AN ACTIVE CONTOUR ALGORITHM
FOR THICK CURVES

Chris A. Davatzikos and Jerry L. Prince

Image Analysis and Communications Laboratory
Department of Electrical and Computer Engineering
The Johns Hopkins University, Baltimore MD 21218

Technical Report JHU/ECE 92-07

A shortened version of this report has been submitted for publication in the IEEE Transactions on Pattern Analysis and Machine Intelligence. Appendices C and D, which begin on page 41 herein, are not included in the published version. When this report references itself Appendices C and D should be consulted.
An Active Contour Algorithm for Thick Curves

Chris A. Davatzikos and Jerry L. Prince

Department of Electrical and Computer Engineering
Johns Hopkins University
Baltimore, MD 21218.

Abstract — Curves with thickness, or thick curves, arise naturally in certain applications such as magnetic resonance imaging of the brain; they can also arise through low-level image processing procedures such as morphological dilation. In this paper we describe and analyze an active contour algorithm for the reconstruction and mapping of the central contour of thick curves. We address two issues of theoretical and practical importance: convergence — i.e., is a unique solution guaranteed? — and fidelity — i.e., will the solution agree with the true central contour? We show that convexity of the active contour model depends strongly on the thick curve itself, but that under certain conditions the problem is convex and a unique globally optimal solution exists. We also find through a frequency domain analysis conditions under which the solution can agree with the true thick curve. Ideally these two conditions prescribe an interval from which the regularization parameter of the problem should be selected. In cases where such an interval does not exist, however, our results suggest a range over which the regularization parameter may be selected to effectively balance global convergence versus fidelity. The active contour algorithm is evaluated on a series of experiments with synthetic curves and with a magnetic resonance image of the human brain.

I. INTRODUCTION

The problem of identifying and describing image contours has been extensively studied in the image analysis literature. In this paper we address the problem of locating and mapping the central contour of objects that we call thick curves. We define a thick curve by its central contour and a constant width, where the width is measured along lines defined by the normal vectors of the central contour (see Section II). Cross-sectional images of real objects may naturally exhibit thick curves. For example, the cortex of the human brain when viewed within cross-sectional magnetic resonance images closely matches this description. It is also possible to construct thick curves using image processing methods — e.g., morphological image processing [1]. We note, however, that dilating an arbitrary (conventional) curve with a circular structuring element does not always produce a thick curve because its thickness may not be constant. In fact, the central contour of a
thick curve has an upper bound on its curvature which is determined by its thickness, a fact that we exploit in our frequency analysis of Section IV.

Our approach is based on an active (or deformable) contour model. Active contour models have been developed by several investigators in recent years and have been used in several applications [2, 3, 4, 5, 6, 7], including medical imaging [8, 9, 10, 11, 12, 13]. Alternate methods to extract contours have generally been based on edge detection operations combined with edge linking [14] or on contour-following procedures [15]. These methods are usually computationally complex; moreover, they are not very robust in the presence of noise or in cases of very convoluted contours. Active contour methods, however, offer several advantages. First, the active contour is modeled directly as a curve, and the iterative process which deforms the contour towards a solution maintains a connected curve at all stages of processing. Thus, characteristics of the desired curve — e.g., its length, curvature, orientation, etc. — and its conformation to the data can be evaluated or imposed as an explicit part of the algorithm. Second, optimality criteria, defined via characteristics of the curve and the relationship between the curve and the data, can be specified, and an optimal solution can be sought. Finally, an explicit map between the curve and the unit interval is generated automatically.

Despite both the diverse optimality criteria imposed on active contour models and the many alternate computational methods proposed to solve these problems, the convergence behavior of the resulting algorithms is rarely discussed in the literature. In our application, we are particularly concerned with two questions: 1) will the numerical implementation of an active contour model converge? and 2) if convergence is guaranteed, will the solution agree with the true contour? Terzopoulos and his co-authors, the pioneers of this field, have given only cursory discussion of these issues [16]. Leymarie [17] and Berger and Mohr [18] have recently identified these problems, but did not give a convergence analysis. In this paper, we present an active contour model and algorithm designed to take advantage of the features of thick curves. We find conditions under which 1) our model will have a unique global minimum and our algorithm will converge to it and 2) the global minimum will be close to the true central contour of the thick curve.

Although this paper concerns only the reconstruction and mapping of thick curves within two-dimensional (2-D) images, our ultimate goal is to reconstruct and map thick surfaces from three-dimensional (3-D) data. Most particularly, we are interested in reconstructing and mapping the cortical surface of the human brain from stacks of magnetic resonance images. Most previous efforts which led toward this or a similar goal first segment each 2-D image into contour and non-contour, and then join the contours into a 3-D representation of the surface [19, 20, 21, 22]. None of these efforts, however, addressed the question of mapping the cortical surface to, say, the unit square. Carman [23] and Schwarz [24, 25] did address this question by seeking maps which are nearly isometric; but, they began their development by assuming that a representation of the cortical contour within each 2-D image existed. In their experiments, they derived these contours by hand.
Our algorithm, therefore, can be viewed as an automated preprocessing stage to either the Carman or Schwartz algorithm. We have demonstrated, however, that our method can be directly extended to 3-D by defining a deformable surface model and applying an analogous optimality criterion [26]. Thus, the methods reported in this paper will ultimately serve to circumvent the two-stage surface reconstruction method which is very common in the literature.

This remainder of this paper is organized as follows. In Section II we describe our data and active contour models. Section III derives conditions under which the energy function of the active contour model is convex and shows how problem parameters may be chosen to satisfy these conditions. We develop a frequency domain analysis in Section IV, and show that energy-minimizing contours will be too smooth unless a regularization coefficient of the problem formulation is small enough. Section V presents experiments involving both real and synthetic data, and finally, in Section VI we provide a discussion of our results and indicate future research directions.

II. ACTIVE CONTOUR MODEL

A. Thick Curves

Let \( \alpha(s) : [0, 1] \rightarrow \mathbb{R}^2 \) be a constant velocity plane curve with unit normal vector \( \mathbf{N}(s) \), as shown in Fig. 1. A thick curve is defined using two boundary curves given by \( \alpha_1(s) = \alpha(s) + w \mathbf{N}(s)/2 \) and \( \alpha_2(s) = \alpha(s) - w \mathbf{N}(s)/2 \), where \( s \in [0, 1] \) and \( w \) is a width parameter satisfying \( w > 0 \). If both \( \alpha_1(s) \) and \( \alpha_2(s) \) are simple curves — i.e., they have no self intersections — and are twice differentiable then the thick curve with central contour \( \alpha(s) \) and thickness \( w \) is defined by the set

\[
\mathcal{C} = \left\{ \mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x} = \alpha(s) + \lambda w \mathbf{N}(s)/2, 0 \leq s \leq 1, -1 \leq \lambda \leq 1 \right\}.
\]  

If either boundary curve fails to be both simple and twice differentiable then \( \alpha(s) \) does not define a thick curve; thus, not every curve can be the central contour of a thick curve. This feasibility condition defines a constraint on the maximum curvature that central contours can possess, a fact that we exploit in the frequency domain analysis of Section IV.

Our goal is to find the central layer of a thick curve using an active contour model. An active contour is a curve that is allowed to deform under internal and external forces. Internal forces reflect an elasticity property which maintains the continuity of the curve; external forces originate from the image data, and provide a mechanism to match the curve to the data. In our problem the external forces originate from points belonging to a thick curve. We assume that a presegmentation step yields a set \( \mathcal{C} \) of points belonging to this thick curve. Denoting a point in the image by \( \mathbf{x} = (x, y) \) we define the mass function as the indicator function of the thick curve:

\[
m(\mathbf{x}) = m(x, y) = \begin{cases} 
1, & \mathbf{x} \in \mathcal{C}, \\
0, & \text{otherwise}. 
\end{cases}
\]
Now consider a circular neighborhood $\mathcal{N}(x)$ of radius $\rho$ centered around $x$, as shown in Fig. 1. The mass included within this neighborhood is given by

$$
\mu(x) = \iint_{(\sigma, \tau) \in \mathcal{N}(x)} m(\sigma, \tau) d\sigma d\tau,
$$

and the center of mass within the neighborhood, which we call the center of mass function, is the point in $\mathcal{N}(x)$ given by $c(x) = (u(x), v(x)) = (u(x, y), v(x, y))$, where

$$
u(x) = \frac{1}{\mu(x)} \iint_{(\sigma, \tau) \in \mathcal{N}(x)} \sigma m(\sigma, \tau) d\sigma d\tau,
$$

$$
v(x) = \frac{1}{\mu(x)} \iint_{(\sigma, \tau) \in \mathcal{N}(x)} \tau m(\sigma, \tau) d\sigma d\tau.
$$

The center of mass function constitutes all of the external forces in our active contour model, as described in the following section.

**B. Variational Formulation**

If the width of a thick curve is exactly $2\rho$ then the central layer of $\mathcal{C}$ is the collection of points satisfying $x = c(x)$. To find this layer we could, in principle, form the collection of all points in the image that satisfy this center of mass balance condition. Such an approach, however, would be extremely unreliable in practice due to discretization, noise, blurring, etc. Furthermore, such a method does not produce a parametrization of the curve. So instead, we specify an active contour model using a variational formulation which attempts to make all points on the optimal curve satisfy this center of mass balance condition, but also includes an elasticity term to maintain connectivity and robustness to discretization and noise. Specifically, we seek the function $\hat{x}(s)$, $s \in [0, 1]$, that minimizes

$$
\mathcal{E} = \mathcal{E}_F + \mathcal{E}_E
$$

where

$$
\mathcal{E}_F = \int_0^1 \|x(s) - c(x(s))\|^2 ds
$$

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

subject to the boundary conditions:

$$
x(0) = \alpha, \quad x(1) = \beta, \quad y(0) = \gamma, \quad y(1) = \delta.
$$

Here, the coefficient $K_0$ is a regularization parameter chosen to tradeoff smoothness with the center of mass balance condition. We refer to this problem as (CVP1), where the letters “CVP” stand for continuous variational problem and the numeral “1” stands for first order. Many of our results
can be extended to higher order variational problems, but to maintain clarity in this paper we concentrate entirely on the first order problem (CVP1).

(CVP1) is an active contour model which behaves like a set of points connected with springs. Because of the energy term $\mathcal{E}_F$, (CVP1) favors solutions that coincide with the center of mass within the neighborhood around them. This reflects the desired center of mass balance property described above. The energy term $\mathcal{E}_E$ serves several purposes. First, it causes (CVP1) to favor maps that have constant velocity, a property which is often a desirable in a mapping since relative distances are preserved. Second, since there is only one constant velocity mapping for any given curve, $\mathcal{E}_E$ is a regularizing term forcing a unique solution where otherwise the problem would be ill-posed [27]. Finally, it maintains continuity of the curve.

The energy term $\mathcal{E}_E$ also has one undesirable property: it may cause the estimated curve to be too smooth. That is, $\hat{x}(s)$ will generally be a smoothed version of the desired curve $\alpha(s)$, the constant velocity mapping of the central layer of the thick curve. Although this property is undesirable in a noise-free situation, it is generally desirable in a noisy environment. In Section IV, we treat the noise-free case, and show that there are inherent tradeoffs between the size of $K_0$, the curve thickness, and the fidelity of our estimate. While these considerations are also present in the noisy case, we would expect to have to increase $K_0$ to retain noise immunity. Detailed study of this case, however, is beyond the scope of this paper. It is worth noting that we desire solutions that are connected curves, both because the desired curve $\alpha(s)$ is connected and because the discrete approximation of the derivatives with finite differences, which is usually done in practice, is valid only when the samples of the active contour are close enough. Thus, we generally wish to discourage discontinuities or splits of the type that are often deliberately sought in other boundary reconstruction approaches [28, 5, 14].

**Relation to Snakes.** The formulation of (CVP1) is similar to that of snakes [16], except that our external forces originate from the center of mass function rather than from the edges. As we shall see, the equilibrium equations developed for a discrete analog to (CVP1) are virtually identical to those of snakes except for this difference in the external forces. We believe that the local averaging inherent to the calculation of the center of mass provides a measure of robustness to noise, which might be exploited in standard snake formulations if the concept of “thick edges” were developed. Such a generalization, however, is beyond the scope of this paper.

**C. Optimality Conditions**

The calculus of variations [29] may be used to find the necessary conditions that the solution $\hat{x}(s)$ must satisfy. For (CVP1), these conditions are given by the following Euler equations:

$$
(x - u)(1 - v_x) - (y - v)v_x - K_0 x'' = 0
$$

(8a)
\[(y-v)(1-v_y)-(x-u)u_y-K_0y''\] = 0, \hspace{1cm} (8b) 

together with the original boundary conditions given in (7). Here we have used the following simplified notation: \(x = x(s), y = y(s), u = u(x(s), y(s)), \) and \(v = v(x(s), y(s))\). Subscripts are used to denote partial derivatives — e.g., \(v_y\) is the partial derivative of \(v\) with respect to \(y\) — and a double-prime superscript means second derivative with respect to \(s\).

In general, these necessary conditions are also sufficient only if (CVP1) is convex, which may not be the case. Therefore, curves that satisfy the Euler equations may be only locally optimal, and they may exhibit undesirable properties. In Section III we describe a criterion for choosing \(K_0\) that guarantees the local convexity of (CVP1) for the discrete analog described below. In Section IV, we discuss (among other things) the behavior of locally optimal solutions which may result from the numerical solution of a non-convex formulation of (CVP1).

**D. Discrete Analog**

To solve (CVP1) numerically, we model a curve using discrete sequences that represent sampled values of the functions \(x(s)\) and \(y(s)\) as follows:

\[x_i = x \left( \frac{i}{N} \right), \quad y_i = y \left( \frac{i}{N} \right), \quad i = 0, 1, ..., N.\]

Since the boundary conditions in (7) fix the values of \(x(s)\) and \(y(s)\) at \(s = 0\) and \(s = 1\), the free values of the discrete curve can be represented by the vector

\[d = [x_1, x_2, ..., x_{N-1}, y_1, y_2, ..., y_{N-1}]^T.\] \hspace{1cm} (9)

Using the following difference approximations of the required derivatives

\[x' \approx \frac{x_{i+1} - x_i}{1/N} = N(x_{i+1} - x_i),\]
\[y' \approx \frac{y_{i+1} - y_i}{1/N} = N(y_{i+1} - y_i),\]
\[x'' \approx N^2(x_{i+1} + x_{i-1} - 2x_i),\]
\[y'' \approx N^2(y_{i+1} + y_{i-1} - 2y_i),\]

a discrete approximation of the energy function \(\mathcal{E}\) is then given by

\[\tilde{\mathcal{E}}(d) = \tilde{\mathcal{E}}_F(d) + \tilde{\mathcal{E}}_E(d),\] \hspace{1cm} (10)

where

\[\tilde{\mathcal{E}}_F(d) = \frac{1}{N} \sum_{i=1}^{N-1} \left[ (x_i - u_i)^2 + (y_i - v_i)^2 \right] \] \hspace{1cm} (11a)
\[+ \frac{1}{2N} \left[ (\alpha - u^0)^2 + (\beta - u^N)^2 + (\gamma - v^0)^2 + (\delta - v^N)^2 \right].\]
\[
\hat{\mathcal{E}}_E(d) = K_0 N \sum_{i=1}^{N-2} \left[ (x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 \right]
+ K_0 N \left[ (x_1 - \alpha)^2 + (x_{N-1} - \beta)^2 + (y_1 - \gamma)^2 + (y_{N-1} - \delta)^2 \right].
\]

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., \(u^i = u(x_i, y_i)\).

For the discrete curve \(d\) to be optimal, it must satisfy the necessary conditions
\[
\frac{\partial \hat{\mathcal{E}}(d)}{\partial x_i} = 0, \quad \frac{\partial \hat{\mathcal{E}}(d)}{\partial y_i} = 0, \quad i = 1, 2, \ldots, N - 1,
\]
which leads directly to the following discrete form of Equation (8):
\[
\begin{align*}
(x_i - u^i)(1 - u_x) - (y_i - v^i)v_x & = 0, \quad i = 1, 2, \ldots, N - 1, \\
K_0 N^2 (2x_i - x_{i+1} - x_{i-1}) & = 0, \quad i = 1, 2, \ldots, N - 1
\end{align*}
\]
\[
\begin{align*}
(y_i - v^i)(1 - v_y) - (x_i - u^i)u_y & = 0, \quad i = 1, 2, \ldots, N - 1, \\
K_0 N^2 (2y_i - y_{i+1} - y_{i-1}) & = 0, \quad i = 1, 2, \ldots, N - 1.
\end{align*}
\]

The following discrete form of the original boundary conditions must also be satisfied:
\[
x_0 = \alpha, \quad x_N = \beta, \quad y_0 = \gamma, \quad y_N = \delta.
\]

In order to derive a more compact form for Equation (12), we introduce the following notation:
\[
b = [\alpha, 0, \ldots, 0, \beta, \gamma, 0, \ldots, 0, \delta]^T,
\]
\[
A = \begin{bmatrix}
B & 0 \\
- & - & - \\
0 & B
\end{bmatrix},
\]
where \(B\) is the following \((N - 1) \times (N - 1)\) tridiagonal matrix:
\[
B = \begin{bmatrix}
2 & -1 & 0 & \ldots & 0 & 0 \\
-1 & 2 & -1 & 0 & \ldots & 0 \\
& \ddots & \ddots & \ddots & \ddots & \ddots \\
& & \ddots & \ddots & \ddots & \ddots \\
0 & \ldots & 0 & -1 & 2 & -1 \\
0 & 0 & \ldots & 0 & -1 & 2
\end{bmatrix}.
\]
We also define the vector function \( \phi(d) = [\phi_1(d), \ldots, \phi_{2N-2}(d)]^T \) where the elements of \( \phi(d) \) are given by

\[
\phi_i(d) = (x_i - u^i) (1 - u^i_z) - (y_i - v^i) v^i_z, \quad i = 1, \ldots, N - 1, \tag{16a}
\]

\[
\phi_{i+N-1}(d) = (y_i - v^i) (1 - v^i_y) - (x_i - u^i) u^i_y, \quad i = 1, \ldots, N - 1. \tag{16b}
\]

Finally we define \( f \) as follows:

\[
f = [u^1, \ldots, u^{N-1}, v^1, \ldots, v^{N-1}]^T. \tag{17}
\]

Using the notation introduced above the discrete energy of (10) can be written as

\[
\hat{\mathcal{E}}(d) = \frac{1}{N} (d - f)^T (d - f) + K_0 N \left[ d^T A d - 2 b^T d \right] + C, \tag{18}
\]

where

\[
C = \frac{1}{2N} \left[ (\alpha - v^0)^2 + (\beta - u^N)^2 + (\gamma - v^0)^2 + (\delta - v^N)^2 \right] + K_0 N \left[ \alpha^2 + \beta^2 + \gamma^2 + \delta^2 \right],
\]

and the necessary conditions given in (12) can be written as

\[
\phi(d) + K_0 N^2 A d = K_0 N^2 b. \tag{19}
\]

We will henceforth refer to this discrete version of (CVP1) as (DVP1).

III. ACTIVE CONTOUR ALGORITHM

In this section, we focus exclusively on the discrete problem (DVP1), since this is the problem that must be solved in practice. Specifically, we want to develop an algorithm for finding the discrete curve \( \hat{d} \) that minimizes \( \hat{\mathcal{E}}(d) \). Since \( \hat{\mathcal{E}}(d) \) is not necessarily convex, a deterministic descent algorithm or, equivalently, an approach designed to find a solution to the necessary conditions of (19), will yield an estimate that may be only locally optimal. We may consider alternate approaches such as serial repetitions of a deterministic algorithm from different initial conditions or stochastic algorithms such as simulated annealing; but the computation time of these approaches is generally excessive, especially since we intend to implement a three-dimensional version in the future. Our approach is therefore based on deterministic descent algorithms, one method a simple gradient descent and another a modified Gauss-Seidel approach. Since we want these algorithms to converge to meaningful solutions, we begin this section by studying the conditions under which \( \hat{\mathcal{E}}(d) \) is convex. We conclude with a discussion of our numerical procedures.
A. Convexity of the Energy Function

The major difficulty with showing the convexity of $\tilde{E}(d)$ is that it depends on the underlying data. This situation is in sharp contrast to the common least squares problem of estimating $x$ under the model $Ax = b$, where convexity is guaranteed regardless of the data $b$. Therefore, in order to discuss the convexity of our problem we must make assumptions about the nature of the true object. In the major theoretical result of this section, we give sufficient conditions for the convexity of $\tilde{E}(d)$ under the premise that each neighborhood $\mathcal{N}(x_i), i = 1, \ldots, N,$ intersects a single boundary curve of $C$ which comprises a circular arc within the intersection. This is often approximately true in practice; however, we stress the fact that condition is only strictly true when the thick curve is a portion of an annulus. We maintain that although our results can be rigorously stated only for this very specific case, they have a broader application in a practical sense. This conclusion is further justified through the simulations given in Section V.

Even in the simple case of an annulus, the proof of convexity is surprisingly complicated. To demonstrate that local minima do in fact exist and to give a flavor of their nature, we provide a simple example. Consider the annulus depicted in Fig. 2. Both of the curves shown in Fig. 2a and Fig. 2b (represented by a sequence of points which sample them) satisfy the necessary conditions of (19) when $K_o = 10^{-7}$. The curve in Fig. 2a, however, has energy 1.4 while the curve in Fig. 2b has energy 7.1; the latter curve is therefore a local minimum. We have found in our experiments that the presence of splits such as those in the curve of Fig. 2b is a characteristic of local minima. We will have more to say about this in Section IV.

Any descent algorithm begins with an initial curve, represented by an initial vector $d_0$ together with the boundary conditions, and modifies this curve iteratively to reduce the energy function. We now consider the convexity of $\tilde{E}(d)$ for initial curves that are sufficiently close to the thick curve $C$. More precisely, letting $x_i = (x_i, y_i), i = 0, \ldots, N,$ be the points on the curve, we require that $\mathcal{N}(x_i) \cap C \neq \emptyset, i = 0, \ldots, N,$ where $\cap$ denotes set intersection and $\emptyset$ denotes the empty set. We see that this condition is equivalent to requiring that the boundary points $(x_0, y_0)$ and $(x_N, y_N)$ are within $\rho$ (the radius of the neighborhood $\mathcal{N}(\cdot)$) of $C$ and that $d \in D$ where

$$D = \{e \in \mathbb{R}^{2N-2} | \mathcal{N}(e_i, e_{i+1}) \cap C \neq \emptyset, i = 1, \ldots, N - 1\}. \tag{20}$$

This condition is also equivalent to assuming that the points in the curve represented by $d$ must come from the dilation of $C$ with a disk of radius $\rho$ centered at the origin. We now restrict our attention to the convexity of $\tilde{E}(d)$ when $C$ is piece of an annulus, which is shown in Fig. 3, and we examine only the case where $\rho \leq w/2$; hence a given neighborhood can only intersect $\alpha_1$ (Fig. 3a) or $\alpha_2$ (Fig. 3b), but not both.

Since each neighborhood $\mathcal{N}(x_i)$ intersects $C$, the center of mass function $c(x_i)$ is well defined. Let $R$ be the radius of $\alpha(s)$. Now consider the polar coordinate system centered at the center of the circle implied by $\alpha(s)$. In this coordinate system the coordinates of $x_i$ are $(\tau_i, \phi_i)$ while those
of $c(x_i)$ are $(\tau_i + n(\tau_i), \phi_i)$, as shown in Fig. 3a and Fig. 3b. It is clear from the geometry that $n(\cdot)$ is only a function of the radial coordinate. To simplify our notation in the sequel we refer to $n(\tau_i)$ as $\bar{n}_i$, and use $n'_i$ and $n''_i$ to denote its first and second derivatives, respectively.

We now turn our attention to the Hessian matrix $H$ of $\tilde{\epsilon}(d)$ and consider points $d$ in the interior of $\mathcal{D}$. For $\tilde{\epsilon}(d)$ to be strictly convex at $d$, $H$ is required to be positive definite at $d$. We now find a condition under which the minimum eigenvalue of $H$, denoted $\lambda_{\min}(H)$, is strictly positive. From (10) we see that

$$H = H_F + H_E,$$

and from standard matrix analysis [30] it follows that

$$\lambda_{\min}(H) = \lambda_{\min}(H_F + H_E) \geq \lambda_{\min}(H_F) + \lambda_{\min}(H_E).$$

From (18) we see that

$$H_E = 2K_0 N A,$$

which has a minimum eigenvalue of $2K_0 N \lambda_{\min}$, where $\lambda_{\min}$ is the minimum eigenvalue of $A$. Note that this term does not depend on the data.

The form of $H_F$ leads to considerable simplification. In particular, from (11a) we see that

$$\frac{\partial^2 \tilde{\epsilon}_F(d)}{\partial x_i \partial x_j} = \frac{\partial^2 \tilde{\epsilon}_F(d)}{\partial y_i \partial y_j} = \frac{\partial^2 \tilde{\epsilon}_F(d)}{\partial x_i \partial y_j} = 0, \quad i \neq j$$

which immediately implies that

$$H_F = [f_{ij}] = \begin{bmatrix} \Sigma & \Gamma \\ \Gamma & \Psi \end{bmatrix},$$

where $\Sigma$, $\Gamma$, and $\Psi$ are diagonal matrices. If $\lambda$ is an eigenvalue of $H_F$ and its corresponding eigenvector is $z = [x^T \ y^T]^T$ then since $H_F z = \lambda z$ we have

$$\Sigma x + \Gamma y = \lambda x,$$

$$\Gamma x + \Psi y = \lambda y.$$

Since $\Sigma$, $\Gamma$, and $\Psi$ are diagonal, the system may be further broken down into $(N - 1) \times 2 \times 2$ equations of the form

$$D_i \begin{bmatrix} x_i \\ y_i \end{bmatrix} = \lambda \begin{bmatrix} x_i \\ y_i \end{bmatrix}, \quad i = 1, 2, ..., N - 1,$$

where

$$D_i = \begin{bmatrix} \sigma_i & \gamma_i \\ \gamma_i & \psi_i \end{bmatrix}.$$

Each $2 \times 2$ system in (25) must be satisfied by the eigenvalue/eigenvector pair of the larger system (24), and this can happen in only one of two ways: either $\lambda$ is an eigenvalue of $D_i$ with corresponding eigenvector $[x_i \ y_i]^T$; or $x_i = y_i = 0$ and $\lambda$ is arbitrary. Since $z$ is an eigenvector,
$z \neq 0$; hence, $\lambda$ is an eigenvalue of $D_i$, for some $i$. We can also argue this point in reverse. Suppose $\lambda$ is an eigenvalue of $D_i$ for some $i$. We can then construct an eigenvector of the larger system by setting $[x_i, y_i]^T$ to the corresponding eigenvector of $D_i$ and setting all other elements of $z$ to zero. Hence, $\lambda$ is also an eigenvalue of the larger system. Thus, we have the important conclusion: the eigenvalues of $H_F$ can be determined by finding the eigenvalues of the $2 \times 2$ matrices $D_i$ for $i = 1, \ldots, N - 1$. Furthermore, the eigenvalues of a $2 \times 2$ system have an explicit form resulting from use of the quadratic formula. We will denote these eigenvalues by $\lambda_{i1}, \lambda_{i2}$, $i = 1, \ldots, N - 1$.

Putting these results together we can now state that $H$ is guaranteed to be positive definite if

$$2K_0N\lambda_{\min} + \min_i \{\lambda_{i1}, \lambda_{i2}\} > 0. \tag{27}$$

In Appendix A we show that for the special case in which the object is an annulus and $\rho \leq w/2$ the eigenvalues are given by $\lambda_{i1} = 2(n_in_i'' + (n_i')^2)/N$ and $\lambda_{i2} = 2(n_in_i'/\tau_i)/N$. The following proposition then follows from (27) immediately:

**Proposition 1** $\hat{\mathcal{E}}(d)$ is strictly convex at each $d$ in $D$ if $K_0 > K_{LB}$ where

$$K_{LB} = \frac{-\min_i \left\{ (n_in_i'' + (n_i')^2), \left(\frac{n_in_i'}{\tau_i}\right) \right\}}{N^2\lambda_{\min}}. \tag{28}$$

The following corollary follows from elementary optimization theory:

**Corollary 2** Let $D = D_1 \cup D_0$ where $D_0$ is a convex region containing $d^\circ$, the global minimum of $\mathcal{E}(d)$ in $D$. Then if $d_0 \in D_0$ and $K_0 > K_{LB}$, a deterministic descent method with appropriately small step-size will converge to $d^\circ$. \qed

Proposition 1 and Corollary 2 together provide conditions that are sufficient to guarantee the convergence of our active contour algorithm. Since these are sufficient conditions only, however, is it possible that $K_0 < K_{LB}$ may still yield a convex problem. Since large $K_0$'s oversmooth the curve, we note that in practice knowledge of $K_{LB}$ provides only a guideline that is rarely observed. We shall have more to say about this in Section V. We conclude this section by exploring further some issues surrounding $K_{LB}$ and $D_0$ and describing the actual numerical algorithms.

**Numerator of $K_{LB}$.** In [31] we derive an expression for $n_i$ for the special case in which the object is an annulus and $\rho \leq w/2$. Although $n_i$ is an explicit function of only $\tau_i$, it is also an implicit function of $R$ and $\rho$; however, as shown in [31], we may treat it as a function of just the ratios $\tau_i/\rho$ and $R/\rho$. In Fig. 4 we plot $n_in_i'' + (n_i')^2$ and $n_in_i'/\tau_i$ as functions of $\tau_i/\rho$ for different values of $R/\rho$, assuming that $\rho \leq w$. From these plots we conclude that $\lambda_{i1}$ is always positive, but that $\lambda_{i2}$ is either positive or negative. Therefore, since by definition $K_0 > 0$, only $\lambda_{i2} < 0$ is of
any concern in the selection of $K_0$. Although it is only suggested by the plots in Fig. 4b we, have identified a grave complication in other more detailed plots: as $R/\rho \to 1$ and $r_i/\rho \to 0$, $\lambda_2 \to -\infty$. This implies that $K_{LB}$ must approach $+\infty$ in order to guarantee convexity. Therefore, although we would never expect to see this exact anomalous condition in practice, we should expect problems when the boundary of the object is wrapped tightly around a neighborhood.

**Denominator of $K_{LB}$**. From (28), it appears that one can make $K_{LB}$ arbitrarily small by increasing $N$. This is not true, however, since $\lambda_{\min}$ also changes with increasing $N$. In fact, the minimum eigenvalue of $A$ is easily found by noting that its its eigenvalues coincide with the eigenvalues of $B$. Then since $B$ is Toeplitz its eigenvalues can be calculated through a recurrence formula [32], yielding

$$\lambda_{\min} = 2 - 2 \cos \left(\frac{\pi}{N}\right).$$  \hspace{1cm} (29)

Then using the series expansion of the cosine it is easily shown that

$$\lim_{N \to \infty} N^2 \lambda_{\min} = \pi^2.$$  \hspace{1cm} (30)

In fact, plots of the quantity $N^2 \lambda_{\min}$ show that it converges fairly fast to a value close to $\pi^2$ and, therefore, in practice we can simplify $K_{LB}$ by setting $N^2 \lambda_{\min} \approx \pi^2$.

**Existence of $D_0$**. In [31] we show that a subset $D_0$ containing the global minimum always exists. In practice, however, we cannot ever guarantee that $d_0$ is in $D_0$. This problem motivates a two-stage algorithm, in which the first stage is designed to seek out a suitable starting point for the second stage, a standard descent algorithm. For the first stage we use a heuristic multiresolution approach that starts with a very general initial configuration, such as a semicircle, and a relatively large neighborhood. The curve is then allowed to deform toward $C$ while the neighborhood size is made progressively smaller until it reaches its final size. The second stage simply deforms the curve according to one of the descent algorithms described in the following section without modification of the neighborhood size. This heuristic approach does not guarantee convergence to the true global minimum since the first stage does not guarantee that the second stage will begin in $D_0$. This is a practical problem of great importance, but is not the primary topic of this paper.

**B. Numerical Solution**

We have used two methods to solve (DVP1) numerically: gradient descent and a modified version of Gauss-Seidel. If $K_0$ is chosen to obey (28) then the energy function is convex and the gradient descent method, with a sufficiently small step-size, is guaranteed to converge to the unique global minimum of $\tilde{E}(d)$ in $D$, provided that $d_0 \in D_0$. Precalculation of $c(x)$ at each point $x$ in the image significantly reduces total computation time since the costly step of averaging within a
neighborhood is avoided. While in principle we could also precalculate the partial derivatives of $c(x)$, this increases the total storage requirements by a factor of five.

The second method we use is a modified Gauss-Seidel (MGS) method. Gauss-Seidel is often used in the solution of linear, sparse systems of high dimensionality. In our case the system of necessary equations is nonlinear; however, if the center of mass function is temporarily assumed to be fixed then the system is linear and Gauss-Seidel can be used. Accordingly, MGS is the iterative application of a single Gauss-Seidel iteration followed by a center of mass update. It can also be shown to be equivalent to a single step of the gradient descent method with step-size equal to $(2K_0N^2 + 1)^{-1}$ if the derivatives of $u$ and $v$ are set equal to zero. This method is significantly faster than the gradient descent since it does not require computation of the partial derivatives of $c(x)$; however, we have no proof of its convergence.

MGS gives our active contour algorithm a physical interpretation. After setting the center of mass derivatives to zero in (19), the Euler equations for $(x_i, y_i)$ can be written as

$$u^i - x_i + K_0N^2(x_{i+1} - x_i) + K_0N^2(x_{i-1} - x_i) = 0,$$

$$v^i - y_i + K_0N^2(y_{i+1} - y_i) + K_0N^2(y_{i-1} - y_i) = 0.$$

These equations are the equilibrium equations for the point $x_i$ under the presence of three forces exerted on it. The first force originates from the center of mass $c(x_i)$, and reflects the tendency of the active contour to move towards $C$. The other two forces originate from the neighbors of $x_i$ and reflect the elasticity property of the deformable curve. Gauss-Seidel simply solves these equations for $(x_i, y_i)$, which is equivalent to finding a new position where the three forces applied to $(x_i, y_i)$ balance. The (MGS) method is faster than gradient descent and offers an intuitive interpretation for the curve deformations. In the next section we will see that it also offers a good interpretation of the curve deformations via frequency domain analysis.

## IV. FREQUENCY DOMAIN ANALYSIS

As described in Section II, an active contour is represented by the coordinate functions $x(s)$ and $y(s)$ on the interval $[0, 1]$, where the values of these functions are fixed at $s = 0$ and $s = 1$. In this section we interpret the action of our active contour algorithm on $x(s)$ and $y(s)$ as one of frequency filtering. It must be understood, however, that this interpretation is technically incorrect since the boundary conditions never change as the algorithm progresses. On the other hand, points that are not near the boundary are modified exactly according to the convolution integral (or sum in the discrete case) provided that the impulse response is short. Thus, while a rigorous interpretation as frequency filtering is not possible, we find that this interpretation provides a great deal of insight into the actual behavior of active contours.
Throughout this section we assume that $x(s)$ and $y(s)$ can be represented by the following Fourier series representation

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

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

(31a)

(31b)

The corresponding first and second order derivatives are given by

$$x'(s) = 2\pi \sum_{l=1}^{\infty} \{-lX_{cl} \sin(2\pi ls) + lX_{sl} \cos(2\pi ls)\}$$

$$y'(s) = 2\pi \sum_{l=1}^{\infty} \{-lY_{cl} \sin(2\pi ls) + lY_{sl} \cos(2\pi ls)\}$$

$$x''(s) = -4\pi^2 \sum_{l=1}^{\infty} l^2 \{X_{cl} \cos(2\pi ls) + X_{sl} \sin(2\pi ls)\}$$

$$y''(s) = -4\pi^2 \sum_{l=1}^{\infty} l^2 \{Y_{cl} \cos(2\pi ls) + Y_{sl} \sin(2\pi ls)\} .$$

(32a)

(32b)

(32c)

(32d)

A. Frequency Domain Interpretation of the Energy Function

A simple frequency interpretation of the energy function is found using Parseval’s theorem. In particular, it can be shown that the functions $x(s)$ and $y(s)$ that minimize

$$\mathcal{E}_E = K_0 \int_0^1 \left[ (x')^2 + (y')^2 \right] ds$$

have Fourier coefficients that minimize

$$\mathcal{E}_E = 2\pi^2 K_0 \sum_{l=1}^{\infty} l^2 F_l^2$$

(33)

(34)

where

$$F_l^2 = (X_{cl})^2 + (X_{sl})^2 + (Y_{cl})^2 + (Y_{sl})^2 .$$

(35)

Since differentiation of the coordinate functions in (33) causes an $l^2$ term to appear in the sum of (34), we see that curves with energy in the high frequencies are penalized while ones with a spectrum concentrated near $l = 0$ are not. Therefore, if we sample all the possible parametrizations of a particular curve with the same number of points, the parametrization that minimizes $\mathcal{E}_E$ will be minimally aliased since the high frequency energy in this curve is smallest.

In practice we only have an estimate of the curve to be parametrized, not the complete family of parametrizations. This estimate is given by the center of mass function $c(x(s)) = (u(s), v(s))$, and is included in the active contour model by adding the term $\mathcal{E}_F$ into the energy function. Applying
Parseval's theorem to the sum of the two energy terms yields
\[
\mathcal{E}_E + \mathcal{E}_F = K_0 \int_0^1 \left[ \left( x' \right)^2 + \left( y' \right)^2 \right] \, ds + \int_0^1 \left[ \left( x - u(x, y) \right)^2 + \left( y - v(x, y) \right)^2 \right] \, ds
\]
\[
= 2\pi^2 K_0 \sum_{l=1}^{\infty} \left[ l^2 F_l^2 + \left( X_0 - U_0 \right)^2 + \left( Y_0 - V_0 \right)^2 \right]
+ \frac{1}{2} \sum_{l=1}^{\infty} \left[ \left( X_{c_l} - U_{c_l} \right)^2 + \left( X_{s_l} - U_{s_l} \right)^2 + \left( Y_{c_l} - V_{c_l} \right)^2 + \left( Y_{s_l} - V_{s_l} \right)^2 \right]
\]
(36)

where \( U_{c_l}, U_{s_l}, V_{c_l}, \) and \( V_{s_l} \) are the Fourier coefficients for the expansions of the center of mass coordinate functions \( u(s) \) and \( v(s) \), defined in an analogous fashion as (31). We now see that the pair \( (x(s), y(s)) \) that minimizes \( \mathcal{E}_E + \mathcal{E}_F \) will tend to have low energy in the high frequency components, but will also tend to match the spectrum of \( (u(s), v(s)) \). This principle shows that the inherent smoothing property of the active contour algorithm can be also be interpreted as the favoring of low frequency components. We will reveal the principle behind this characteristic even more clearly in the following section.

**B. Filtering Interpretation of the Active Contour Algorithm**

As mentioned in Section III, a single iteration of the modified Gauss-Seidel (MGS) algorithm assumes that the center of mass function is fixed. Fixing the center of mass function implies
\[
u_x(s) = 0, \quad u_y(s) = 0, \quad v_x(s) = 0, \quad v_y(s) = 0,
\]
and (8a) and (8b) become
\[
\left( x(s) - u(s) \right) - K_0 x''(s) = 0
\]
\[
(38a)
\]
\[
( y(s) - v(s) ) - K_0 y''(s) = 0.
\]
\[
(38b)
\]

Substituting the Fourier expansions of \( x(s), y(s), u(s), \) and \( v(s) \) into (38) and simplifying yields \( X_0 = U_0, Y_0 = V_0, \) and
\[
X_{c_l} = H(l) \, U_{c_l}, \quad l = 0, 1, \ldots, \infty
\]
\[
(39a)
\]
\[
X_{s_l} = H(l) \, U_{s_l}, \quad l = 0, 1, \ldots, \infty
\]
\[
(39b)
\]
\[
Y_{c_l} = H(l) \, V_{c_l}, \quad l = 0, 1, \ldots, \infty
\]
\[
(39c)
\]
\[
Y_{s_l} = H(l) \, V_{s_l}, \quad l = 0, 1, \ldots, \infty
\]
\[
(39d)
\]

where
\[
H(l) = \frac{1}{1 + K_0 \pi^2 l^2}
\]
(40)

\( H(l) \) is the transfer function of a first order discrete Butterworth low pass filter with cutoff frequency
\[
l_c = \frac{1}{2\pi \sqrt{K_0}}.
\]
(41)
In order to actually implement this filter, MGS would have to iterate a large number of times with the same center of mass function, which is not done. Instead, we apply just one iteration and then update the center of mass function given the new curve. Nevertheless, a single iteration of MGS acts somewhat like a low pass filter applied to the current center of mass function. 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. Therefore, the active contour algorithm behaves like a sequence of filtering operations with intermediate resampling of the center of mass function. It is important to note that high frequencies reappear after each iteration because of this resampling; therefore, it is not appropriate to view the complete algorithm as a cascade of Butterworth filters.

It is natural to ask whether a true filter could replace the single MGS iteration. The answer is: in general, no. The reason is that the fixed boundary conditions cannot be modified, and a true filter would inevitably modify them. One notable exception is when the active contour is closed — i.e., $x(0) = x(1), y(0) = y(1)$. In this case the boundary conditions can be excluded, since the two ends of the active contour are tied together, and (DVP1) can be solved by performing a series of actual filtering operations on the center of mass function. Furthermore, the filter can be implemented using the FFT, providing a fast alternative to the MGS iteration.

Since the MGS iteration performs a first step in a Butterworth filter, we are led to the conclusion that if the coordinate functions of $\alpha$ have frequencies outside the cutoff frequency the final curve is guaranteed to be too smooth, while if they have frequencies primarily within the bandwidth of the filter we may be able to reconstruct it. This basic idea motivates the following three sections, starting off with a study of the spectral properties of $\alpha$.

C. Bandwidth of a Thick Curve

Let $\alpha(s) = (x(s), y(s))$ be the constant velocity mapping of $[0, 1]$ onto a curve which has length $L$ and yields a thick curve with thickness $w$. In this section we will derive an upper bound for the bandwidth of the function $\alpha(s)$.

Our first task is to derive two fundamental constraints on the Fourier coefficients of the central contour of a thick curve. Let $g$ be the constant velocity of $\alpha(s)$ — i.e., $g = |\alpha'(s)|$. Since the length of the curve is $L$ it is easy to show that $L = g$, and squaring both sides and integrating over $[0, 1]$ gives

$$L^2 = \int_0^1 \left[ \left( \frac{dx}{ds} \right)^2 + \left( \frac{dy}{ds} \right)^2 \right] ds,$$

which is just $2\pi^2 \sum_{i=1}^{\infty} |F_i|^2$. The first fundamental constraint follows immediately:

$$\sum_{i=1}^{\infty} |F_i|^2 \leq \frac{L^2}{2\pi^2}. \quad (43)$$

The requirement that $C$ has certain thickness and that does not cross itself imposes a constraint on the degree to which it can bend and, hence, to the maximum curvature of $\alpha(s)$. More specifically,
it is not hard to see that the smallest radius of curvature for a thick curve with thickness \( w \) is \( w/2 \), yielding a maximum curvature of

\[
\kappa_{\text{max}} = \frac{2}{w} .
\]  

(44)

In order to relate the curvature of \( \alpha(s) \) to its Fourier coefficients, we start by noting that the unit tangent of \( \alpha(s) \) is \( T(s) = \alpha'(s)/g \). The derivative of \( T(s) \) with respect to arclength is defined as the curvature \( \kappa(s) \) times the unit normal \( N(s) \), but here \( s \) is not arclength. Since the curve has constant velocity, however, the arclength is just \( gs \), and a simple application of the chain rule yields

\[
\kappa(s)N(s) = \frac{\alpha''(s)}{g^2}.
\]

Since \( N(s) \) is a unit vector by definition, we also have

\[
|\alpha''(s)| = g^2|\kappa(s)|,
\]

and it follows from (44) and the fact that \( g = L \) that

\[
|\alpha''(s)| \leq \frac{2L^2}{w}.
\]

(46)

Squaring both sides of (46) and integrating over \([0,1]\) gives

\[
\int_0^1 \left[ (x''(s))^2 + (y''(s))^2 \right] \, ds \leq \int_0^1 \left( \frac{2L^2}{w} \right)^2 \, ds
\]

which, after application of (32) and (35) followed by some rearrangement yields

\[
\sum_{l=1}^\infty l^4 F_l^2 \leq \frac{L^4}{2\pi^4 w^2}.
\]

(48)

This constitutes our second fundamental constraint on the Fourier coefficients of \( \alpha(s) \).

Our second task is use these two constraints to show that \( \alpha(s) \) is bandlimited. In Appendix B we show that if the 3-dB point \( l_0 \) of the series expansion of \( \alpha(s) \) is defined to be the largest \( l \) for which

\[
\frac{F_l^2}{\max_l F_l^2} \geq \frac{1}{2},
\]

(49)

then under certain mild approximations

\[
l_0 \leq \frac{1}{\sqrt{2}} \left( \frac{L}{\pi w} \right)^{\frac{3}{2}}.
\]

(50)

The bound given in (50) is not sharp, so in practice we use a smaller upper bound found using the following heuristic argument. Suppose that only one frequency, say at index \( k \), has energy —
i.e., $F_l = 0$, $l \neq k$. Then it is easy to show that $\alpha(s)$ must be a circle. Ignoring possible overlap, translation, and the initial starting point, the coordinate functions must be of the form

$$\begin{align*}
\pi(s) &= \frac{1}{\kappa} \cos (L\kappa s) \\
y(s) &= \frac{1}{\kappa} \sin (L\kappa s).
\end{align*}$$

Because of the constraint on $\kappa$, the maximum frequency at which these functions can exist is $L\kappa_{\text{max}}/2\pi$. Using (44) we conclude that a practical constraint on the bandlimit of $\alpha(s)$ is provided by

$$l_\alpha \leq \frac{L}{\pi w}. \quad (51)$$

Furthermore, it is clear that the two bounds given by (50) and (51) are approximately equal if $(L/4\pi w)^{1/4} \approx 1$.

D. Oversmoothing and Aliasing

Our active contour algorithm cannot reconstruct $\alpha(s)$ accurately if $l_c < l_\alpha$, since in this case the high frequencies will be filtered out. At best, the result will be an oversmoothed version of $\alpha(s)$. We may, however, select $K_0$ to guarantee that $l_c > l_\alpha$. Accordingly, use of (41) and (51) leads to the requirement that $K_0 < K_{\text{UB}}$ where

$$K_{\text{UB}} = \frac{1}{4\pi^2 l_\alpha^2}. \quad (52)$$

Selecting a small $K_0$ to satisfy $K_0 < K_{\text{UB}}$ forces $l_c$ to be large; therefore, high frequencies will not be filtered out in the MGS step of the active contour algorithm. But the presence of these high frequencies may cause sampling problems. In particular, Nyquist’s theorem specifies that in order to adequately sample $\alpha(s)$ the sampling frequency $l_s$ must be at least $2l_c$. The active contour itself, however, may have higher frequencies present due to higher $l_c$ or to the reintroduction of high frequency components from center of mass sampling. Ideally, the MGS iteration will impose a bandlimit of $l_c$, and therefore $l_s = 2l_c$ provides an adequate sampling frequency. This is not necessarily true, however, since the center of mass resampling process may insert new high frequency components at every stage (and since MGS is just a single iteration of a non-ideal low pass filter). Furthermore, the very act of resampling the center of mass can introduce aliasing that will never be filtered out.

We see now that selecting $K_0$ too small virtually guarantees that some aliasing will take place, and that we can expect to see some odd behavior in the active contours, most particularly in the frequencies just under the MGS cutoff frequency $l_c$. Most commonly, we see this type of behavior near sharp folds in $C$, where splits in the active contour occur (see Section V). These contours ultimately converge to some final solution that matches the Euler equations, and we have shown in our experiments that these represent local minima. Therefore, in this case we have been able to demonstrate the equivalence of the presence of local minima with aliasing in the MGS iterations.
V. EXPERIMENTAL RESULTS

In this section we present a set of experiments with both synthetic and real data, which demonstrate the behavior of the active contour algorithm for different choices of $K_0$ and $N$. We show solutions to (DVP1) that demonstrate nearly optimal performance as well as those that correspond to local minima or are too smooth.

A. Verification of Bounds

In order to discuss the validity of the lower and the upper bounds $K_{LB}$ and $K_{UB}$ derived in Sections III and IV, we present a set of experiments in which the object is an annulus. Specifically, the object is created using a semicircle with radius $R = 60$ and thickness $w = 6$; the neighborhood size is $\rho = 4$ and the number of points in the contour is $N = 50$ (all physical dimensions in this section are in units of pixels). For this annulus we find from (28) that $K_{LB} \approx 10^{-3}$ and from (52) that $K_{UB} = 25 \times 10^{-5}$. Even in this simple case, therefore, there is no $K_0$ that will satisfy both theoretical bounds. We note, however, that $K_{UB}$ is computed on the basis that the thick curve obtains its maximum curvature, which this annulus does not. For this annulus we can compute the true bandwidth, which leads to an effective upper bound of about $1 \times 10^{-1}$. In this case we conclude that choosing $K_0$ from the interval $[1 \times 10^{-3}, 1 \times 10^{-1}]$ will guarantee convexity while simultaneously prohibiting aliasing.

To give an indication of the validity of these bounds we ran a series of experiments for different values of $K_0$. As a measure of error we use the fractional change in energy of the final active contour with respect to the energy of the true curve $\alpha(s)$:

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

Fig. 5 plots $\mathcal{F}$ as a function of $K_0$. The solid curve represents the result when the initial configuration $d_0$ is a semicircle of radius $R - 3$ with uniformly spaced points, while the dotted line results from choosing a $d_0$ that is in $\mathcal{D}$ but contains a split.

Fig. 5 reveals several interesting characteristics of the active contour algorithm. First we see that for $K_0 < \sim 5 \times 10^{-5}$ both curves go sharply positive. This means that $\mathcal{E}$ is much larger than $\mathcal{E}_\alpha$, indicating that the algorithm terminated at a local minimum — i.e., the energy function is not convex for $K_0$ in this region. From $K_0 \approx 5 \times 10^{-5}$ to $K_0 \approx 1.3 \times 10^{-3}$ both curves show a gradual decrease in $\mathcal{F}$, ending up with an actual energy which is only slightly smaller than $\mathcal{E}_\alpha$. In this range $\alpha(s)$ is not the global minimum, but it is not too far away. As $K_0$ grows larger than about $1.3 \times 10^{-3}$, the dotted curve, which corresponds to the initialization containing a split, drops abruptly negative indicating that $\alpha$ is suddenly very far away from the true global minimum. For these values of $K_0$, the minimization procedure leads the starting split configuration outside the boundary of $\mathcal{D}$. In this region the center of mass function is not defined. Therefore, the only forces
that act on the portion of the active contour which lies outside \( D \) are the elasticity ones, and this portion eventually collapses into a line segment, the configuration of minimum energy. This is a case where \( d_0 \) is not within \( D_0 \) and \( D \) itself is not convex. Thus, although \( E \) is convex on \( D \) a simple descent algorithm cannot find the global minimum because of the nonconvexity of \( D \).

We deduce from Fig. 5 that the useful range of \( K_0 \) is \([5 \times 10^{-5}, 1.3 \times 10^{-3}]\), a large range indeed, but one that falls largely below the theoretical interval determined above. It is worth noting however, that for \( K_0 = K_{LB} \) we get an almost perfect reconstruction for both initializations considered in this experiment. We can draw several important conclusions. First, our theoretical lower bound is not very tight. This is no surprise given the loose inequalities used in its derivation. Second, the total range over which the problem is convex and nonaliasing may not be useful due to the nonconvexity of \( D \). Third, the lower bound, although loose, is a good choice for \( K_0 \) for this simple example. Even for more convoluted contours, where one would expect that the lower bound is even looser, we expect that \( K_0 = K_{LB} \) is a good choice that will bring the active contour close to the solution. Finally, for a careful initialization with points uniformly spaced along the active contour, we obtain a fairly broad range of values for \( K_0 \) that yield a good reconstruction, a range which largely coincides with the theoretically determined.

B. Overall Performance

In this section we examine the overall qualitative and quantitative performance of our active contour model on a more realistic object. In this section the thick curve, 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 central contour is

\[
\begin{align*}
\pi(s) &= x_0 - [R + (r_o + a \cos(b \pi s)) \cos(2M \pi s)] \cos \pi s \\
y(s) &= y_0 + [R + (r_o + a \cos(b \pi s)) \cos(2M \pi s)] \sin \pi s,
\end{align*}
\]

where in this set of experiments the parameters were \( r_o = 7, a = 3, b = 2, M = 6 \), \( x_0 = 125 \), and \( y_0 = 50 \). The initial curve is a constant velocity semicircle with the true end-points, as shown in Fig. 6a. The neighborhood size is not large enough to guarantee that \( N(x, y) \cap C \neq \emptyset \) for all points in the curve. Therefore, as described in Section III, we use a larger neighborhood size in the initial iterations and lower its size in increments until the curve lies within \( D \) for neighborhood size \( \rho = 3 \). Thereafter the neighborhood size remains constant.

Fig. 6b shows the MGS solution of (DVP1) for \( K_0 = 8 \times 10^{-7} \) and \( N = 230 \), which together represent empirically “good” parameter choices for this thick curve. Fig. 7 present the results of \( K_0 \) too large and too small for the same \( N \). In Fig. 7a we show the solution when \( K_0 = 25 \times 10^{-6} \), which corresponds to an over-regularized (CVPl). 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. 7b shows the solution when \( K_0 = 10^{-7} \), which specifies an under-regularized 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 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. 7b, especially close to the sharp foldings of the thick curve. 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. The continuous variational problem is the same for both figures, with \( K_0 = 40 \times 10^{-7} \); however, in Fig. 8a the active contour is sampled with 100 points and in Fig. 8b with only 60. The curve in Fig. 8a is a reasonably good result with some evidence of a problem at the two sharp folds nearest the boundary points. In Fig. 8b, however, the final curve contains two strong splits at these sharp folds, which are aliased and never filtered out. The reason these splits did not occur in Fig. 8a 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.

From the results in Figs. 6–8, we can clearly see that the choice of \( K_0 \) and \( N \) strongly affects the ability of our active contour algorithm to converge to the correct curve. As a quantitative measure of performance, we introduce two measures: 1) the total area between the true curve and the final active contour, and 2) the maximum distance in the radial direction between the true curve 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. In both cases we assume that the points in the active contour are connected by line segments, and to make these computations the active contour is upsampled on each of these line segments so that no two points are separated by more than a prespecified small distance.

Our quantitative 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 Figs. 9a and 9b, respectively. A general interpretation of these results is easiest for large \( K_0 \) since in this case the energy function is convex (at least for large \( N \)). For a convex problem, convergence to the global minimum is guaranteed, and if \( N \) is sampled well 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 Figs. 9a and 9b. 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. 9b.

Both Figs. 9a and 9b 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 much larger value, \( K_0 = 80 \times 10^{-7} \), achieves better performance in both cases. Aliasing causes this behavior since splits (combined with linear interpolation between the points) give large area and maximum distance errors. A lower cutoff frequency of the MGS iteration provided by the large \( K \) 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.

Figs. 9c and 9d provide plots of the area and maximum distance errors, respectively, as a function of \( K_0 \) for various \( N \)'s. These plots demonstrate quite dramatically the existence of optimal \( K_0 \)'s. In fact, we readily see from Fig. 9d that the optimal \( K_0 \) decreases with increasing \( N \). The curves for \( N = 230 \) are perhaps the most interesting. Here we see that as \( K_0 \) decreases both performance measures steadily improve, and then suddenly at around \( K_0 = 10 \times 10^{-7} \) both measures sharply rise. The other curves (for different \( N \)'s) reveal a similar behavior on what appears to be 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 very drastic performance threshold. Since, as we have discussed in Sections III and IV, this performance threshold is object-dependent, we would not choose such a low \( K_0 \) if we desired robust performance across a family of different objects.

In the previous examples the MGS algorithm was used; We have run the same experiments using the GD method and the resulting curves are similar, in most of the cases. Hence, MGS is preferred because it is substantially faster.

C. Magnetic Resonance Data

Cross-sectional magnetic resonance (MR) images of the brain cause the cortex to appear as a thick curve with a fairly uniform thickness. In this experiment we model the cortex as a thick curve and apply our active contour algorithm. MR data suffers from both noise and partial volume averaging — i.e., where pixels include more than one tissue type — causing the segmentation of an image into cortex and non-cortex to be very difficult. Although it is possible to proceed with a crude segmentation and rely on the robustness of the active contour algorithm to ignore misclassified points, we choose instead to redefine the mass function altogether. Since this method is not of primary interest in this paper, we provide only a sketch of our approach (see also [33]).

Let \( \lambda_i, i = 1, \ldots, K \), denote the tissue classes present in an image and let \( I(x, y) \) represent the image intensity. We define the mass function to be the posterior probability that \((x, y)\) belongs to
\[ m(x,y) = Pr[(x,y) \in \lambda_1 | I(x,y)] = \frac{\sum_{i=1}^{K} f[I(x,y)|\lambda_i]Pr[\lambda_i]}{f[I(x,y)|\lambda_1]Pr[\lambda_1]} , \] 

where the probability density function \( f[\cdot] \) is assumed to be Gaussian and of the form 

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

After determining the means \( \mu_i \) and variances \( \sigma_i^2 \) for any particular data set through partially interactive statistical methods, our mass function is fully specified. The fact that the range of the mass function no longer simply the set \( \{0,1\} \) but rather the interval \( [0,1] \) does not change the implementation of our active contour algorithm in any significant way. In fact, pixels that are either noisy or partial volume averaged have small masses, and hence, do not significantly influence the deformations of the active contour. This gives the algorithm a robustness to noise and partial volume averaging that a rigid presegmentation step could not do. On the down side, the convergence analysis of Section III and the frequency analysis of Section IV are no longer strictly valid, and can only serve as theoretical guides.

Application of the active contour algorithm to MR data is demonstrated in Fig. 10 using \( K_0 = 16 \times 10^{-6} \) and \( N = 200 \). We measured the thickness of the cortical cross-section to be approximately 6, and accordingly used \( w = 6 \). Use of a larger neighborhood and sequential reduction of its size to \( \rho = 3 \) was used in this example, as in the simulations above. We can see that for the most part, the solution falls very nearly on the central layer of the cortical cross-section. It fails to follow the sharpest foldings, however, and completely avoids any migration into the interior portion of the cortex. One could claim that \( K_0 \) was chosen too large; however the true reason is more subtle. In particular, very sharp foldings occur at the so-called sulci of the brain; this is where the cortex folds deep into the brain cavity. At the sulci the effective thickness of the curve becomes \( 2w \) since two layers of the thick curve are juxtaposed, which allows for the possibility of many configurations that will satisfy the center of mass criterion exactly. Of these configurations, the curve most favored by the tension term in the energy function will be the smoothest curve, which is also the curve closest to the outer surface of the brain.

VI. Summary and Discussion

In this paper we introduced the concept of thick curves and proposed a variational formulation for determining their central contours. The variational formulation comprises two terms, one which causes the solution to conform to the data and the other which encourages a constant velocity mapping. The discrete analog and numerical solution together make our approach an active
contour model in which the curve can be viewed as a physical object with applied forces that cause it to undergo deformations leading it to its final minimum energy configuration. Our main result is an analysis of the conditions under which the active contour will converge to the desired solution. We derived a lower bound on the regularization coefficient that guarantees convexity of the energy function in the special case in which the object is a portion of an annulus. We derived an upper bound on the regularization coefficient for an arbitrary thick curve that guarantees no oversmoothing of the final solution. Finally, we showed through simulations and experiments the validity of our bounds and the performance of the algorithm.

Although the lower bound $K_{LB}$ technically applies only to annuli, one could in principle approximate the local boundary of a thick curve as a portion of an annulus. Thus, with some additional knowledge of the nature of the curvature of a thick curve it is possible to carry through the calculation of an approximate lower bound for an arbitrary curve. We saw in Section V, however, that $K_{LB}$ is not a sharp bound; therefore, we would expect to select $K_0$ less than $K_{LB}$ anyway. Choosing $K_0 < K_{LB}$ does not prohibit convergence to the global minimum. In particular, since it is a loose bound, the problem can be expected to remain convex over some (unknown) range below the bound. Also, as in all continuous but non-convex optimization problems, the global solution will have a region around it in which the energy function is convex, and if the initial curve starts within this region our algorithm will converge to the global minimum. Therefore, our two-stage algorithm, in which the neighborhood size is gradually decreased as the active contour approaches the thick curve, is ultimately an important part of the overall algorithm. Since the energy function itself changes as the neighborhood size decreases this approach is clearly related to such well-known methods as graduated non-convexity [28, 34] and simulated annealing [35]. It is interesting to note, however, that the non-convexity of our problem is introduced by the data, and is not an inherent part of the objective function, as is generally the case in these other methods.

Our upper bound $K_{UB}$ on the regularization coefficient is also not tight since it is based on a worst case analysis. Therefore when we find that $K_{UB} < K_{LB}$, as in the case of the annulus, we need not conclude that no $K_0$ exists that will simultaneously guarantee convexity and nonaliasing. Choosing $K_0$ between these bounds may in fact satisfy our condition; but even if it doesn't, such a selection makes a good practical choice that finds a balances between two undesirable consequences.

Adaptive modification of the regularization coefficient in active contour models has been previously proposed in the literature [36, 10], but without theoretical justification. Our work, while intended strictly for thick curve models, also provides some insights in this area. For example, Section III showed that the eigenvalues of the Hessian matrix of the energy function arise independently from separate locations on the curve. Therefore, if $K_0$ were allowed to be a function of the position on the curve, one can imagine using a local analysis of the thick curve to adaptively select an appropriate $K_0$ at each position which would guarantee global convexity. A careful treatment of this approach would require a reformulation of the problem that treats $K_0$ as a function of $s$. In
summary, a fuller development of both adaptive modification of neighborhood size and of adaptive specification of $K_0$ as a function of curve position promises to yield further insights and improved algorithms.

**APPENDIX A**

In this appendix, we derive formulas for the eigenvalues of the Hessian matrix $H_F$ of $\tilde{\mathcal{E}}_F(d)$. This matrix was shown in section III to be a sparse, banded matrix. Using the polar coordinate system introduced in Section III we can see that

$$\tilde{\mathcal{E}}_F(d) = \frac{1}{N} \sum_{i=1}^{N-1} n_i^2 + \frac{1}{2N} (n_0^2 + n_N^2).$$

(57)

We see that the energy function is considerably simplified in the polar coordinate system. The gradient of $\tilde{\mathcal{E}}_F(d)$ is given by the following equations:

$$\frac{\partial \tilde{\mathcal{E}}_F(d)}{\partial x_i} = \frac{\partial \tilde{\mathcal{E}}_F(d)}{\partial n_i} \frac{\partial n_i}{\partial x_i} = \frac{2}{N} n_i \left( \frac{\partial r_i}{\partial x_i} + \frac{\partial n_i}{\partial \phi_i} \frac{\partial \phi_i}{\partial x_i} \right),$$

(58)

$$\frac{\partial \tilde{\mathcal{E}}_F(d)}{\partial y_i} = \frac{\partial \tilde{\mathcal{E}}_F(d)}{\partial n_i} \frac{\partial n_i}{\partial y_i} = \frac{2}{N} n_i \left( \frac{\partial r_i}{\partial y_i} + \frac{\partial n_i}{\partial \phi_i} \frac{\partial \phi_i}{\partial y_i} \right).$$

(59)

But $n_i$ is only a function of $r_i$, hence, we get

$$\frac{\partial \tilde{\mathcal{E}}_F(d)}{\partial x_i} = \frac{2}{N} n_i n_i' \cos \phi_i,$$

(60)

$$\frac{\partial \tilde{\mathcal{E}}_F(d)}{\partial y_i} = \frac{2}{N} n_i n_i' \sin \phi_i.$$

(61)

Similarly we get

$$\frac{\partial^2 \tilde{\mathcal{E}}_F(d)}{\partial x_i^2} = \frac{2}{N} \left[ (n_i n_i'' + (n_i')^2) \cos^2 \phi_i + \frac{n_i n_i'}{r_i} \sin^2 \phi_i \right],$$

(62)

$$\frac{\partial^2 \tilde{\mathcal{E}}_F(d)}{\partial x_i \partial y_i} = \frac{2}{N} \left[ (n_i n_i'' + (n_i')^2) - \frac{n_i n_i'}{r_i} \right] \cos \phi_i \sin \phi_i,$$

(63)

$$\frac{\partial^2 \tilde{\mathcal{E}}_F(d)}{\partial y_i^2} = \frac{2}{N} \left[ (n_i n_i'' + (n_i')^2) \sin^2 \phi_i + \frac{n_i n_i'}{r_i} \cos \phi_i \right].$$

(64)

Hence, $H_F = [f_{ij}]$ is a sparse matrix, with only three diagonal arrays of non-zero elements:

$$f_{ii} = \frac{2}{N} \left\{ \begin{array}{ll} (n_i n_i'' + (n_i')^2) \cos^2 \phi_i + \frac{n_i n_i'}{r_i} \sin^2 \phi_i, & i < N \\ (n_i n_i'' + (n_i')^2) \sin^2 \phi_i + \frac{n_i n_i'}{r_i} \cos \phi_i, & i \geq N \\ 0, & \text{otherwise} \end{array} \right. $$

(65)

$$f_{ij} = \frac{2}{N} \left\{ \begin{array}{ll} \left[ (n_i n_j'' + (n_j')^2) - \frac{n_j n_i'}{r_i} \right] \cos \phi_i \sin \phi_j, & |i - j| = N - 1 \\ 0, & \text{otherwise} \end{array} \right.$$  

(66)
Using these formulas, we can now calculate the eigenvalues of the \((N - 1) \times 2 \times 2\) matrices \(D_i\), which, as it was shown in section III, coincide with the eigenvalues of \(H_F\). It is easy to show using (26) that the eigenvalues of \(D_i\) are the solution to the following quadratic equation:

\[
\lambda^2 - (\sigma_i + \psi_i)\lambda + \sigma_i\psi_i - \gamma_i^2 = 0
\]  

(67)

Introducing the formulas of (66) and omitting the details, we finally conclude that the solutions to (67) are given by

\[
\lambda_{i_1} = \frac{2}{N}\left(n_i n_i'' + (n_i')^2\right), \quad \lambda_{i_2} = \frac{2 n_i n_i'}{N r_i}.
\]  

(68a, 68b)

APPENDIX B

In this appendix we prove Equation (50). Upon defining \(G_i^2 = I_i^2 F_i^2\) Equations (43) and (48) can be written as

\[
\sum_{l=1}^{\infty} G_i^2 = \frac{L^2}{2\pi^2} 
\]  

(69a)

\[
\sum_{l=1}^{\infty} l^2 G_i^2 \leq \frac{L^4}{2\pi^4 w^2}.
\]  

(69b)

Solving (69a) for \(G_{i_0}^2\) and substituting the result into (69b) yields

\[
l_0^2 C^2 - \sum_{l \neq l_o} \beta_l (C^2 - l^2)
\]  

(70)

where \(\beta_l = l^2 G_l^2 / G_{i_0}^2\) and \(C = L / \pi w\).

The bandwidth \(l_0\) is bounded by the inequality in (70), and the largest possible \(l_0\) (over all possible curves) is found by choosing the collection of \(\beta_l\)'s that maximize the right hand side of (70). It is straightforward to show using (49) that the \(\beta_l\)'s are bounded by \(\beta_l \leq 2l^2\). Therefore, the right hand side of (70) is maximized by choosing

\[
\beta_l = \begin{cases} 
0 & l > C, l \neq l_o \\
2l^2 & l \leq C, l \neq l_o,
\end{cases}
\]  

(71)

and Equation (70) now becomes

\[
l_0^2 (C^2 - \sum_{l=1}^{C} l^2 - 2\sum_{l=1}^{C} l^4.
\]

(72)

Although exact expressions for the partial sums appearing in (72) are known, a simpler approximative expression can be found by assuming that \(L \gg w\). In particular, this assumption implies
that $C \gg 1$, which gives the following approximations

$$\sum_{l=1}^{C} l^2 \approx \frac{C^3}{3} \quad (73a)$$

$$\sum_{l=1}^{C} l^4 \approx \frac{C^5}{5} \quad (73b)$$

and (72) becomes

$$\left( l_0^2 \right)^3 - C^2 \left( l_0^2 \right) - \frac{4}{15} C^5 \leq 0. \quad (74)$$

This expression has the roots

$$l_0^2 = \frac{1}{2} C^2 \pm \frac{1}{2} \sqrt{C^4 + \frac{16}{15} C^5}, \quad (75)$$

which, since $C$ is assumed to be large, are given approximately by $\pm \frac{1}{2} C^\frac{5}{2}$. Hence, we have the requirement that

$$-\frac{1}{2} C^\frac{5}{2} \leq l_0^2 \leq \frac{1}{2} C^\frac{5}{2}, \quad (76)$$

which immediately implies (50).

**References**


Figure 1: A thick curve $C$ is defined by its central layer $\alpha(s)$ and its thickness $w$. The point $x$ is close enough to the curve so that its neighborhood $N(x)$ intersects $C$, and its center of mass function $c(x)$ is defined.
Figure 2: Two locally optimal curves with energies (a) 1.4 and (b) 7.1. The curve in (b) represents a local minimum.
Figure 3: Two possible geometries for the intersection of a small neighborhood with an annulus.
Figure 4: Plots showing (a) $\frac{N}{2} \lambda_{11}$ and (b) $\frac{N}{2} \lambda_{22}$. 
Figure 5: The fractional change in energy, $\mathcal{F}$, as a function of the regularization constant $K_0$. 
Figure 6: (a) Initial configuration of the active contour. (b) Final curve for a good parameter selection.
Figure 7: Poor selection of $K_0$: (a) $K_0$ too large (oversmoothing); (b) $K_0$ too small (introducing splits).
Figure 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).
Figure 9: Error Plots: (a) Area error measure, as a function of $N$, 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 choices of $N$. (d) Maximum distance error measure, as a function of $K_0$, for different choices of $N$. 
Figure 9: Error Plots: (a) Area error measure, as a function of $N$, 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 choices of $N$. (d) Maximum distance error measure, as a function of $K_0$, for different choices of $N$. 
Figure 10: A typical MR image of a cross-section of the brain and the reconstruction of the visible part of the cortex with our active contour algorithm.
APPENDIX C

Assume that locally the thick contour can be modelled by a convex thick arc, as in Fig. 3a. Let \( r \) be the radial coordinate of the point \( A \), which is the center of the circular neighborhood, in the polar coordinate system located at the center of the circular arc. Let \( R \) be the radius of curvature of the outermost arc of the thick arc, which is closer to the point \( A \). Let also \( \rho \) be the neighborhood radius. From Fig. 3a we see that

\[
R^2 - y^2 = \rho^2 - (r - y)^2
\]

which gives

\[
y = \frac{R^2 + r^2 - \rho^2}{2r}
\]

We define the following functions:

\[
f_1(r) = \int_\tau^\rho 2z \sqrt{\rho^2 - (r - z)^2} \, dz \tag{77}
\]

\[
f_2(r) = \int_y^R 2z \sqrt{R^2 - z^2} \, dz \tag{78}
\]

\[
g_1(r) = \int_\tau^\rho 2\rho^2 - (\tau - z)^2 \, dz \tag{79}
\]

\[
g_2(r) = \int_y^R 2\sqrt{R^2 - z^2} \, dz \tag{80}
\]

We can see from the geometry of the problem that

\[
n(r) = \frac{f_1(r) + f_2(r)}{g_1(r) + g_2(r)} - r
\]

We will explicitly calculate these integrals, using the following facts:

\[
\int \sqrt{a^2 - z^2} \, dz = \frac{1}{2} \left( z\sqrt{a^2 - z^2} + a^2 \arcsin \left( \frac{z}{a} \right) \right) \tag{81}
\]

\[
\int z\sqrt{a^2 - z^2} \, dz = \frac{1}{3} \left( a^2 - z^2 \right)^{\frac{3}{2}}. \tag{82}
\]

Carrying the integrations through yields

\[
f_1(r) = r \left[ \frac{\pi}{2} \rho^2 - (r - y)\sqrt{\rho^2 - (r - y)^2} - \rho^2 \arcsin \left( \frac{r - y}{\rho} \right) \right] - \frac{2}{3} (\rho^2 - (r - y)^2)^{\frac{3}{2}} \tag{83}
\]

\[
g_1(r) = \frac{\pi}{2} \rho^2 - (r - y)\sqrt{\rho^2 - (r - y)^2} - \rho^2 \arcsin \left( \frac{r - y}{\rho} \right) \tag{84}
\]

\[
f_2(r) = \frac{2}{3} (R^2 - y^2)^{\frac{3}{2}} \tag{85}
\]

\[
g_2(r) = \frac{\pi}{2} R^2 - y\sqrt{R^2 - y^2} - R^2 \arcsin \left( \frac{y}{R} \right) \tag{86}
\]
The range of \( r \) is \( (R - \rho, R + \rho) \) in the above equations.

Up to now we have assumed that the function \( n(\cdot) \) is only a function of \( r \), that is parametrized by \( R, \rho \). In order to be able to plot it, we rewrite it as a function of three variables, namely \( r, R, \rho \):

\[
n(r, R, \rho) = \frac{f_2(r, R, \rho) + f_1(r, R, \rho)}{g_2(r, R, \rho) + g_1(r, R, \rho)} - r
\]

where the formulas giving \( f_1, f_2, g_1, g_2 \) are the same as before. From these formulas it is easy to see that

\[
n(r, R, \rho) = \rho n_0\left(\frac{r}{\rho}, \frac{R}{\rho}, 1\right) = \rho n_0(\rho, R')
\]

(87)

where

\[
r' = \frac{r}{\rho}
\]

(88)

\[
R' = \frac{R}{\rho}
\]

(89)

\[
n_0(\rho, R') = n_0\left(\frac{r}{\rho}, \frac{R}{\rho}, 1\right).
\]

(90)

Now we can see that

\[
\frac{\partial n(r, R, \rho)}{\partial r} = \rho \frac{\partial n_0(\rho, R')}{\partial r} \frac{\partial r'}{\partial r} = \rho \frac{\partial n_0(\rho, R')}{\partial r} \frac{1}{\rho} = \frac{\partial n_0(\rho, R')}{\partial r}
\]

\[
\frac{\partial^2 n(r, R, \rho)}{\partial r^2} = \frac{\partial}{\partial r}\left(\frac{\partial n(r, R, \rho)}{\partial r}\right) = \frac{\partial}{\partial r}\left(\frac{\partial n_0(\rho, R')}{\partial r}\right) = \frac{\partial^2 n_0(\rho, R')}{\partial (\rho')^2} \frac{1}{\rho}
\]

(91)

Hence, we can consider the function \( n_0(\cdot) \) as a function of \( \rho ' \) parametrized by \( R' \) and denote its partial derivative with respect to \( \rho ' \) by \( n_0'(\rho ') \). In this case we get

\[
n(\rho) n'(\rho) = \rho n_0(\rho) n'_0(\rho) = \frac{n_0(\rho) n'_0(\rho)}{\rho r'}
\]

\[
n(\rho) n''(\rho) + (n'(\rho))^2 = \rho \frac{1}{\rho} n''(\rho) + (n_0'(\rho))^2 = n_0(\rho) n''(\rho) + (n_0'(\rho))^2.
\]

Let us assume now that the thick contour is locally a concave circular arc, as in Fig. 3b. We see that

\[
R^2 - y^2 = \rho^2 - (r - y)^2
\]
which yields

\[ y = \frac{R^2 + r^2 - \rho^2}{2r} \]

We now define the following functions:

\[ f_1(r) = \int_y^R 2z\sqrt{R^2 - z^2} \, dz \]

\[ f_2(r) = \int_y^R 2z\sqrt{\rho^2 - (z - r)^2} \, dz \]

\[ g_1(r) = \int_y^R 2\sqrt{R^2 - z^2} \, dz \]

\[ g_2(r) = \int_y^R 2\sqrt{\rho^2 - (z - r)^2} \, dz \]

It is clear from Fig. 3b that

\[ n(r) = \frac{f_2(r) - f_1(r)}{g_2(r) - g_1(r)} - r \]

Carrying the integrations through, we get

\[ f_1(r) = \frac{2}{3} \left( R^2 - y^2 \right)^{\frac{3}{2}} \]

\[ f_2(r) = r \left[ \frac{\pi}{2} \rho^2 - (y - r)\sqrt{\rho^2 - (y - r)^2} - \rho^2 \arcsin \left( \frac{y - r}{\rho} \right) \right] + \frac{2}{3} \left( \rho^2 - (y - r)^2 \right)^{\frac{3}{2}} \]

\[ g_1(r) = \frac{\pi}{2} R^2 - y\sqrt{R^2 - y^2} - R^2 \arcsin \left( \frac{y}{R} \right) \]

\[ g_2(r) = \frac{\pi}{2} \rho^2 - (y - r)\sqrt{\rho^2 - (y - r)^2} - \rho^2 \arcsin \left( \frac{y - r}{\rho} \right) \]

**APPENDIX D**

In this appendix we show that there is always a convex subset \( D_0 \) in \( D \) containing the global minimum. Let \( d^* = [x_1^*, \ldots, x_{N-1}^*, y_1^*, \ldots, y_{N-1}^*]^T \) be a global minimum of \( \mathcal{E}(d) \) in \( D \), and let \( S_i^\circ \) be a convex set in \( C \oplus D_\rho \) containing \((x_i^*, y_i^*)\).

**Lemma 3** If \( d_1 = [x_1^1, \ldots, x_{N-1}^1, y_1^1, \ldots, y_{N-1}^1]^T \) where \((x_i^1, y_i^1) \in S_i^\circ \) \( i = 1, \ldots, N - 1 \) and \( d_2 = [x_1^2, \ldots, x_{N-1}^2, y_1^2, \ldots, y_{N-1}^2]^T \) where \((x_i^2, y_i^2) \in S_i^\circ \) \( i = 1, \ldots, N - 1 \) then \( d = ad_1 + (1-a)d_2 \) is in \( D \) whenever \( 0 \leq a \leq 1 \).

**Proof** \( d = ad_1 + (1-a)d_2 \) \( 0 \leq a \leq 1 \) forms a convex combination of the respective elements of \( d_1 \) and \( d_2 \). Since these elements comprise points in \( S_i^\circ \) \( i = 1, \ldots, N - 1 \), which are convex sets in \( C \oplus D_\rho \), \( d \) is in \( D \) also.

The following lemma shows the existence of convex sets \( S_i^\circ \).
Lemma 4 At least one $S_i^|$ exists for each $i = 1, \ldots, N - 1$.

Proof Let $q_i$ be any point in $N(x_i^|) \cap C$ not equal to $x_i^|$. Since $q_i$ is in the neighborhood of $N(x_i^|)$, the distance between the two points $d(q_i, x_i^|)$ is no greater than $\rho$. Let $S_i$ be the set of points $p$ such that $d(p, x_i^|) + d(p, q_i) \leq \rho$. Clearly $S_i$ is an ellipse and by construction the $\rho$-neighborhood of any point within $S_i$ intersects both $q$ and $x_i^|$. Therefore $S_i$ is a convex set in $C \oplus D_{\rho}$ containing $x_i^|$ and $S_i^| = S_i$ satisfies the required conditions.

Putting these two lemmas together leads immediately to the following proposition:

Proposition 5 There exists a convex set $D_0$ satifying $D_0 \subset D$ and $d^\circ \in D_0$.

Proof Follows immediately from previous two lemmas.