

An Active Contour Model for Mapping the Cortex

Chris A. Davatzikos and Jerry L. Prince

Abstract—A new active contour model for finding and mapping the outer cortex in brain images is developed. A cross-section of the brain cortex is modeled as a ribbon, and a constant speed mapping of its spine is sought. A variational formulation, an associated force balance condition, and a numerical approach are proposed to achieve this goal. The primary difference between this formulation and that of snakes is in the specification of the external force acting on the active contour. A study of the uniqueness and fidelity of solutions is made through convexity and frequency domain analyses, and a criterion for selection of the regularization coefficient is developed. Examples demonstrating the performance of this method on simulated and real data are provided.

I. INTRODUCTION

DETERMINING the location of the brain surface from MR images is often a first step in brain visualization [1], [2], quantitative analysis of brain properties, multimodal registration [3], and mapping and unfolding the cortex [4]–[6]. Contours representing the brain outer surface are usually determined either automatically through edge detection and linking [7], [8] or contour following [9], [10], or are drawn by hand [4]. Typically, edge detection methods suffer from the identification of either too few or too many edge points, creating problems in edge linking. Contour following methods are typically not very robust to noise and cannot easily be made to satisfy known boundary conditions. Finally, hand-drawn contours require a great deal of time and are generally not reproducible.

An alternate and potentially very powerful approach based on active contours [11]–[16] has been adapted to brain surface extraction by several investigators in recent years [17]–[20]. In this approach, an initial curve is placed somewhere near the cortex. By modeling this initial curve as a physical object, which is called an *active contour*, and the data as an external force to which the object is attracted, an iterative procedure is initiated to cause the active contour to move toward the data and ultimately conform to it. An active contour algorithm is characterized by three parts: 1) a model of the internal forces, e.g., elasticity and bending moments, which describes

the active contour as a physical object; 2) a model of the external forces which describes how the active contour is attracted to the data; and 3) an iterative procedure which attempts to find the configuration that best matches both the internal and external forces. This contour should satisfy a force balance condition, which are the Euler equations of a variational formulation, and should also minimize the energy function of this same formulation.

Active contour methods offer several advantages. First, an active contour is modeled directly as a curve and is maintained as a curve throughout the iterative process that deforms it toward the final solution. Thus, characteristics of the desired curve such as its length, curvature, and conformation to the data can be evaluated or imposed as an explicit part of the algorithm. Second, an optimality criterion involving both intrinsic properties of the curve and the curve's relationship to the data is specified, and an optimal solution is sought. Third, an explicit map between the curve and the unit interval is generated automatically. Such a map can be used to determine properties of the cortex such as lengths, tangents, normals, and curvature using the theory of differential geometry. A final benefit of this approach is that it can be readily extended to 3-D by defining deformable surfaces instead of curves. Some work along these lines has already been reported [21]–[25].

A potential weakness of active contour methods is that their convergence properties are poorly understood. The two central questions are: will the active contour algorithm converge to a unique solution, and if so, will this solution be near the truth? Some grasp of the answers to these questions is essential if active contours are to be used for either a scientific or clinical objective. Several authors have begun to study convergence (cf., [12], [26]–[28]). It has been determined that active contour models are nonconvex and that solutions are often local rather than global minima.

In this paper, we present a new active contour algorithm for ribbons (ACAR), which is specifically designed for mapping the human cortex. The model uses only elastic internal forces and is therefore classified as a *string* model. The more common *snake* model [12] has bending moments, which we believe are undesirable for brain mapping where conforming to the sharp folds of the sulci is highly desirable. The external force used in ACAR is motivated by the fact that the cortex has a nearly constant width (around 5 mm) throughout its extent. We require that if a small disk centered at a point on the active contour rests entirely within the cortex, it experiences no external force; if, on the other hand, a portion of the disk intersects adjacent tissue, the disk experiences a force drawing it back toward the cortex. This external force represents a significant departure from that of snakes for two

Manuscript received July 28, 1993; revised May 25, 1994. This work was supported in part by an NSF Presidential Faculty Grant (MIP-9350336) and by NIH Grant NICHD P50HD 25806-03. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was J. Duncan.

C. A. Davatzikos was with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218 USA. He is now with the Department of Radiology and Radiological Science, Johns Hopkins Medical School, Baltimore, MD 21287 USA; e-mail: hristos@welchlink.welch.jhu.edu.

J. L. Prince is with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218 USA; e-mail: prince@machie.ece.jhu.edu.

IEEE Log Number 9408763.

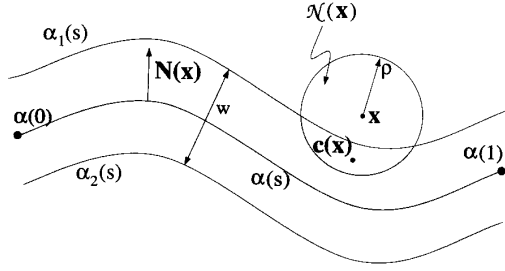


Fig. 1. A ribbon C is defined by its central layer $\alpha(s)$ and its thickness w . The point \mathbf{x} is close enough to the curve so that its neighborhood $\mathcal{N}(\mathbf{x})$ intersects C , and its center of mass function $\mathbf{c}(\mathbf{x})$ is defined.

reasons. First, it is derived from an integration of the image data rather than a differentiation. This integration leads to a certain degree of robustness to noise which we demonstrate in simulations. Second, our external force is designed to draw the active contour toward the spine of a ribbon, not toward a boundary between two objects as in snakes. This is a fundamental difference which precludes direct comparison of the performance of ACAR to that of snakes.

The remainder of this paper is organized as follows. In Section II, we introduce the terminology and notation of ribbons and define a new external force designed for ribbons. In Section III, we present our new algorithm ACAR, which is motivated by an intuitive force balance condition and completed by the development of an associated variational principle. A condition on the regularization coefficient which guarantees both uniqueness and fidelity of ACAR solutions is developed in Section IV. In Section V, we give examples of the performance of ACAR using both simulated and real data, and finally in Section VI, we summarize our results and discuss directions for future research.

II. PRELIMINARIES

A. Ribbons

In a cross-sectional image of the brain, the band of cortex beneath the outer cortical surface is continuous and has nearly constant thickness. This type of object has been studied in the computer vision literature [29], [30], where it is called a *ribbon*. A ribbon can be defined using its spine, which we define as the constant speed plane curve $\alpha(s)$, and a thickness $w(s)$, both defined on the unit interval $[0, 1]$. As depicted in Fig. 1, the ribbon is the set between two boundary curves, and is given by

$$C = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x} = \alpha(s) + \lambda w(s) \mathbf{N}(s)/2, \\ 0 \leq s \leq 1, -1 \leq \lambda \leq 1\},$$

where $\mathbf{N}(s)$ is the unit normal vector of $\alpha(s)$.

The goal of our algorithm is to find the spine of a ribbon using an active contour algorithm. The following observations are relevant to this goal:

- 1) The spine of a ribbon is a curve parameterizing its skeleton. It is therefore possible to approach this problem as one of skeletonization followed by parameterization.

Most skeletonization methods, however, are sensitive to noise and presegmentation errors (cf., [31]–[34]), both of which are inevitable in any medical imaging problem. Under these conditions, it is highly likely that a skeleton could not be parameterized as a simple curve. In contrast, an active contour is a simple curve by definition and maintains this topology throughout the numerical procedure.

- 2) The constant speed curve is unique among all possible parameterizations of a curve. In addition to providing a unique solution, the constant speed parameterization is highly desirable in brain mapping and registration applications since it preserves relative distances. This simplifies the calculation of intrinsic metric properties of the cortex such as length and curvature.
- 3) The spine does not lie on a boundary between two objects. This means that snakes are not directly applicable to this problem since their external forces seek a boundary.

B. A New External Force

Consider the point \mathbf{x} with circular neighborhood $\mathcal{N}(\mathbf{x})$ of radius ρ as shown in Fig. 1. The intersection of the neighborhood with the ribbon forms a crescent-shaped area with center of mass $\mathbf{c}(\mathbf{x})$. We define the external force acting on \mathbf{x} to be the vector $\mathbf{c}(\mathbf{x}) - \mathbf{x}$. This force draws \mathbf{x} toward the ribbon. If the neighborhood is completely within the ribbon, then the force is zero and the point is not compelled to move. In this case, if the diameter of the neighborhood is equal to the thickness of the ribbon, then \mathbf{x} is, by definition, on the ribbon's skeleton. In this way, our new external force attracts points to the skeleton of a ribbon.

There are several points to consider in this definition of external force. First, in order that there is a center of mass, there must first be a mass. It is implied above that a ribbon has constant mass on the ribbon and zero mass off the ribbon. In this case, it is sufficient to define the *mass function* as just the indicator function of a ribbon. This, in turn, implies that a brain image is first segmented into cortex (the ribbon) and noncortex regions. It is certainly possible to define a ribbon in this fashion and to run our algorithm on this type of presegmented data. But, is also possible to run our algorithm directly on the raw image by defining a different mass function. For example, in our magnetic resonance experiments (see Section V), we define the mass function at each pixel as the posterior probability that pixel belongs to gray matter. In the ideal case of perfect observations, this reverts to the indicator function; but with noisy observations this defines a kind of “soft” mass function having values between zero and one. In either case, the center of mass is well defined and provides a valid definition of external force.

A second consideration involves the choice of neighborhood radius ρ . Ideally, to have the external force become zero only on the spine, we require either a variable radius which conforms to the local ribbon thickness or a constant-thickness ribbon with thickness $w = 2\rho$. It is also important to note that the center of mass is undefined if the neighborhood

does not intersect the ribbon. For our theoretical analysis in Section IV, we assume that the ribbon has constant width w , the neighborhood radius is $\rho = w/2$, and that the active contour is close enough to the ribbon so that the centers of mass are defined. For our experiments in Section V, however, we use a multiresolution approach wherein the neighborhoods start out large and then shrink as the active contour approaches the ribbon. The final neighborhood size is constant, an approximation to the average width of the ribbon.

A final point to consider is the way in which the external force on an active contour point changes as the point moves toward the ribbon. Most importantly, the center of mass moves as the point moves; thus, the point of attraction of a given point is always changing. This is really just a consequence of the fact that the collection of forces defined on all the points in the plane comprise a vector field defined in *spatial* coordinates rather than material coordinates. In this way, our external forces are analogous to the forces implied by the gradient field of filtered images used in [12].

Kass *et al.* [12] also introduced the notion of control points, in which stationary points within the image attract particular points on the active contour. This defines a *material* field, which is handled in quite a different analytical and computational fashion. In Section V, we introduce a new approach to control points that augments the spatial forces rather than the material forces. This approach shows improved mapping of the sulci and opens up a new area for future research.

III. ACTIVE CONTOUR ALGORITHM FOR RIBBONS (ACAR)

Active contour algorithms are traditionally developed by formulating an energy function and then using calculus of variations to find the Euler equations representing its equilibrium force balance condition [12]. Here, we work backwards, finding a force balance condition and then deducing the required energy function. Although the force balance condition is all that is required for implementation, the variational framework provides physical intuition, allows us to select fundamental constants, and permits us to contrast this approach with that of snakes.

A. Force Balance Condition

We start by modeling an active contour as a collection of points $\mathbf{x}_i = [x_i, y_i]^T$, $i = 0, \dots, N$, where $x_i = x(i/N)$ and $y_i = y(i/N)$. The positions of the endpoints are assumed to be fixed and known, i.e., Dirichlet boundary conditions, and given by $\mathbf{x}_0 = (\alpha, \gamma)$ and $\mathbf{x}_N = (\beta, \delta)$. Now consider a free (nonboundary) point \mathbf{x}_i on the active contour. If the sole force exerted on this point were the external force directing it toward its center of mass $\mathbf{c}_i = \mathbf{c}(\mathbf{x}_i)$, then it would move directly toward the ribbon independently of the other points on the active contour. But what of the connectivity of curve and the desired isometry of the result? To encourage these two properties, spring forces acting between consecutive points on the active contour are introduced. Two terms per free active contour point, these forces act in the direction of the neighbors with strength proportional both to the neighbor's distance and

to a spring constant K . These forces comprise the internal forces of the active contour.

Given the definitions of both the external and internal forces, we define the solution to ACAR by the requirement that each free point on the active contour obeys the force balance condition

$$[\mathbf{c}_i - \mathbf{x}_i] + [K(\mathbf{x}_{i+1} - \mathbf{x}_i)] + [K(\mathbf{x}_{i-1} - \mathbf{x}_i)] = 0, \quad i = 1, \dots, N-1. \quad (1)$$

It is convenient to restate the collection of force balance conditions in matrix form. The free points of a discrete curve are represented by the vector

$$\mathbf{d} = [x_1, y_1, \dots, x_{N-1}, y_{N-1}]^T = [\mathbf{x}_1^T, \dots, \mathbf{x}_{N-1}^T]^T, \quad (2)$$

and their centers of mass by

$$\mathbf{f}(\mathbf{d}) = [\mathbf{c}_1^T, \dots, \mathbf{c}_{N-1}^T]^T. \quad (3)$$

Then, (1) for $i = 1, \dots, N-1$ can be written as

$$\mathbf{d} - \mathbf{f}(\mathbf{d}) + K\mathbf{A}\mathbf{d} = K\mathbf{b}, \quad (4)$$

where

$$\mathbf{b} = [\alpha, \beta, 0, \dots, 0, \gamma, \delta]^T, \quad (5)$$

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & -1 & 0 & \cdots & 0 \\ 0 & 2 & 0 & -1 & \ddots & \vdots \\ -1 & 0 & 2 & 0 & \ddots & 0 \\ 0 & -1 & 0 & 2 & \ddots & -1 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 0 & 2 \end{bmatrix}. \quad (6)$$

Because \mathbf{f} depends on \mathbf{d} , this equation is nonlinear in \mathbf{d} and cannot be solved by standard large-scale matrix inversion techniques.

B. Modified Gauss-Seidel (MGS) Solution

The numerical procedure we use to solve (4) is a modification of the standard Gauss-Seidel approach for solving large linear systems. While certainly not the only approach for solving this nonlinear equation, it has both intuitive and theoretical appeal which should become clear during the following discourse.

Consider an iterative approach in which $\mathbf{f}(\mathbf{d})$ is temporarily fixed for a given \mathbf{d} . In this case, the force-balance equation becomes

$$\mathbf{d} - \mathbf{f} + K\mathbf{A}\mathbf{d} = K\mathbf{b}, \quad (7)$$

which is linear in the unknown \mathbf{d} . While it is now possible to solve for \mathbf{d} in a single step, it is also important to recognize that (7) represents a crude linearization of the true force-balance equation and is only valid for \mathbf{d} near the one for which \mathbf{f} was fixed. Therefore, upon fixing \mathbf{f} , our approach takes only a single Gauss-Seidel iteration toward the solution to (7). At the k -th iteration, this step is given by (cf., [35])

$$\mathbf{x}_i^{k+1} = (2K + 1)^{-1}[\mathbf{c}_i^k + K(\mathbf{x}_{i+1}^k + \mathbf{x}_{i-1}^{k+1})], \quad (8)$$

for $i = 1, \dots, N-1$. After all the points are so updated, the center of mass function is resampled, i.e., \mathbf{f} is recalculated given the new active contour. This two-step procedure, which we call *modified Gauss-Seidel*, is repeated until convergence.

C. A Variational Framework for ACAR

Physically, a force-balance equation is a necessary condition that minimizes an energy [36]. It is both interesting and useful to determine the variational principle that gives rise to ACAR. For example, it is interesting to see how the external energy term differs from that of snakes. It is useful to find the fundamental regularization parameter K_0 which underlies the problem and to examine the conditions under which the problem is convex.

It can be shown that MGS is a so-called *explicit step* (cf., [37]) in the solution of the following pair of nonlinear partial differential equations:

$$x(s) - u[x(s), y(s)] - K_0 x(s) = 0 \quad (9a)$$

$$y(s) - v[x(s), y(s)] - K_0 y(s) = 0 \quad (9b)$$

where $K = K_0 N^2$ and $[u(x(s), y(s)), v(x(s), y(s))]$ is the center of mass in the neighborhood of the point $[x(s), y(s)]$. We can therefore view MGS as the solution of a variational problem for which (9) are the Euler equations. The constant K_0 is a fundamental (regularization) constant in this formulation, independent of N , the number of points used to represent the continuous active contour.

Euler equations arise as the *variational derivative* of an energy function \mathcal{E} which depends on the functions $x(s)$ and $y(s)$ used to represent a continuous active contour (cf., [38]). It can be shown that (9a) and (9b) are Euler equations if

$$\mathcal{E} = \mathcal{E}_F + \mathcal{E}_E \quad (10)$$

where

$$\mathcal{E}_F = 2 \int_0^1 P[x(s), y(s)] ds \quad (11a)$$

$$\mathcal{E}_E = K_0 \int_0^1 \left\{ \left(\frac{dx(s)}{ds} \right)^2 + \left(\frac{dy(s)}{ds} \right)^2 \right\} ds, \quad (11b)$$

and

$$\frac{\partial P(x, y)}{\partial x} = x - u(x, y), \quad (12a)$$

$$\frac{\partial P(x, y)}{\partial y} = y - v(x, y). \quad (12b)$$

The first term \mathcal{E}_F is an energy term corresponding to the external forces, and the second term \mathcal{E}_E is an energy term corresponding to the internal elastic forces.

From (12), we see that $P(x, y)$ is a potential function whose negative gradient defines our external force field. But it is not always possible to define $P(x, y)$ given arbitrary data $u(x, y)$ and $v(x, y)$. This is true since $\partial^2 P(x, y)/\partial x \partial y = \partial^2 P(x, y)/\partial y \partial x$ must be true, which implies that (12) is valid only if

$$u_y(x, y) = v_x(x, y). \quad (13)$$

Another way to state this problem is to note that any vector field which is the gradient of a potential function must be irrotational, which is not true in general. This difficulty is strictly of theoretical concern since ACAR is defined by its Euler equations, not by a variational formulation. We have, however, determined that the field is approximately irrotational—and therefore that this variational framework is accurate—in two circumstances: 1) when the active contour is very close to the spine and the neighborhood diameter is approximately equal to the width of the ribbon; and 2) when the ribbon has no sharp bends and its intersections with a circular neighborhood are circular arcs. In our study of convexity in Section V, we implicitly assume that one of these conditions holds.

Although P does not need to be calculated to solve ACAR, it is instructive in our simulations to determine the final energy E of an active contour. (We note that since the final active contour is near the spine in this case, the field is nearly irrotational and the variational framework is accurate.) From (12), we can deduce that the total differential of P is

$$dP(x, y) = [x - u(x, y)] dx + [y - v(x, y)] dy.$$

Then, from vector integral calculus (cf., [39]), it follows that

$$P(x, y) = P(x_0, y_0) + \int_{\mathcal{C}} [g - u(g, h)] dg + \int_{\mathcal{C}} [h - v(g, h)] dh, \quad (14)$$

where \mathcal{C} is any path originating at $\mathbf{x}_0 = (x_0, y_0)$ and ending at $\mathbf{x} = (x, y)$. Thus, P is completely determined by knowledge of P at a single point \mathbf{x}_0 and of u and v along any path connecting \mathbf{x}_0 to \mathbf{x} . Suppose we choose $\mathbf{x}_0 = \alpha(0)$ and set $P(x_0, y_0) = 0$; then, it is easy to deduce from (14) that P is zero on the entire spine. Therefore, $P(x, y)$ is determined by any path leading from any point on the spine to \mathbf{x} .

A particularly useful path that takes \mathbf{x} to a point on the spine is given by the so-called *line of flow* of the vector force field $\mathbf{F} = \mathbf{c}(\mathbf{x}) - \mathbf{x}$ originating from \mathbf{x} . This is the path that a particle at \mathbf{x} would follow if \mathbf{F} were a time-invariant velocity field. A parametric expression for the line of flow originating from \mathbf{x} can be found by solving

$$\frac{dg(t)}{dt} = F_x[g(t), h(t)], \quad (15a)$$

$$\frac{dh(t)}{dt} = F_y[g(t), h(t)], \quad (15b)$$

for $g(t)$ and $h(t)$ subject to the boundary condition $[g(0), h(0)] = (x, y)$. Thus, $P(x, y)$ can be determined using (15) to find \mathcal{L} and (14) to integrate the field backwards over this line of flow. A numerical procedure can be readily developed to combine these operations into a single iterative procedure that is applied to each point (x, y) .

IV. UNIQUENESS AND FIDELITY

The selection of the regularization constant K_0 , essentially a spring constant, is critical to the overall performance of ACAR. If chosen too large, the active contour may not be

able to conform to the ribbon, and fidelity is lost. If chosen too small, the points become separated, and both isometry and uniqueness are lost. In this section, we analyze both uniqueness and fidelity of ACAR as a function of K_0 , and provide a guideline for its selection.

A. Convexity

In this section, we consider the convexity of $\tilde{\mathcal{E}}$, a discrete approximation to \mathcal{E} , noting that if $\tilde{\mathcal{E}}$ is convex, then it has a unique global minimum. In the common least squares problem of estimating x under the model $Ax = b$, convexity is guaranteed regardless of the data b . In contrast, since the Euler equation (4) is nonlinear, the convexity of \mathcal{E} depends on the data. Therefore, in order to discuss convexity, it is necessary to make assumptions about the nature of the true object. Our guidelines were to make the least restrictive yet workable assumptions, and under these assumptions, our main theoretical result of this section is that \mathcal{E} is guaranteed to be convex if $K_0 \geq K_{LB}$, where K_{LB} is determined by the ribbon and is defined precisely below. The inevitable deviations from our assumptions make this lower bound approximate in practice. However, our experience has shown very good agreement between theory and practice, even under nonideal circumstances. Experimental evidence supporting this observation is given in Section V.

Here are our assumptions. First, we assume that the ribbon has constant thickness w and that its indicator function is known and used for center of mass calculations. Second, we assume that each neighborhood intersects a single boundary curve of the ribbon. Third, we assume that the boundary curve intersecting a given neighborhood is exactly a circular arc within the neighborhood.

Using finite differences to approximate the required derivatives, a discrete approximation of \mathcal{E} is given by $\tilde{\mathcal{E}}(\mathbf{d}) = \tilde{\mathcal{E}}_F(\mathbf{d}) + \tilde{\mathcal{E}}_E(\mathbf{d})$ where

$$\tilde{\mathcal{E}}_F(\mathbf{d}) = \frac{2}{N} \sum_{i=0}^N P_i \quad (16a)$$

$$\begin{aligned} \tilde{\mathcal{E}}_E(\mathbf{d}) = K_0 N \sum_{i=1}^{N-2} [(x_{i+1} - x_i)^2 \\ + (y_{i+1} - y_i)^2] \\ + K_0 N [(x_1 - \alpha)^2 + (x_{N-1} - \beta)^2 \\ + (y_1 - \gamma)^2 + (y_{N-1} - \delta)^2]. \end{aligned} \quad (16b)$$

In these equations and for the remainder of the paper, we adopt the convention that a superscript i applied to u , v , or any of their derivatives means that the function is to be evaluated at (x_i, y_i) , e.g., $u^i = u(x_i, y_i)$. This is done so that we may use the subscripts x and y to denote partial derivatives.

For $\tilde{\mathcal{E}}(\mathbf{d})$ to be strictly convex at \mathbf{d} , the Hessian matrix \mathbf{H} of $\tilde{\mathcal{E}}(\mathbf{d})$ must be positive definite at \mathbf{d} . From (16), we see that

$$\mathbf{H} = \mathbf{H}_F + \mathbf{H}_E,$$

and from standard matrix analysis [40], it follows that

$$\lambda_{\min}(\mathbf{H}) \geq \lambda_{\min}(\mathbf{H}_F) + \lambda_{\min}(\mathbf{H}_E). \quad (17)$$

Since \mathbf{H} is positive definite if its minimum eigenvalue is greater than zero, a sufficient condition for the strict convexity of $\tilde{\mathcal{E}}(\mathbf{d})$ is given by

$$\lambda_{\min}(\mathbf{H}_F) + \lambda_{\min}(\mathbf{H}_E) > 0. \quad (18)$$

Finding expressions for $\lambda_{\min}(\mathbf{H}_E)$ and $\lambda_{\min}(\mathbf{H}_F)$ and solving (18) for the regularization coefficient K_0 provides our convexity condition.

From (16a), we see that

$$\mathbf{H}_E = 2K_0 N \mathbf{A}', \quad (19)$$

which has a minimum eigenvalue of $\lambda_{\min}(\mathbf{H}_E) = 2K_0 N \lambda_{\min}(\mathbf{A}')$. To find $\lambda_{\min}(\mathbf{A}')$, we note that we can write the free variables of an active contour using the vector $\mathbf{d}' = [x_1, \dots, x_{N-1}, y_1, \dots, y_{N-1}]^T$. With this ordering, the Hessian becomes $\mathbf{H}'_E = 2K_0 N \mathbf{A}'$, where

$$\mathbf{A}' = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}, \quad (20)$$

and \mathbf{B} is a symmetric Toeplitz tridiagonal matrix whose first row is $[2, -1, 0, \dots, 0]$. It is readily demonstrated that the eigenvalues of \mathbf{A}' coincide with those of \mathbf{B} , and that those of \mathbf{A}' coincide with those of \mathbf{A} . Also, since \mathbf{B} is tridiagonal and Toeplitz, its eigenvalues can be calculated through a recurrence formula [41]. Putting these facts together yields

$$\lambda_{\min}(\mathbf{A}') = 2 - 2 \cos(\pi/N), \quad (21)$$

and therefore

$$\lambda_{\min}(\mathbf{H}_E) = 2K_0 N (2 - 2 \cos(\pi/N)). \quad (22)$$

Turning now to \mathbf{H}_F , we note that since

$$\frac{\partial P_i}{\partial x_j} = \frac{\partial P_i}{\partial y_j} = 0 \quad i \neq j,$$

it follows that \mathbf{H}_F is block diagonal with blocks

$$D_i = \begin{bmatrix} \sigma_i & \gamma_i \\ \gamma_i & \psi_i \end{bmatrix}, \quad i = 1, \dots, N-1.$$

The elements of D_i are given by

$$\begin{aligned} \sigma_i &= 2 \frac{\partial(x_i - u^i)}{\partial x_i} = 2(1 - u_x^i), \\ \gamma_i &= 2 \frac{\partial(x_i - u^i)}{\partial y_i} = 2 \frac{\partial(y_i - v^i)}{\partial x_i} = -2u_y^i = -2v_x^i, \\ \psi_i &= 2 \frac{\partial(y_i - v^i)}{\partial y_i} = 2(1 - v_y^i). \end{aligned}$$

The eigenvalues of \mathbf{H}_F can be determined by finding the eigenvalues λ_{i1} and λ_{i2} of the $N-1$ matrices D_i by solving

$$\lambda^2 - (\sigma_i + \psi_i)\lambda + \sigma_i\psi_i - (\gamma_i)^2 = 0, \quad i = 1, \dots, N-1.$$

Because σ_i , γ_i , and ψ_i depend on the gradient of \mathbf{c}_i , the local shape of the ribbon must be considered. Here, we use the assumptions that a neighborhood intersects only one boundary curve of the ribbon and that within this intersection the boundary is exactly a section of a circular arc. Consider

the local geometry in which the center of this circular arc is the origin, as shown in Fig. 2. It follows that the spine is also a circular arc with a fixed radius, say R , and that the point \mathbf{x}_i and its center of mass \mathbf{c}_i lie on the same radial line. As is shown in Fig. 2, the polar coordinates of \mathbf{x}_i and \mathbf{c}_i in this coordinate system are (r_i, ϕ_i) and $(r_i + n(r_i), \phi_i)$, respectively, where $n(\cdot)$ is a function depending only on the radial coordinate. To simplify our notation, we refer to $n(r_i)$ as n^i , and use n_r^i and n_{rr}^i to denote its first and second derivatives with respect to r_i . Using this approximation and notation, it can be shown after some algebra that

$$\lambda_{i1} = \frac{2}{N} (-n_r^i), \quad (23a)$$

$$\lambda_{i2} = \frac{2}{N} \left(-\frac{n^i}{r_i} \right), \quad (23b)$$

for $i = 1, \dots, N-1$. It follows that the minimum eigenvalue of \mathbf{H}_F is

$$\lambda_{\min}(\mathbf{H}_F) = \min \{ \lambda_{i1}, \lambda_{i2}, \quad i = 1, \dots, N-1 \}. \quad (24)$$

Combining (24) and (22) with (18), we have that $\tilde{\mathcal{E}}$ is convex if

$$\frac{2}{N} \min_i \{ -n_r^i, -n^i/r_i \} + 2K_0N[2 - 2\cos(\pi/N)] > 0.$$

Therefore, $\tilde{\mathcal{E}}(\mathbf{d})$ is convex if $K_0 > K_{LB}$ where

$$K_{LB} = \frac{-\min_i \{ -n_r^i, -n^i/r_i \}}{N^2(2 - 2\cos(\pi/N))}. \quad (25)$$

The following remarks are made to clarify and interpret this result:

- **Sharp Lower Bound:** If $K_0 > K_{LB}$, then from the above argument $\tilde{\mathcal{E}}(\mathbf{d})$ is guaranteed to be convex. But it is also possible that K_0 chosen smaller than the bound will yield a convex problem. Thus, our convexity condition is sufficient, but not necessary. In practice, the sharpness of this bound depends on the ribbon.
- **Numerator of Lower Bound:** The radial function $n(r_i)$ is implicitly a function of both R and ρ , as well as r_i , but in fact may be considered to be a function of just the ratios r_i/ρ and R/ρ . Using formulas derived in [42], in Fig. 3 we plot $-n_r^i$ and $-n^i/r$ as functions of r_i/ρ for different values of R/ρ , assuming that $\rho \leq w/2$. We conclude that λ_{i1} is always positive and that λ_{i2} can be either positive or negative. Since K_0 is greater than zero by definition, from (25) we see that positive eigenvalues pose no constraint whatsoever. Therefore, K_0 is constrained only when λ_{i2} is negative. From Fig. 3, we see that this occurs only when the ribbon boundary curves toward the neighborhood (as in Fig. 2(b)) and that λ_{i2} goes increasingly negative with increasing distance between the spine and the neighborhood. Also, the slope is largest when the ribbon curvature is largest. We therefore conclude that K_{LB} is determined at the ribbon's strongest curve by the most distant inside neighborhood.
- **Denominator of the Lower Bounds:** From the denominator of (25) it appears at first glance that the lower bound

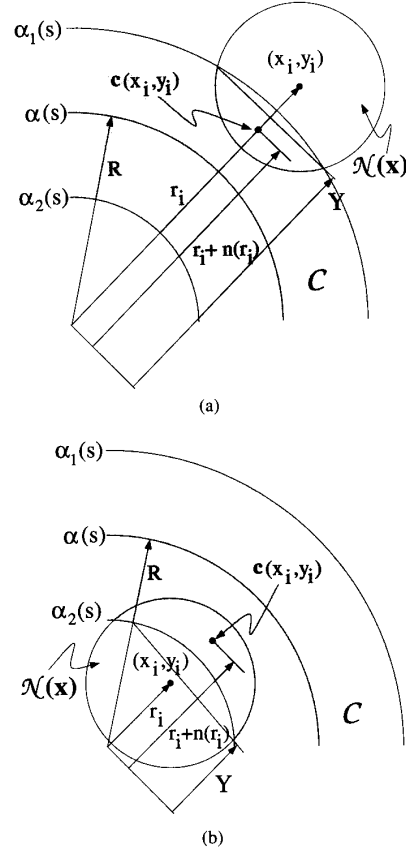


Fig. 2. Two possible geometries for the intersection of a small neighborhood with an annulus.

can be made arbitrarily small by increasing N . This is not true, however, as it can be shown using a series expansion of the cosine function that $\lim_{N \rightarrow \infty} N^2[2 - 2\cos(\pi/N)] = \pi^2$. In fact, this sequence converges quite rapidly so that for typical N we can approximate K_{LB} in practice by

$$K_{LB} \approx -\pi^{-2} \min_{1 \leq i \leq N-1} \{ -n_r^i, -n^i/r_i \}. \quad (26)$$

- **Nonconvex Domain:** We have shown that the energy function $\tilde{\mathcal{E}}(\mathbf{d})$ is convex when $K_0 > K_{LB}$. Since $\tilde{\mathcal{E}}(\mathbf{d})$ is only defined when the neighborhoods intersect the ribbon, however, this restricts \mathbf{d} to a domain \mathcal{D} , and \mathcal{D} is generally not convex. Therefore, despite the convexity of the energy function, a descent algorithm (of which ACAR is one) will generally “get stuck” at the boundary and not reach the global minimum. To get around this problem, we have implemented two approaches based on temporarily redefining the energy function itself. The first approach allows the active contour to go outside of \mathcal{D} by redefining the external forces so that a point whose neighborhood does not intersect the ribbon simply feels no external force. The second approach redefines \mathcal{D} by increasing the neighborhood size so the external forces are always defined. In either case, we expect the active

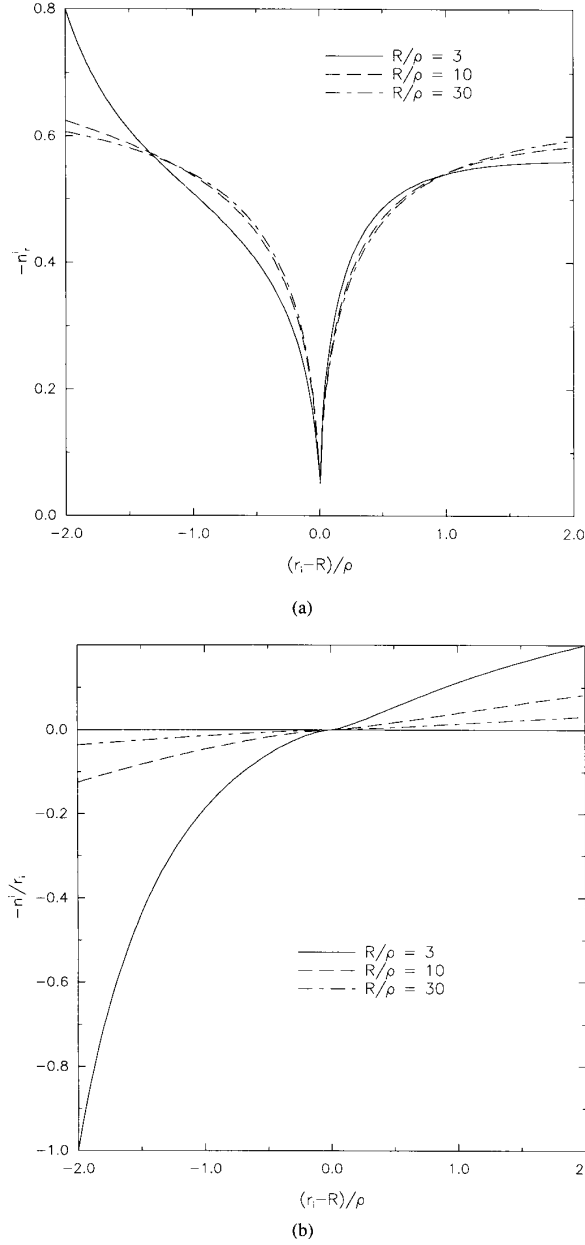


Fig. 3. Plots showing (a) $(N/2) \lambda_{i1}$ and (b) $(N/2) \lambda_{i2}$.

contour to continue to move closer to the ribbon so that these changes will be temporary. Through this temporary action, the active contour is allowed to relocate itself into a different region of \mathcal{D} , which may be closer to the global minimum.

B. Frequency Domain Analysis

The convexity result above shows how to choose K_0 to assure a unique solution. But how close will this solution be to the truth? If K_0 is chosen to be very large then the problem is convex, but in this case the active contour acts like

a very tight rubber band, resisting all effort to stretch it to conform to the ribbon. Clearly, if K_0 is too large then fidelity is lost, meaning that the solution is too smooth and the active contour points do not sit on the spine. In this section, we study ACAR in the frequency domain and show that it can be interpreted as a sequence of lowpass filtering operations. The effective cutoff frequency is determined in part by K_0 , where a large K_0 dictates a small cutoff frequency. Thus, if K_0 is too large, the cutoff frequency may be lower than the highest frequencies in the ribbon, precluding the possibility of accurate reconstruction. This principle allows us to determine an upper bound K_{UB} on K_0 below which fidelity is maintained.

Recall that the numerical solution to ACAR involves a two-step procedure, MGS, which linearizes the nonlinear Euler equation, takes one step toward the solution of the resulting linear equation, resamples the center of mass function, and repeats. The linearized Euler equation (7) has the following continuous equivalent:

$$x(s) - u(s) - K_0 x(s) = 0, \quad (27a)$$

$$y(s) - v(s) - K_0 y(s) = 0. \quad (27b)$$

Using the following Fourier series representations for $x(s)$ and $y(s)$:

$$x(s) = X_0 + \sum_{l=1}^{\infty} \{X_{cl} \cos(2\pi ls) + X_{sl} \sin(2\pi ls)\} \quad (28a)$$

$$y(s) = Y_0 + \sum_{l=1}^{\infty} \{Y_{cl} \cos(2\pi ls) + Y_{sl} \sin(2\pi ls)\}, \quad (28b)$$

and analogous representations for $u(s)$ and $v(s)$, it is readily shown that

$$X_{cl} = H(l) U_{cl}, \quad l = 0, 1, \dots, \infty \quad (29a)$$

$$X_{sl} = H(l) U_{sl}, \quad l = 0, 1, \dots, \infty \quad (29b)$$

$$Y_{cl} = H(l) V_{cl}, \quad l = 0, 1, \dots, \infty \quad (29c)$$

$$Y_{sl} = H(l) V_{sl}, \quad l = 0, 1, \dots, \infty \quad (29d)$$

where

$$H(l) = \frac{1}{1 + K_0 4\pi^2 l^2}. \quad (30)$$

These equations describe a frequency filter with transfer function $H(l)$, a first-order discrete Butterworth low-pass filter, applied to $u(s)$ and $v(s)$. Therefore, we see that MGS performs one step toward the implementation of a Butterworth filter. The new curve will be similar to the current center of mass function, but will have its frequency content more concentrated in the low frequencies. A similar filtering interpretation of the snake active contour formulation was derived by Whitten in [43]. However, in [43], this frequency interpretation was not related to the selection of K_0 , as we now do.

Assume that the active contour is initialized exactly on the spine, $\alpha(s)$. If $\alpha(s)$ has spectral energy outside the pass-band of $H(l)$, one step of ACAR will filter some of these frequencies out, causing the active contour to move away from $\alpha(s)$. Resampling the center of mass will offset this loss to some extent, but subsequent iterations will continue to move

the active contour away from the spine. If, on the other hand, the spectrum of $\alpha(s)$ falls within the pass-band of $H(l)$, then an MGS iteration will leave the active contour unchanged, and $\alpha(s)$ will be the solution of ACAR.

Because $H(l)$ is not a perfect low-pass filter and because only one step toward its solution is made per iteration, the preceding argument is only approximate. In fact, since the filter is applied repeatedly, the effective cutoff frequency of the cumulative filter is far lower than that of a standard Butterworth filter. We define this effective cutoff frequency l_c in terms of a parameter $\mu \ll 1$ as follows:

$$H(l_c) = (1 - \mu)H(0). \quad (31)$$

From (31) and (30), and using the fact that $\mu \ll 1$, we obtain

$$l_c = \sqrt{\frac{\mu}{1 - \mu}} \frac{1}{2\pi\sqrt{K_0}} \approx \frac{\sqrt{\mu}}{2\pi\sqrt{K_0}}. \quad (32)$$

If the spectrum of $\alpha(s)$ falls largely below this effective cutoff frequency, then ACAR will maintain fidelity. Although the bandwidth of $\alpha(s)$ may be known or assumed, because it is the spine of a ribbon it also has a maximum possible bandwidth. In particular, it is not difficult to see using a simple geometric argument that spine of a ribbon with minimum thickness w_{\min} cannot have a radius of curvature smaller than $w_{\min}/2$. Therefore, the curvature, which is the reciprocal of radius of curvature, is bounded above by

$$\kappa_o = \frac{2}{w_{\min}}. \quad (33)$$

Now suppose that only one frequency has energy. Then ignoring possible overlap, translation, and the initial starting point, the coordinate functions must be of the form

$$x(s) = \frac{1}{\kappa} \cos(L\kappa s) \quad (34a)$$

$$y(s) = \frac{1}{\kappa} \sin(L\kappa s), \quad (34b)$$

where L is the total length of the spine. Since $\kappa \leq \kappa_o$, the maximum frequency at which these functions can exist is $L\kappa_o/2\pi$. Therefore, we conclude that a practical constraint on the bandlimit of $\alpha(s)$ is provided by

$$l_o \leq \frac{L\kappa_o}{2\pi}. \quad (35)$$

Substituting (33) into (35) yields

$$l_o \leq \frac{L}{\pi w_{\min}}. \quad (36)$$

A more accurate bound that allows the possibility of energy in all Fourier coefficients is developed in [44].

We are now in position to specify K_{UB} . To make sure that ACAR will not filter out the frequencies present in $\alpha(s)$, we must have $l_o < l_c$. Using (32) and (36), we find that K_0 should be selected to satisfy

$$K_0 < K_{UB}, \quad (37)$$

where

$$K_{UB} = \frac{\mu w_{\min}^2}{4L^2}. \quad (38)$$

If the maximum curvature κ_{\max} is known to be less than the upper bound determined by the width of the ribbon, a larger upper bound can be found by using (32) and (35), yielding

$$K_{UB} = \frac{\mu}{L^2 \kappa_{\max}^2}. \quad (39)$$

From the above result, we see that, in order to maintain fidelity, K_0 should be small. We know from the previous section that if K_0 is too small convexity will be lost and the result will be a local minimum. Interestingly, it turns out that this phenomenon can also be interpreted in the frequency domain as aliasing. To show how aliasing occurs when K_0 is too small, we consider what sampling frequencies are required to sample the spine, the active contour, and the center of mass function. Nyquist's theorem specifies that in order to adequately sample $\alpha(s)$, the sampling frequency l_s must be at least $2l_o$. At some point during the iterative process, however, the active contour itself may contain higher frequencies than the spine, so we must, in general, choose a sampling frequency higher than that required by the spine. If the filter imposed by ACAR were ideal, then a sampling frequency of $l_s = 2l_c$ adequately samples the active contour at the output of this filter. But even this sampling frequency may not be high enough because the center of mass function of this same active contour may have even higher frequencies present. We note that if aliasing of the resampled center of mass function occurs—and this depends on N , K_0 , and the ribbon itself—high frequencies will be mapped into low frequencies and the effect will never be filtered out. Therefore, aliasing causes convergence to a local minimum.

C. Selecting the Regularization Constant

From the previous two sections, we know that convexity is maintained if $K_0 > K_{LB}$ and fidelity is maintained if $K_0 < K_{UB}$. Clearly, the ideal circumstance would have $K_{LB} < K_0 < K_{UB}$. But we have found that K_{LB} is less than K_{UB} only when the ribbon turns very gently. Thus, in practice, we cannot expect to ever satisfy both conditions simultaneously, and a trade-off must be made. In this case, we can take $K_{UB} < K_0 < K_{LB}$. If K_0 is selected near K_{UB} then the active contour points will move very near to the spine, but might be separated and isometry might be lost. If K_0 is selected near K_{LB} then the points will be equally spaced but the active contour will be too tight, an overly smooth version of the spine. We note that since the lower bound is a sufficient condition, it is possible to choose K_0 lower than K_{LB} and still have a convex problem. However, if $K_0 > K_{UB}$, then oversmoothing is guaranteed to occur. Thus, if $K_{LB} < K_{UB}$ then K_0 should be chosen close to K_{UB} rather than close to K_{LB} .

V. EXPERIMENTAL RESULTS

In this section, we present a set of experiments with both synthetic and real data. First, we experimentally confirm our theoretical lower and upper bounds for ACAR. We then demonstrate its overall performance for different choices of K_0 and N , showing the effects of oversmoothing and aliasing.



Fig. 4. An annulus of thickness $w = 6$ and radius $R = 60$ used in the experiment of Section V-A.

We then adapt the mass function to gray scale images and show how ACAR is robust to noise. Finally, we apply ACAR to magnetic resonance brain images, demonstrating its use in mapping of the outer cortex.

A. Confirmation of Bounds

In this section, we examine the performance of ACAR on a very simple ribbon, an annulus created using a semicircular spine with radius $R = 60$ dilated to obtain a ribbon with thickness $w = 6$, as shown in Fig. 4. (All physical dimensions in this section are in units of pixels). The number of points in the contour was $N = 50$ and the neighborhood radius was $\rho = 3$. The bounds for this ribbon are $K_{LB} = 1.7 \times 10^{-3}$ and $K_{UB} = 1.0 \times 10^{-3}$, assuming $\mu = 0.01$ (see (32)). Clearly, there is no K_0 that will simultaneously satisfy both bounds.

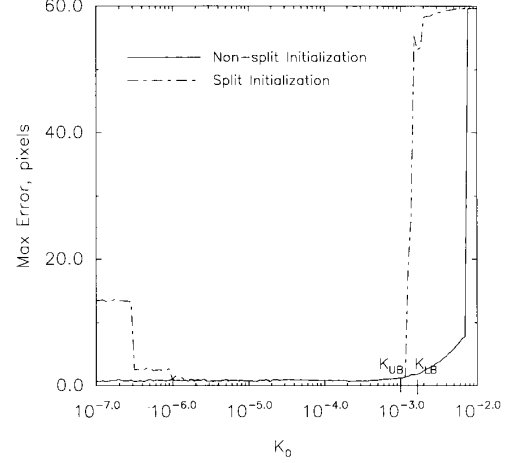
We ran a series of simulations using different values of K_0 for two different initial configurations: a nonsplit and a split configuration. The nonsplit configuration has its points almost uniformly spaced on a semicircle of radius 57, except that its endpoints are fixed to the truth. This is a curve that is very close to the true spine. The split configuration is also on a semicircle of radius 57 with endpoints fixed to the truth; but its points are split so that nearly half of the semicircle contains no points at all. Since this configuration is far from the spine, we expect the active contour to be trapped in local minima for small K_0 's. We note that this initial active contour is unlikely in practice and was chosen here for theoretical verification.

We use two measures of error to assess the performance of ACAR. The first measure is the fractional change in energy of the final active contour with respect to the energy of the true spine

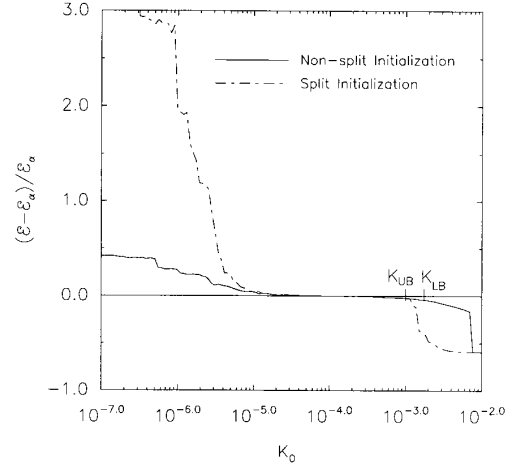
$$\mathcal{F} = (\mathcal{E} - \mathcal{E}_\alpha) / \mathcal{E}_\alpha.$$

We note that if this measure is positive, then we can conclude that the energy function is not convex; if it is negative, then we can conclude that $\alpha(s)$ is not the global minimum. Our second measure is the maximum distance in the radial direction between the spine and the final active contour. To make this computation, we assume that active contour points are connected by line segments and resampled so that no two points are separated by more than a prespecified small distance. This measure represents a worst-case performance since it focuses locally on the poorest fit between the final active contour and the spine.

Performance results are presented in Fig. 5, where we see three distinct regions corresponding to K_0 small, medium, and large. In the central range where $10^{-5} < K_0 < 10^{-3}$,



(a)



(b)

Fig. 5. (a) Plots of the maximum error between the spine and the ACAR solution as a function of K_0 for a nonsplit and a split initialization of the active contour. (b) Plots of the fractional change in energy \mathcal{F} as a function of the regularization constant K_0 for the same two initializations.

the maximum error is constant and small as is the fractional energy change \mathcal{F} . In this region, the active contour performs very well. As K_0 drops below 10^{-5} , \mathcal{F} begins to grow positive for both initializations. This indicates that the active contour is converging to local minima. However, the max error does not begin to grow until K_0 approaches 10^{-6} , and then only for the split initialization. This tells us that local minima exhibit themselves primarily as a loss of isometry, i.e., splits between the points begin to occur. The nonsplit configuration does not suffer this because of its very particular initialization.

When $K_0 > 10^{-3}$, several interesting phenomena take place. First, the split initialization curve makes an abrupt transition in both the maximum distance error and the fractional energy change plots. The nonsplit curve makes a similar transition at about 10^{-2} . This transition means that the active contour has left the domain \mathcal{D} in which the energy function is defined, that is, one or more neighborhoods no longer in-

intersect the ribbon. When this happens, our algorithm has been programmed to provide zero external force on the affected neighborhoods and zero external energy contribution to the overall energy. As explained in an earlier section, it is hoped that in doing this the active contour may drift back into \mathcal{D} and approach $\alpha(s)$ again. This does not happen for very large K_0 's, however. Instead, both curves ultimately collapse to a line segment joining the endpoints of the ribbon. The presence of these sharp transitions represent a catastrophic failure, a strong divergence from the truth, and a situation to be avoided at all costs.

The theoretical bounds are also indicated on the plots in Fig. 5. We first note that K_{UB} sits right at the point where catastrophic errors begin to take place. It is a very good upper bound on K_0 for this example. We next note that K_{LB} sits in the region of catastrophic failure. However, since this bound was derived as a sufficient condition, it is possible that the problem is also convex below this bound. In fact, from Fig. 5(b), we can conclude that the problem is definitely nonconvex when $K_0 < 10^{-4}$ (since \mathcal{F} is positive), but is possibly still convex above this value. For this example, any value between $0.1K_{UB}$ and K_{UB} is a good choice for K_0 .

B. Overall Performance

In this section, we examine the overall qualitative and quantitative performance of ACAR on a more realistic object. In this section, the ribbon, as shown in Figs. 6–8, is a modulated semicircle with nominal radius $R = 60$ and thickness $w = 6$, both in units of pixels. The formula for the spine is

$$x(s) = x_o - \{R + [r_o + a \cos(b\pi s)] \cdot \cos(2M\pi s)\} \cos \pi s \quad (40)$$

$$y(s) = y_o + \{R + [r_o + a \cos(b\pi s)] \cdot \cos(2M\pi s)\} \sin \pi s, \quad (41)$$

where, in this set of experiments, the parameters were $r_o = 7$, $a = 3$, $b = 2$, $M = 6$, $x_o = 125$, and $y_o = 50$. The initial curve is a constant velocity semicircle with the true end-points, as shown in Fig. 6(a). The neighborhood size is not large enough to guarantee intersection with the ribbon for all points in the curve. Therefore, we use a larger neighborhood size in the initial iterations and lower this size incrementally until all neighborhoods intersect with radius $\rho = 3$. Thereafter, the neighborhood size remains constant.

1) *Qualitative Assessment:* Fig. 6(b) shows the ACAR solution for $K_0 = 8 \times 10^{-7}$ and $N = 230$, which together represent empirically “good” parameter choices for this ribbon. Fig. 7 presents the results of K_0 too large and too small for the same N . In Fig. 7(a), we show the solution when $K_0 = 1 \times 10^{-4}$, which corresponds to an overregularized ACAR. The oversmoothing is apparent, especially near the sharp foldings of the boundary. As explained in Section IV, this oversmoothing occurs because high frequencies, which were shown to correspond to high curvatures, are filtered out. Fig. 7(b) shows the solution when $K_0 = 3 \times 10^{-8}$, which specifies an underregularized problem. Because the energy function is not convex in this case, the active contour is trapped in a local minimum. We have found through numerous

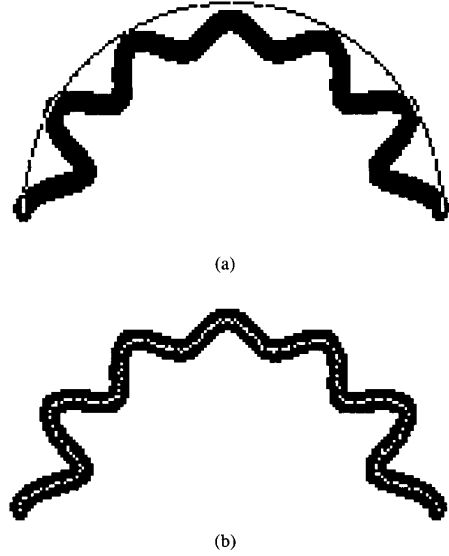


Fig. 6. (a) The initial configuration of the active contour superimposed on the synthetic ribbon. (b) Final curve for $K_0 = 8 \times 10^{-7}$.

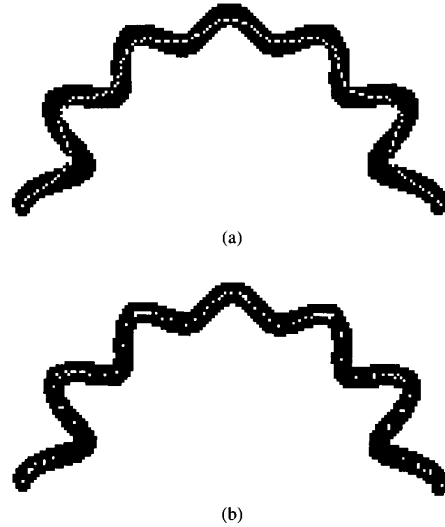


Fig. 7. (a) The final curve obtained through ACAR for $K_0 = 1 \times 10^{-4}$. This curve corresponds to an overregularized problem. (b) The final curve obtained through ACAR for $K_0 = 3 \times 10^{-8}$. This curve corresponds to an underregularized problem.

simulations that local minima correspond to curves that contain splits, i.e., points that are separated by large gaps. We can clearly see this kind of behavior in Fig. 7(b), especially close to the sharp foldings of the ribbon. One practical solution to this problem would be to redistribute the points near each split. This new curve will have higher energy, but resampling the center of mass function from this new position will allow the curve to move toward the desired solution, thus bypassing the local minimum.

In Fig. 8, we show the effect of undersampling the active contour. Using $K_0 = 4 \times 10^{-6}$ for both examples, the result of sampling the active contour with $N = 100$ points

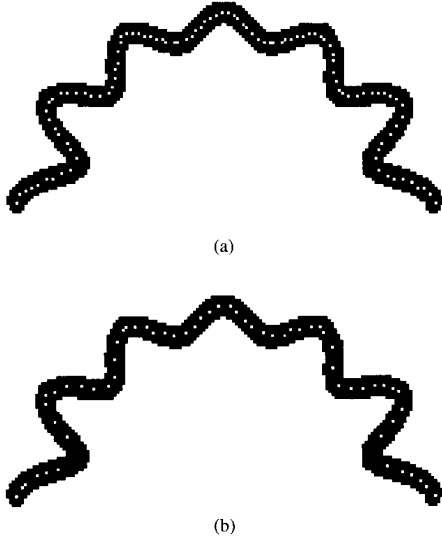


Fig. 8. The problem of undersampling and aliasing: K_0 is the same for both (a) and (b), but the active contour was sampled with 100 points in (a) and with 60 points in (b).

is shown in Fig. 8(a), and the result when $N = 60$ is shown in Fig. 8(b). The curve in Fig. 8(a) is a reasonably good result with some evidence of a problem at the two sharp folds nearest the boundary points. In Fig. 8(b), however, the final curve contains two strong splits at these sharp folds. The reason these splits did not occur in Fig. 8(a) is because the higher frequencies introduced at each center of mass resampling stage did not cause aliasing. Furthermore, these high frequencies were largely filtered out by the subsequent MGS iteration, preventing a cascade of subsequent aliased curves.

2) *Quantitative Assessment:* From the results in Figs. 6–8, we see that both K_0 and N affect the performance of ACAR. To gain additional understanding into the relationship between these two parameters, we ran a series of simulations and computed two measures of error: the maximum distance error and the total area between the true spine and the final active contour. The area measure gives a global measure of performance while the maximum distance measure gives a measure of the worst local performance.

The results are shown in Fig. 9. The area error and maximum distance error as functions of N for different K_0 's are shown in Fig. 9(a) and (b), respectively. A general interpretation of these results is easiest for large K_0 since in this case the energy function is convex. For a convex problem, convergence to the global minimum is guaranteed, and if N is large enough, the discrete solution should be a good approximation of the continuous solution. Hence, the errors should approach a constant as N increases, a behavior that is confirmed in the curves for $K_0 = 250 \times 10^{-7}$ in Fig. 9(a) and (b). For smaller K_0 , the prospect of aliasing increases, especially for small N ; therefore, the irregular behavior of the curves for small N as seen in these figures confirms our expectation that these curves will be examples of various local minima. The smaller K_0 is the longer we would expect the

presence of local minima to dominate as N increases, a feature that is confirmed fairly well, especially in Fig. 9(b).

Both Fig. 9(a) and (b) confirm that for $N = 230$ the best performance is achieved by $K_0 = 8 \times 10^{-7}$, i.e., one that is neither too large nor too small. When $N = 100$, however, a larger value, $K_0 = 80 \times 10^{-7}$, achieves better performance in both cases. Aliasing causes this behavior since splits give large area and maximum distance errors. A lower cut-off frequency of the MGS iteration provided by the large K_0 prevents these grave results, at least at this point. As N is made even smaller, however, larger K_0 's simply oversmooth the active contour, leaving no possibility of approaching the true curve.

Fig. 9(c) and (d) shows plots of the area and maximum distance errors, respectively, as a function of K_0 for various N 's. These plots demonstrate the existence of optimal K_0 's and the fact that the optimal K_0 decreases with increasing N . The curves for $N = 230$ are the most dramatic, showing an optimal $K_0 = 10 \times 10^{-7}$ and rapid degradation on either side. The other curves for different N 's reveal a similar behavior on a different scale. We can see then that, in designing an active contour algorithm for a specific N , our goal to improve performance would be to lower K_0 as far as possible without crossing a drastic performance threshold. Since this threshold depends on the object, we would not choose such a low K_0 if we desired robust performance across a family of different objects.

3) *Solution Through FFT:* It was shown in Section IV that ACAR can be interpreted as a sequence of filters applied to the center of mass function with intermediate resampling. When the contour is closed, this filter can be implemented efficiently using the FFT. Fig. 10 shows a result from this FFT implementation applied to a closed ribbon. The initial configuration is shown in Fig. 10(a), and two intermediate configurations of the active contour and its final one are shown in Fig. 10(b)–(d).

C. Grayscale Images and Robustness

One way to treat grayscale images within our current framework is to segment the image into two objects: an estimated ribbon and the background (everything else). In this scenario, noise is simply the misclassification of ribbon as background or background as ribbon. One can expect that our algorithm is fairly robust to this kind of error since the direction of movement of the active contour is determined by integration. The effects of an occasional mistake would largely cancel out. Still, the success of this approach is ultimately linked to the performance of the segmentation algorithm, which makes both analysis and generalization difficult.

An alternate way to treat grayscale images is to redefine the mass function, allowing numbers between zero and one rather than just zeros and ones. The mass may reflect the probability that the pixel is in the ribbon, or it may reflect a partial volume contribution from the ribbon, or some combination thereof. In short, this allows a great deal of flexibility not permitted by hard segmentation. We sketch here an approach based on probabilistic classification.

Let λ_i , $i = 1, \dots, K$, denote the tissue classes present in an image and let $I(x, y)$ represent the image intensity. We define

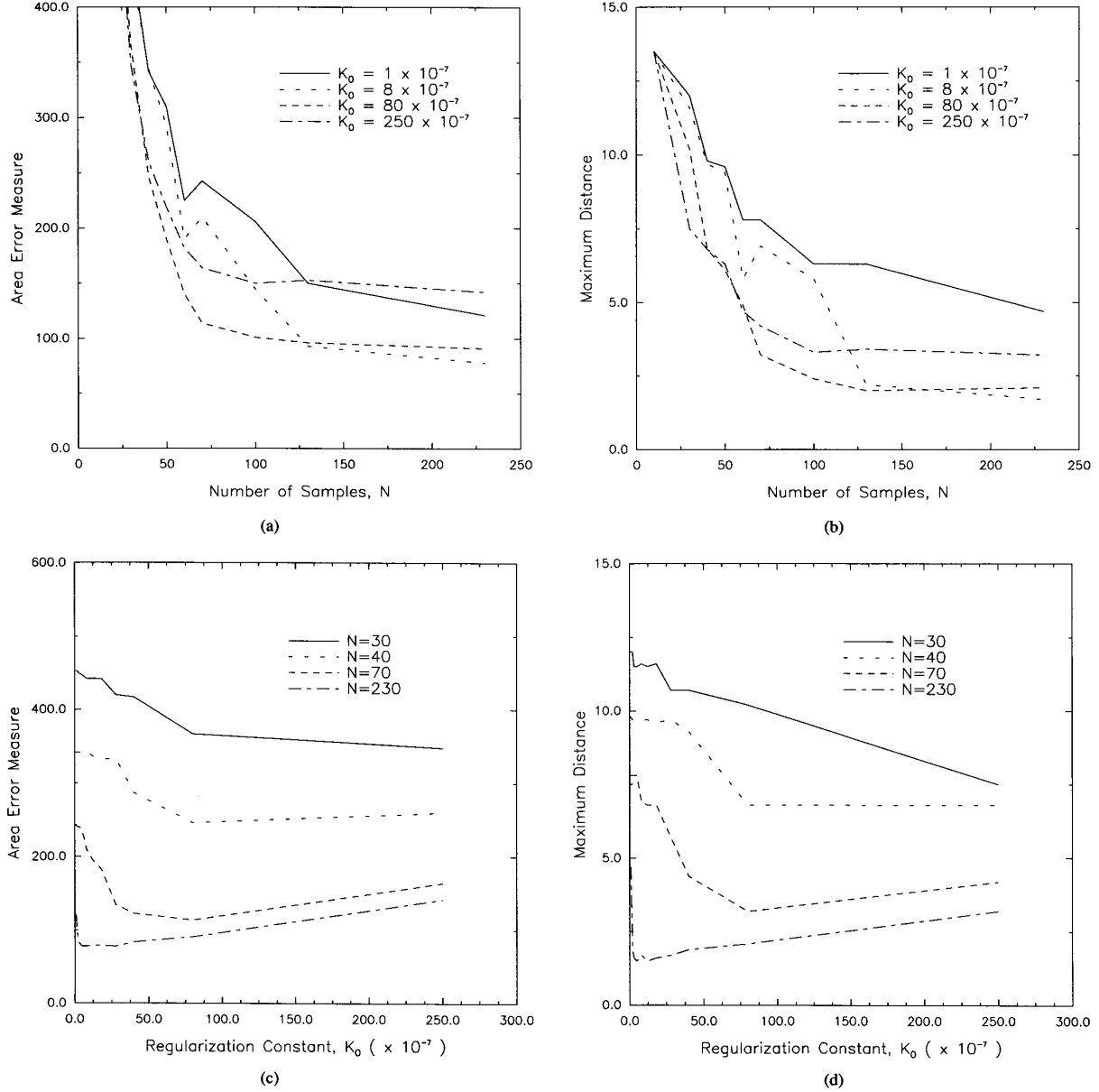


Fig. 9. Error plots: (a) Area error measure as a function of the number of samples N , for different choices of K_0 . (b) Maximum distance error measure as a function of N , for different choices of K_0 . (c) Area error measure as a function of K_0 , for different N 's. (d) Maximum distance error measure as a function of K_0 , for different N 's.

the mass function to be the posterior probability that (x, y) belongs to λ_1 , the gray matter of the cortex

$$m(x, y) = \Pr[(x, y) \in \lambda_1 | I(x, y)] = \frac{f[I(x, y) | \lambda_1] \Pr(\lambda_1)}{\sum_{i=1}^K f[I(x, y) | \lambda_i] \Pr(\lambda_i)}$$

In our experiments, we assumed the conditional probability

density functions $f(\cdot | \cdot)$ are Gaussian and of the form

$$f[I(x, y) | \lambda_i] = \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left\{ -\frac{[I(x, y) - \mu_i]^2}{2\sigma_i^2} \right\}.$$

The means μ_i and variances σ_i^2 might be known, as they are in our simulations, or not known, in which case they can be determined through partially interactive statistical methods.

With the above mass function, pixels that are either very noisy or are heavily partial-volume-averaged should have small masses; hence, they should not enter significantly into

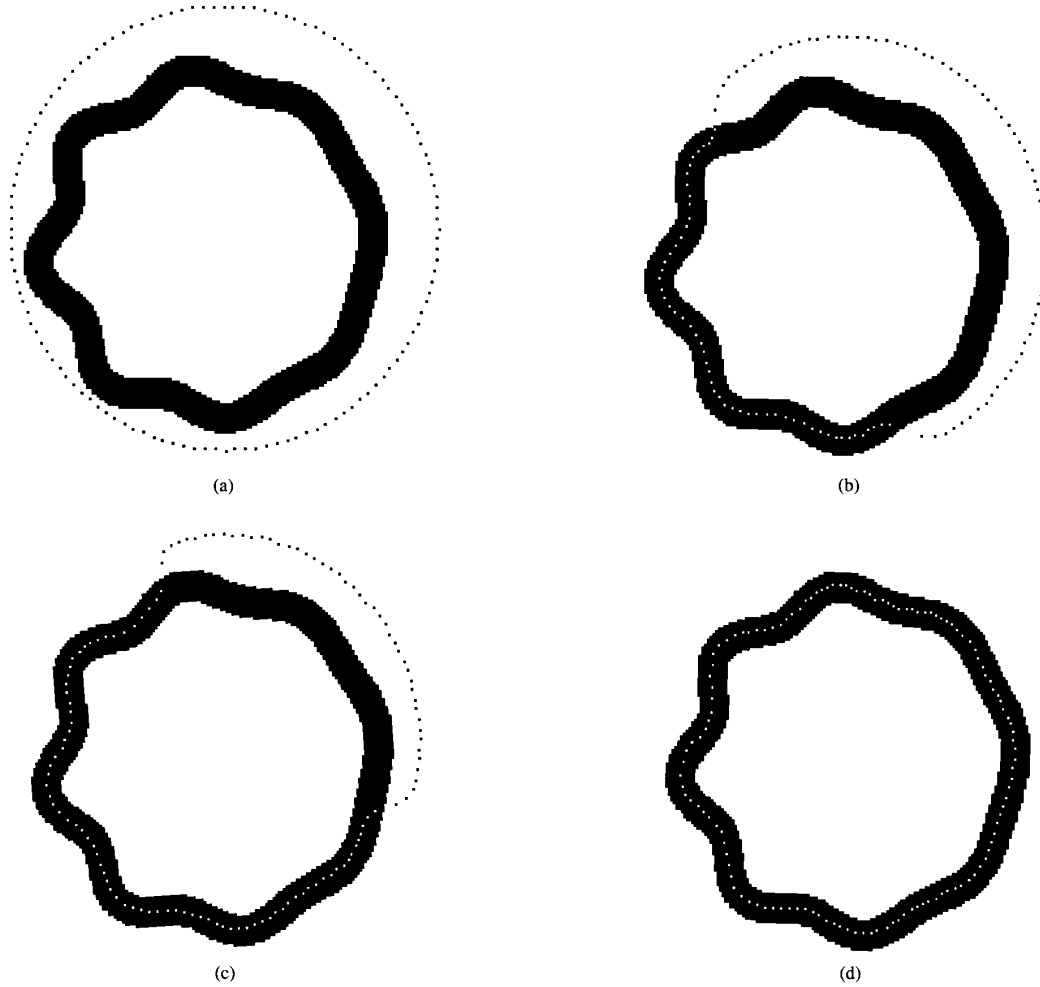


Fig. 10. The use of FFT for closed ribbons: (a) The initial configuration. (b), (c) Two intermediate configurations. (d) The final curve.

the center of mass calculation or thereby influence the deformations of the active contour. In fact, pixels with the highest probability of being cortex have the strongest effect on the active contour. Because of this property, we would expect to see a certain robustness to noise and partial volume averaging. This robustness is demonstrated via simulations in Figs. 11 and 12. In Fig. 11(a), we show a noisy image of a simulated cross-section of the brain. The standard deviation of the additive Gaussian noise is equal to 25 while the difference between the intensity values of the white matter (brighter interior) and the grey matter (ribbon) is 50, yielding an SNR of 6 db. Using the same initial curve as in Fig. 6 and $K_0 = 1.9 \times 10^{-6}$, ACAR converges to the final active contour shown in Fig. 11(b).

To demonstrate robustness quantitatively, we applied ACAR to the same underlying image with different additive noise levels. Fig. 12 shows a plot of the maximum distance between the final contour and the truth as a function of noise standard deviation. We see that up to a standard deviation of about 30—or equivalently down to an SNR of about 4.4 dB—the performance is nearly the same. For noise levels above 30,

the performance deteriorates very rapidly. This plot shows the typical characteristic for robust estimators, in which the estimate is largely independent of noise in high SNR scenarios but deteriorates rapidly below a certain threshold. This supports our claim that our active contour algorithm is robust to the effects of noise.

D. Magnetic Resonance Data

The result of applying ACAR to MR brain data is shown in Fig. 13. Here we used $K_0 = 2 \times 10^{-7}$, $N = 300$, $w = 6$, $\rho = 3$, and a probabilistic mass function was used. A larger neighborhood was used to start, and was reduced to its final value. We can see that for the most part, the solution falls very near to the spine of the outer cortex. It fails to follow the sharpest foldings of the sulci, however, and completely avoids any migration into the interior portion of the cortex. Because of this, one might conclude that K_0 was too large. But the true reason for this failing is more subtle: it is because the effective width of the cortex at the sulci is $2w$ since two layers of the cortex are juxtaposed. This allows for the possibility of many configu-

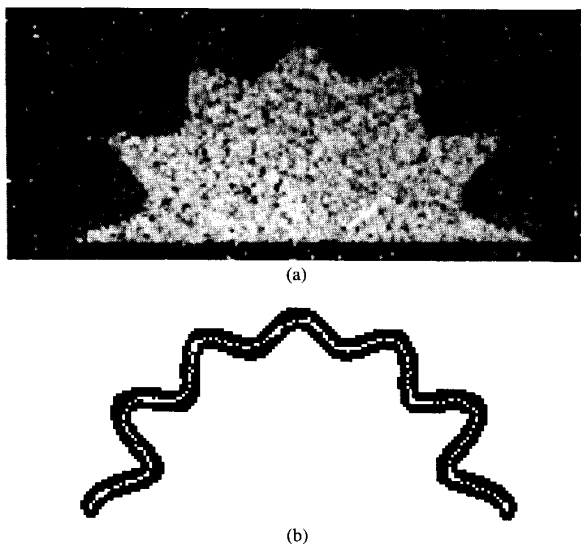


Fig. 11. The performance of ACAR on noisy data: (a) The corrupted by noisy ribbon. (b) The obtained result for the initialization of Fig. 6(a).

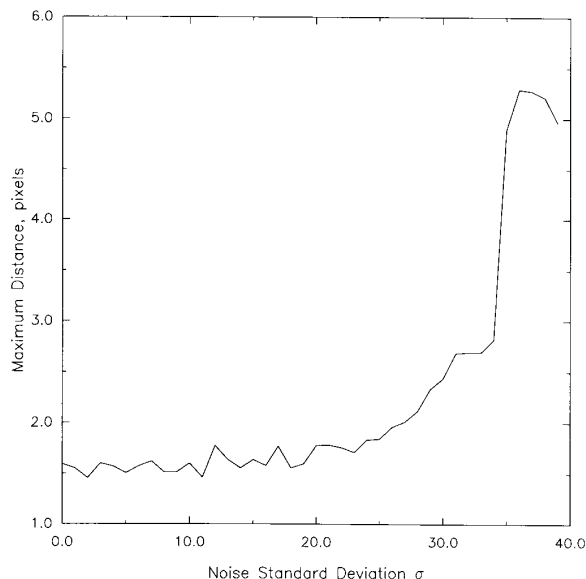


Fig. 12. The performance of ACAR as a function of the noise standard deviation.

rations that will satisfy the center of mass criterion exactly. Of these configurations, the curve most favored by the elasticity term in the energy function will be the smoothest curve, which is also the curve closest to the outer surface of the brain. For this reason, our active contour model, as it has been presented, is suitable only for reconstruction of the outer cortex.

It should be noted that stripping the skull in brain images alleviates the problem of the active contour initialization. In particular, in the experiment of this section, the active contour was initialized at a circular configuration surrounding the brain. However, in images of poor resolution or low contrast, a careful initialization of the active contour close to the cortex would be required to allow convergence to the cortical spine.

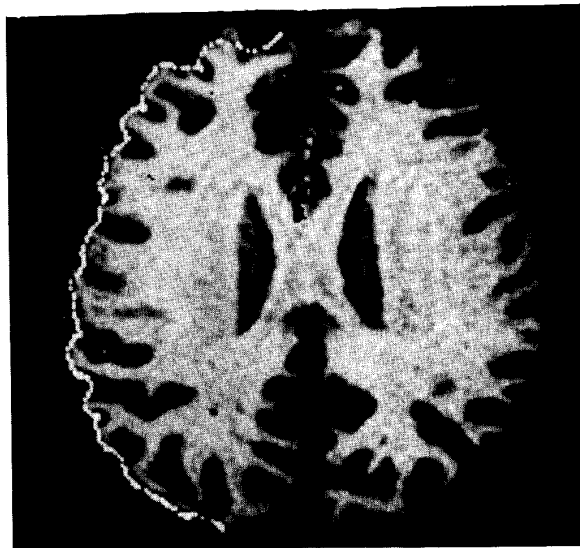


Fig. 13. The final curve obtained through ACAR: The outer cortex has been extracted, but the active contour failed to follow the sharp cortical folds.

E. Control Points

One of the great strengths of active contour models is in their ability to accommodate constraints or additional information which will influence their convergence behavior in order to meet a new objective. In this section, we sketch such a modification, with the goal of improving the previous MR results.

In order to force the active contour to track the sulci, one might imagine applying new external forces originating from within the sulci. A particularly effective approach is based on the idea of *control points*, originally proposed by Kass, Witkin, and Terzopoulos in [12]. In their model, the control points are fixed, manually provided points, each of which is connected through a spring to a fixed point of the active contour. In our application, this model can lead to dramatic failures, since we can not know in advance which active contour point should be attracted by a given control point. This problem becomes even more severe because of the convoluted shape of the cortex. Clearly, a different solution is required.

Here, we propose a new control point algorithm that differs from that proposed in [12] in two key ways (see also [45]). First, our control points are allowed to move with the active contour. Second, we do not impose a fixed relationship between a given control point and a specific point on the active contour. Combined, these two properties allow a control point to lead the active contour into the sulci, rather than force a particular point on the contour toward a particular point in the image. This attractive force is provided by defining a circular neighborhood around each control point. All points on the active contour within this neighborhood experience a new force toward the control point in proportion to their separation.

To be effective, the control points must move into the sulci. To accomplish this, we initialize them close to the initial configuration of the active contour and then let each of them move under the presence of two forces. The first force, F_c , is a constant magnitude force originating from a central point

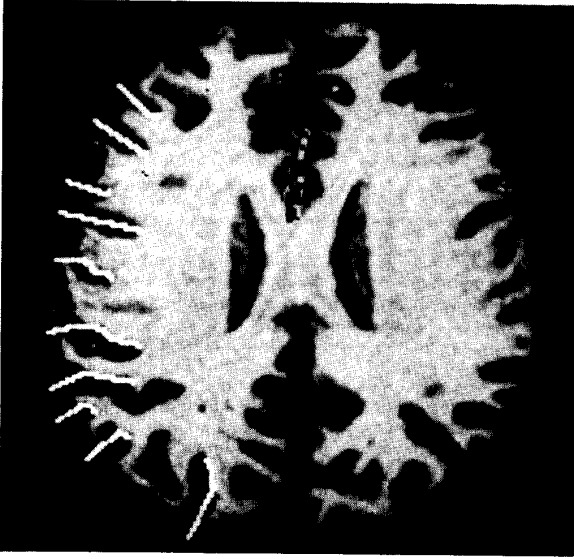


Fig. 14. The trajectories of the control points.

$\mathbf{x}_c = (x_c, y_c)$ of the brain, which is provided manually. This force reflects the tendency of the sulci to fold toward the interior of the brain. The second force, F_w , exerted on each control point is a repulsive force originating from the white matter underneath the cortex. This second force becomes active only when a circular neighborhood around a control point intersects the white matter.

Under the presence of these two forces, a control point moves radially toward \mathbf{x}_c until the repulsive force F_w from the white matter balances F_c . This force balance condition is satisfied at the deep parts of the sulci. An example of the trajectories that the control points follow is shown in Fig. 14. As the control points move toward the interior of the brain, they attract the active contour and force it to sharply fold following the shape of the sulci. In Fig. 15, we show the final active contour configuration using control points. From this figure, we see that the accuracy of the reconstruction and mapping of the cortical contours has dramatically increased over that shown in Fig. 13; some deep lying structures are now captured where they were otherwise completely missing.

VI. SUMMARY AND DISCUSSION

In this paper, we introduced the concept of ribbons for modeling the outer cortex in cross-sectional brain images and proposed an active contour algorithm (ACAR) for determining the spine of such ribbons. Our active contour model has an external force derived from an integration of the data and internal elasticity forces. We developed an iterative algorithm (MGS) to solve the nonlinear force balance equation and derived a variational formulation which ACAR solves. We then analyzed the convexity of ACAR, demonstrating that it can be made convex by choosing the regularization coefficient K_0 larger than a specified lower bound. We next analyzed ACAR in the frequency domain, showing that spines have a maximum bandwidth and that K_0 should be selected below a certain upper bound in order to preserve the fidelity of the

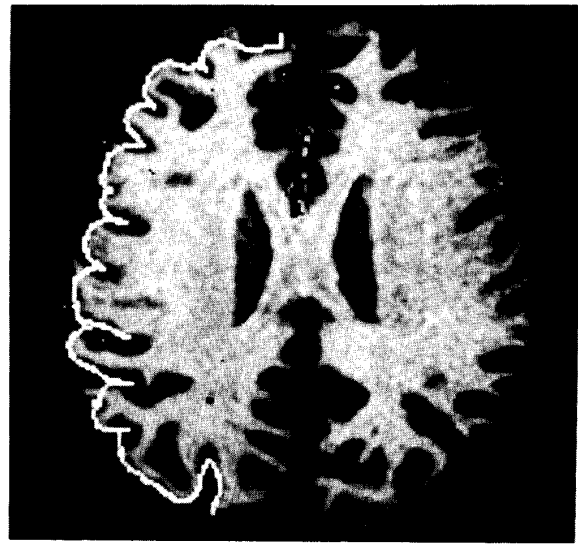


Fig. 15. The final curve obtained using ACAR and control points.

spine. Finally, we tested the validity of these bounds and gave several simulations and experimental results.

Several important insights can be drawn from the results of this paper. To start, we observed that, in general, there is no regularization constant K_0 which will simultaneously guarantee both convexity and fidelity of this active contour formulation. But, we also learned that the problem is made nonconvex by points far away from the spine on the inside of the sharpest bend in the ribbon. These observations suggest that adaptive modification of K_0 might be advantageous (see also [46]–[48]). For example, K_0 could be spatially adapted so that it was large only where the sharp bends occur. In addition, K_0 could grow smaller as the active contour approaches the spine. A combination of these two could lead to an optimal balance between isometry and fidelity, and make the result largely independent of the initial configuration.

As formulated herein ACAR penalizes only the first-order derivative of the active contour coordinate functions. One might ask whether it would be advantageous to also penalize higher order derivatives. For example, if both first- and second-order derivatives were included, then there would be two regularization constants K_0 and K_1 . Accordingly, the convexity and fidelity of the algorithm would depend on both K_0 and K_1 . Whether such an extension would prove useful in the brain-mapping application is still an open question, but our work herein provides some insight. For example, since large K_1 would restrict the bending of the active contour, one might expect poorer fidelity in general. This notion is bolstered by noting that the Butterworth filter $H(l)$ becomes second-order in this case, implying a sharper transition region and more high frequency filtering. If the spine is truly low-pass, then this might be advantageous; otherwise it may be a disadvantage. As for convexity, one can expect that for a given K_0 , increasing K_1 could turn a nonconvex problem in one that is convex.

Modification and extension of ACAR is relatively straightforward. For example, in this paper, we demonstrated an

extension to grayscale images and to closed contours. We also demonstrated an extension to allow detection of sulci using control points. Modifications to make K_0 adaptive and to extend the approach to three dimensions (by defining an active surface analog) are also exciting possibilities. Our theoretical results in this paper do not directly apply to any of these extensions, unfortunately; but it is hoped that new methods can be developed to extend the theory as well. In any event, the active contour framework has enormous potential in brain imaging and should become an important tool in the future.

ACKNOWLEDGMENT

We are grateful to Dr. R. N. Bryan for guidance and support during this research and Dr. J. Links for his careful reading and insightful comments during manuscript revision.

REFERENCES

- [1] S. P. Raya and J. K. Udupa, "Shape-based interpolation of multidimensional objects," *IEEE Trans. Med. Imag.*, vol. 9, no. 1, pp. 32-42, 1990.
- [2] J. D. Boissonnat, "Shape reconstruction from planar cross-sections," *Comput. Vision, Graphics, Image Processing*, vol. 44, pp. 1-29, 1988.
- [3] C. A. Pelizzari, G. T. Y. Chen, D. R. Spelbring, R. R. Weichselbaum, and C. T. Chen, "Accurate three-dimensional registration of CT, PET, and/or MR images of the brain," *J. Comput. Assist. Tomog.*, vol. 13, no. 1, pp. 20-26, 1989.
- [4] G. J. Carman, "Mappings of the cerebral cortex," Ph.D. dissertation, California Institute of Technology, 1990.
- [5] E. L. Schwartz and M. Merker, "Computer-aided neuroanatomy: Differential geometry of cortical surfaces and an optical flattening algorithm," *IEEE Comput. Graphics Applicat.*, vol. 11, no. 9, pp. 36-44, 1986.
- [6] E. L. Schwartz, A. Shaw, and E. Wolfson, "A numerical solution to the generalized mapmaker's problem: Flattening nonconvex polyhedral surfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 9, pp. 1005-1008, 1989.
- [7] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*. New York: Academic Press, 1982.
- [8] D. Geman, S. Geman, C. Graffigne, and P. Dong, "Boundary detection by constrained optimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 609-628, 1990.
- [9] G. P. Ashkar and J. W. Modestino, "The contour extraction problem with biomedical applications," *Comput. Graphics, Image Processing*, vol. 7, pp. 331-355, 1978.
- [10] R. Cedeberg, "Chain-link coding and segmentation for raster scan devices," *Comput. Graphics, Image Processing*, vol. 10, pp. 224-234, 1979.
- [11] D. Terzopoulos and K. Fleischer, "Deformable models," *The Visual Computer*, vol. 4, pp. 306-331, 1988.
- [12] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vision*, vol. 1, pp. 321-331, 1988.
- [13] D. Terzopoulos, A. Witkin, and M. Kass, "Constraints on deformable models: Recovering shape and nonrigid motion," *Artificial Intelligence*, vol. 36, pp. 91-123, 1988.
- [14] F. Leymarie and M. Levine, "New method for shape description based on an active contour model," *SPIE Proc., Intelligent Robots and Computer Vision VIII: Algorithms and Techniques*, vol. 1192, pp. 536-547, Nov. 1989.
- [15] F. Leymarie and M. D. Levine, "Tracking deformable objects in the plane using an active contour model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 6, pp. 617-634, 1993.
- [16] L. Liu, G. Chumuck, and C. R. Meyer, "Optimal contour approximation by deformable piecewise cubic splines," in *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, pp. 638-643, 1990.
- [17] N. Ayache *et al.*, "Building highly structured volume representations in 3-D medical images," in *Proc. 3rd Int. Sympos. Comput. Assist. Radiology, CAR '89*, 1989.
- [18] I. Carlom, D. Terzopoulos, and K. Harris, "Reconstructing and visualizing models of neuronal dendrites," in *Scientific Visualization of Physical Phenomena*. New York: Springer-Verlag, 1991, pp. 623-638.
- [19] S. R. Sandor and R. Leahy, "Automatic recognition of brain regions from magnetic resonance images," in *Proc. 26th Asilomar Conf. Signals, Syst., Comput.*, 1992.
- [20] I. Carlom, D. Terzopoulos, and K. M. Harris, "Reconstructing and visualizing models of neuronal dendrites," Digital Equipment Corp., Cambridge Research Lab., Tech. Rep. CRL 90144, 1990.
- [21] D. Terzopoulos and K. Waters, "Physically-based facial modeling, analysis, and animation," *J. Visualization Comput. Animat.*, vol. 1, pp. 73-80, 1990.
- [22] D. Terzopoulos, "The computation of visible surface representations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 4, pp. 417-438, 1988.
- [23] L. D. Cohen and I. Cohen, "A finite element method applied to new active contour models and 3-D reconstruction from cross-sections," *Inst. Nat. de Recherche en Informatique et en Automatique, Cedex, France*, Tech. Rep. 1245, 1990.
- [24] I. Cohen, L. D. Cohen, and N. Ayache, "Introducing new deformable surfaces to segment 3-D images," in *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, pp. 738-739, 1991.
- [25] ———, "Introducing deformable surfaces to segment 3-D images and infer differential structures," *Inst. Nat. de Recherche en Informatique et en Automatique, Cedex, France*, Tech. Rep. 1403, 1991.
- [26] F. Leymarie, "Tracking and describing deformable objects using active contour models," McGill University, Tech. Rep. TR-CIM-90-9, 1990.
- [27] M. O. Berger and R. Mohr, "Toward autonomy in active contour models," in *Proc. 10th Int. Conf. Pattern Recog.*, pp. 847-851, 1990.
- [28] D. J. Williams and M. Shah, "A fast algorithm for active contours and curvature estimation," *CVGIP: Image Understanding*, vol. 55, pp. 14-26, 1992.
- [29] A. Rosenfeld, "Axial representations of shape," *Comput. Vision, Graphics, Image Processing*, vol. 33, pp. 156-173, 1986.
- [30] R. Mohan and R. Nevatia, "Perceptual organization for scene segmentation and description," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 6, pp. 616-635, 1992.
- [31] E. Dougherty and C. Giardina, *Image Processing—Continuous to Discrete*. Englewood Cliffs, NJ: Prentice-Hall, 1987, vol. 1.
- [32] F. Leymarie and M. D. Levine, "Simulating the grassfire transform using an active contour model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 1, pp. 56-75, 1992.
- [33] Y. Xia, "Skeletonization via the realization of the fire-front's propagation and extinction in digital binary shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, pp. 1076-1086, 1989.
- [34] C. W. Niblack, P. B. Gibbons, and D. W. Capson, "Generating skeletons and centerlines from the distance transform," *CVGIP: Graphical Models Image Processing*, vol. 54, pp. 420-437, 1992.
- [35] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins University Press, 1983.
- [36] G. Strang, *Introduction to Applied Mathematics*. Wellesley, MA: Wellesley-Cambridge Press, 1986.
- [37] H. B. Keller, *Numerical Methods for Two-Point Boundary Value Problems*. Waltham, MA: Blaisdell, 1968.
- [38] F. B. Hildebrand, *Methods of Applied Mathematics*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1965.
- [39] W. Kaplan, *Advanced Calculus*. Reading, MA: Addison-Wesley, 1973.
- [40] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1985.
- [41] U. Grenander and G. Szego, *Toeplitz Forms and Their Applications*. Berkeley, CA: University of California Press, 1958.
- [42] C. A. Davatzikos and J. L. Prince, "An active contour model for mapping the cortex—Part I: Convexity analysis," Johns Hopkins University, Tech. Rep. JHU/ECE 93-06, 1993. (To be published.)
- [43] G. Whitten, "Scale space tracking and deformable sheet models for computational vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 7, pp. 697-706, 1993.
- [44] C. A. Davatzikos and J. L. Prince, "An active contour model for mapping the cortex—Part II: Frequency domain analysis," Johns Hopkins University, Tech. Rep. JHU/ECE 93-07, 1993.
- [45] ———, "Segmentation and mapping of highly convoluted contours with applications to medical images," in *Proc. ICASSP'92, IEEE Conf. Acoust., Speech, Signal Processing*, pp. 569-572, 1992.
- [46] D. Terzopoulos and M. Vasilescu, "Sampling and reconstruction with adaptive meshes," in *Proc. IEEE Conf. Vision Pattern Recog.*, pp. 70-75, 1991.
- [47] R. Samadani, "Adaptive snakes: Control of damping and material parameters," *SPIE Proc., Geometric Methods in Computer Vision*, vol. 1570, pp. 202-213, 1991.
- [48] C. A. Davatzikos and J. L. Prince, "Adaptive active contour algorithms for extracting and mapping thick curves," in *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, pp. 524-529, 1993.