Zucker, On the evolution of curves via a function of curvature. I. The classical case, *J. Math. Anal. Appl.* **163**(2), 1992.
10. B. B. Kimia, A. Tannenbaum, and S. W. Zucker, Shapes, shocks and deformations. I. The components of shape and the reaction–diffusion space, *Int. J. Comput. Vision,* 1992.
11. B. B. Kimia, A. Tannenbaum, and S. W. Zucker, Entropy scale-space, in *Visual Form* (C. Arcelli, *et al.,* Ed.), Plenum, New York, 1991.
12. S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Gradient flows and geometric active contour models, in *Fifth International Conference on Computer Vision, 1995.*
13. K. Koffka, *Principles of Gestalt Psychology,* Harcourt Brace, New York, 1935.
14. R. Malladi, J. A. Sethian, and B. C. Vemuri, Shape modeling with front propagation: A level set approach, *IEEE Trans. Pattern Anal. Mach. Intell.* **17,** 1995.
15. D. Mumford and J. Shah, Boundary detection by minimizing functionals, in *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, 1985.*
16. D. Mumford and J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, *Commun. Pure Appl. Math.* **XLII**(5), July 1989, 577–684.
17. R. L. Ogniewicz, Skeleton-space: A multiscale shape description combining region and boundary information, in *IEEE Conference on Vision and Pattern Recognition, 1994.*
18. S. Osher and J. Sethian, Fronts propagating with curvature dependent speed: Algorithms based on the Hamilton–Jacobi formulation, *J. Comp. Phys.* **79,** 1988.
19. P. J. Olver, G. Sapiro, and A. Tannenbaum, Differential invariant signatures and flows in computer vision: A symmetry group approach,

in *Geometry Driven Diffusion in Computer Vision,* Kluwer, Dordrecht, 1994.

20. J. Sethian, *An Analysis of Flame Propagation,* Ph.D. Thesis, Univ. of California, Berkeley, CPAM Report 79, 1982.
21. J. Shah, Uses of elliptic approximations in computer vision, in *Variational Methods for Discontinuous Structures* (R. Serapioni and F. Tomarelli, Eds.), Birkhäuser, Basel, 1996.
22. J. Shah, Shape recovery from noisy images by curve evolution, in *IASTED International Conference on Signal and Image Processing, Nov. 1995.*
23. J. Shah, A common framework for curve evolution, segmentation and anisotropic diffusion, in *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, 1996.*
24. K. Siddiqi and B. B. Kimia, Parts of visual form: Computational aspects, in *IEEE Conference on Vision and Pattern Recognition, 1993.*
25. K. Siddiqi and B. B. Kimia, A shock grammar for recognition, in *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, 1996.*
26. S. Tari, J. Shah, and H. Pien, A computationally efficient shape analysis via level sets, in *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, June 1996.*
27. H. Tek and B. B. Kimia, Image segmentation by reaction–diffusion bubbles, in *Fifth International Conference on Computer Vision, 1995.*
28. S. Wang, A. Rosenfeld, and A. Y. Wu, *A Medial Axis Transformation for Grayscale Pictures,* TR-843, Computer Vision Lab, Computer Science Center, University of Maryland, December, 1979.