Convexity Analysis of Active Contour Problems

Christos Davatzikos

Department of Radiology
Johns Hopkins University
Baltimore, MD 21287
hristos@welchlink.welch.jhu.edu

Jerry L. Prince

Department of Electrical and Computer Engineering
Johns Hopkins University
Baltimore, MD 21218
jprince@jhu.edu

Abstract

A general active contour formulation is considered and a convexity analysis of its energy function is presented. Conditions under which this formulation has a unique solution are derived; these conditions involve both the active contour energy potential and the regularization parameters. This analysis is then applied to four particular active contour formulations, revealing important characteristics of their convexity, and suggesting that external potentials involving center of mass computations may be better behaved than the ususal potentials based on image gradients. Most importantly, our analysis provides an explanation for the poor convergence behavior at concave boundaries and suggests an alternate algorithm for approaching these types of boundaries.

I. Introduction

Active contours, originally described by Kass, Witkin, and Terzopoulos [1], have been successfully used in a wide variety of applications. Their main advantage is that they are topologically isomorphic to the features they seek, namely object boundaries. As a result, no edge linking is required, and they are robust to low contrast, noise, and gaps or spurious branches in boundaries. A main disadvantage is that their convexity properties are poorly understood. Specifically, it has been noted in the past that active contour models are nonconvex, and that solutions are often locally rather than globally optimal solutions, often involving discontinuities or "splits" in the final contour [2]. Our main goal in this paper is to understand the nature of the convergence problem, through a study of the convexity of the active contour energy function. In particular, we consider the following question: under what conditions is the active contour energy function convex? We start by stating a fairly general active contour formulation in the continuum. We then discretize this formulation and derive a convexity condition for the discrete model. This condition involves the regularization parameters of the problem — these prescribe the elasticity and rigidity of the active contour — and local object characteristics such as curvature. The fact that the object itself partly determines the convexity of the formulation is unavoidable, and is partly what accounts for the often-observed excellent convergence in some regions and abysmal convergence in other regions.

To demonstrate the utility of our convexity analysis, we apply it to four active contour formulations differing only in their definition of external energy. Since the primary application of our research is to cortical brain mapping [2], our four models focus on the detection of ribbons (which are good models for the cortex in two-dimensional magnetic resonance images of the brain). This framework includes the usual boundary mapping problem as a special case since the ribbons can be made very thin. Therefore, boundaries resulting from convolutional or morphological edge detection operations can also be studied within this framework. The four active contour formulations include two that are modifications of the potentials commonly used in the "snake" models [1, 3] and two that were introduced in [2, 4, 5] in the context of brain mapping and were subsequently used in [6] for estimating the central axis of tube-like objects.

II. Discrete Active Contour Model

An active contour is a curve $\mathbf{x}(s) = (x(s), y(s))$, $s \in [0, 1]$, that moves through image data to minimize an energy $\mathcal{E} = \mathcal{E}_{\rm I} + \mathcal{E}_{\rm E}$, where $\mathcal{E}_{\rm I}$ is the *internal energy* and $\mathcal{E}_{\rm E}$ is the *external energy*, defined as

$$\mathcal{E}_{\rm I} = K_0 \int_0^1 \|\mathbf{x}'(s)\|^2 ds + K_1 \int_0^1 \|\mathbf{x}''(s)\|^2 ds, \quad (1)$$

$$\mathcal{E}_{\rm E} = \int_0^1 P(x(s), y(s))\, ds, \quad (2)$$

where $P(x, y)$ is the *potential function*, and K_0 and K_1 are the elasticity and rigidity parameters of the active contour. It can be shown using calculus of variations that a curve minimizing $\mathcal{E}$ must satisfy the Euler equations

$$-\nabla_{\mathbf{x}} P(\mathbf{x}(s)) + 2K_0 \mathbf{x}''(s) - 2K_1 \mathbf{x}''''(s) = 0. \quad (3)$$

To fully specify the model, we assume that $\mathbf{x}(\cdot)$ and its derivatives are known for $s = 0$ and $s = 1$.

In this paper, we consider a discrete active contour model defined using finite difference approximations of the derivatives in (1). Let the active contour be modeled as a collection of points $\mathbf{x}_i = [x_i, y_i]^{\rm T}$, $i = 0, \ldots, N$, where $x_i = x(i/N)$, $y_i = y(i/N)$. The free values of the discrete curve — i.e., those not fixed by the boundary conditions — are represented by the vector $\mathbf{d} = [\mathbf{x}_1^{\rm T}, \ldots, \mathbf{x}_{N-1}^{\rm T}]^{\rm T}$. Then we obtain the discrete approximation of the energy function $\mathcal{E}$,

$$\tilde{\mathcal{E}}(\mathbf{d}) = \tilde{\mathcal{E}}_{\rm I}(\mathbf{d}) + \tilde{\mathcal{E}}_{\rm E}(\mathbf{d}) \quad (4)$$

1063-6919/96 $5.00 © 1996 IEEE

where the exact expressions for $\mathcal{E}_I$ and $\mathcal{E}_E$ can be found in [4].

The necessary condition for $\mathbf{d}$ to minimize $\tilde{\mathcal{E}}(\mathbf{d})$ is

$$\begin{aligned}\nabla\tilde{\mathcal{E}}(\mathbf{d}) &= K_0N(2\mathbf{A}_1\mathbf{d}-\mathbf{b}_1)+K_1N^3(2\mathbf{A}_2\mathbf{d}-\mathbf{b}_2)+\mathbf{p}\\ &= 0,\end{aligned} \tag{5}$$

where $\mathbf{b}_1$ and $\mathbf{b}_2$ are $2(N-1)$-vectors related to the boundary conditions,

$$\mathbf{p}=\left[\frac{\partial P}{\partial x_1},\frac{\partial P}{\partial y_1},\ldots,\frac{\partial P}{\partial x_{N-1}},\frac{\partial P}{\partial y_{N-1}}\right]^{\mathrm{T}}, \tag{6}$$

and $\mathbf{A}_1$ =diag$\{\mathbf{B}_1\}$ and $\mathbf{A}_2$ =diag$\{\mathbf{B}_2\}$, where $\mathbf{B}_1$ is a symmetric Toeplitz tridiagonal matrix whose first row is $[2,-1,\cdots,0]$ and $\mathbf{B}_2$ is a symmetric tridiagonal matrix (see [4] for details). If $\tilde{\mathcal{E}}(\mathbf{d})$ is strictly convex then (5) is also a sufficient condition, and $\mathbf{d}$ is the unique minimizer of $\tilde{\mathcal{E}}(\mathbf{d})$.

In what follows, the domain where the potential P is defined is denoted by $\mathcal{R}$. The domain $\mathcal{D}$ in which $\mathbf{d}$ is defined is then given by

$$\mathcal{D}=\{\mathbf{e}=[\mathbf{e}_1^{\mathrm{T}},\ldots,\mathbf{e}_{N-1}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{2N-2}\mid \mathbf{e}_i\in\mathcal{R},i=1,\ldots,N-1\}.$$

The domains $\mathcal{R}$ and $\mathcal{D}$ depend on the potential P, which in turn is derived from the image data.

III. Convexity Analysis

The question we address in this section is this: when is $\tilde{\mathcal{E}}(\mathbf{d})$ strictly convex? This question is equivalent to the question: when is a solution satisfying (5) the unique minimizer of $\tilde{\mathcal{E}}(\mathbf{d})$. Practically, if we know that $\tilde{\mathcal{E}}(\mathbf{d})$ is convex then any solution satisfying (5) is also the globally optimal solution.

Our goal in this section is to find a sufficient condition for the strict convexity of $\tilde{\mathcal{E}}(\mathbf{d})$. First, we note that if the Hessian matrix $\mathbf{H}$ of $\tilde{\mathcal{E}}(\mathbf{d})$ is positive definite at each point in $\mathcal{D}$ then $\tilde{\mathcal{E}}(\mathbf{d})$ is strictly convex. Since $\mathbf{H}$ is guaranteed to be positive definite if the eigenvalues of $\mathbf{H}$ are strictly positive, we have that $\tilde{\mathcal{E}}(\mathbf{d})$ is strictly convex if the smallest eigenvalue of $\mathbf{H}$ is greater than zero, i.e.,

$$\lambda_{\min}(\mathbf{H})>0.$$

Now letting $\mathbf{H}_E$ and $\mathbf{H}_I$ denote the Hessian matrices of $\tilde{\mathcal{E}}_E(\mathbf{d})$ and $\tilde{\mathcal{E}}_I(\mathbf{d})$, respectively, we have from (4) that

$$\mathbf{H}=\mathbf{H}_I+\mathbf{H}_E. \tag{7}$$

Using the fact that the smallest eigenvalue of the sum of two symmetric matrices is greater than or equal to the sum of the smallest eigenvalues of the two matrices, we have the following sufficient condition: $\tilde{\mathcal{E}}(\mathbf{d})$ is strictly convex if

$$\lambda_{\min}(\mathbf{H}_I)+\lambda_{\min}(\mathbf{H}_E)>0. \tag{8}$$

In the following sections we find expressions for the minimum eigenvalues of $\mathbf{H}_E$ and $\mathbf{H}_I$.

A. Minimum Eigenvalue of $\mathbf{H}_I$

It can be shown (see also (5)) that the eigenvalues of $\mathbf{H}_I$ coincide with those of the matrix $2K_0N\mathbf{A}_1+2K_1N^3\mathbf{A}_2$. Moreover, since $\mathbf{A}_1$ is block diagonal, its eigenvalues coincide with those of $\mathbf{B}_1$. Similarly, the eigenvalues of $\mathbf{A}_2$ coincide with those of $\mathbf{B}_2$. Therefore, we conclude that

$$\lambda_{\min}(\mathbf{H}_I)\geq 2K_0N\lambda_{\min}(\mathbf{B}_1)+2K_1N^3\lambda_{\min}(\mathbf{B}_2). \tag{9}$$

Since $\mathbf{B}_1$ is Toeplitz tri-diagonal, its eigenvalues can be found through a recursion formula [4], giving the following minimum eigenvalue

$$\lambda_{\min}(\mathbf{B}_1)=2(1-\cos(\pi/N)). \tag{10}$$

The eigenvalues of $\mathbf{B}_2$ satisfy a double recursion formula which does not have an explicit solution. A lower bound can be readily found, however, as follows. It is straightforward to show that

$$\mathbf{B}_2=\mathbf{B}_1^2+\frac{1}{2}\mathbf{z}_1\mathbf{z}_1^{\mathrm{T}}+\frac{1}{2}\mathbf{z}_2\mathbf{z}_2^{\mathrm{T}},$$

where $\mathbf{z}_1=[1,0,\ldots,0]^{\mathrm{T}}$ and $\mathbf{z}_2=[0,\ldots,0,1]^{\mathrm{T}}$. Since $\lambda_{\min}(\mathbf{z}_1\mathbf{z}_1^{\mathrm{T}})=\lambda_{\min}(\mathbf{z}_2\mathbf{z}_2^{\mathrm{T}})=0$, the minimum eigenvalue of $\mathbf{B}_2$ must satisfy $\lambda_{\min}(\mathbf{B}_2)\geq\lambda_{\min}(\mathbf{B}_1^2)$. Since $\mathbf{B}_1$ is positive definite, we have $\lambda_{\min}(\mathbf{B}_1^2)=(\lambda_{\min}(\mathbf{B}_1))^2$ which yields

$$\lambda_{\min}(\mathbf{B}_2)\geq(\lambda_{\min}(\mathbf{B}_1))^2. \tag{11}$$

Substituting (10) and (11) into (9) gives the final result

$$\lambda_{\min}(\mathbf{H}_I)\geq 4K_0N(1-\cos(\pi/N))+8K_1N^3(1-\cos(\pi/N))^2. \tag{12}$$

B. Minimum Eigenvalue of $\mathbf{H}_E$

Taking second partial derivatives of $\tilde{\mathcal{E}}_E(\mathbf{d})$ and using the fact that $\partial P(x_i,y_i)/\partial x_j=\partial P(x_i,y_i)/\partial y_j=0$, for $i\neq j$, it is not hard to see that $\mathbf{H}_E$ is a block diagonal matrix with diagonal entries

$$D_i=\frac{1}{N}\begin{bmatrix}P_{xx}(\mathbf{x}_i) & P_{xy}(\mathbf{x}_i)\\ P_{xy}(\mathbf{x}_i) & P_{yy}(\mathbf{x}_i)\end{bmatrix}\quad i=1,\ldots,N-1,$$

where $\mathbf{x}_i=(x_i,y_i)$. Thus, the eigenvalues of $\mathbf{H}_E$ can be determined by finding the eigenvalues of these 2×2 matrices for $i=1,\ldots N-1$. A direct solution yields

$$\begin{aligned}\lambda_{i1}&=\left(P_{xx}(\mathbf{x}_i)+P_{yy}(\mathbf{x}_i)\right)/2N+\\ &\quad\left(\sqrt{(P_{xx}(\mathbf{x}_i)-P_{yy}(\mathbf{x}_i))^2+4P_{xy}^2(\mathbf{x}_i)}\right)/2N,\end{aligned} \tag{13a}$$

$$\begin{aligned}\lambda_{i2}&=\left(P_{xx}(\mathbf{x}_i)+P_{yy}(\mathbf{x}_i)\right)/2N-\\ &\quad\left(\sqrt{(P_{xx}(\mathbf{x}_i)-P_{yy}(\mathbf{x}_i))^2+4P_{xy}^2(\mathbf{x}_i)}\right)/2N,\end{aligned} \tag{13b}$$

for $i=1,\ldots,N-1$. The minimum eigenvalue is

$$\lambda_{\min}(\mathbf{H}_E)=\min\{\lambda_{i1},\lambda_{i2},\ i=1,\ldots,N-1\}. \tag{14}$$

C. Convexity Condition

Using (12) and (14) in (8) and multiplying both sides by $N/2$ gives the following general convexity condition: $\tilde{\mathcal{E}}(\mathbf{d})$ is strictly convex on $\mathcal{D}$ if

$$\frac{1}{2}\min_{1\le i\le N-1}\{N\lambda_{i1}, N\lambda_{i2}\} + 2K_0N^2(1-\cos(\pi/N)) + 4\,K_1N^4(1-\cos(\pi/N))^2 > 0\,. \quad (15)$$

From (13), we see that the terms $N\lambda_{i1}$ and $N\lambda_{i2}$ are independent of N. We can go a step further to eliminate the dependence of this expression on the specific location of the points $\mathbf{x}_i$, $i = 1, \ldots, N-1$. To do this we define

$$h_1(\mathbf{x}) = \left(P_{xx}(\mathbf{x}) + P_{yy}(\mathbf{x})\right)/4 + \left(\sqrt{(P_{xx}(\mathbf{x}) - P_{yy}(\mathbf{x}))^2 + 4(P_{xy}(\mathbf{x})^2}\right)/4, \quad (16a)$$

$$h_2(\mathbf{x}) = \left(P_{xx}(\mathbf{x}) + P_{yy}(\mathbf{x})\right)/4 - \left(\sqrt{(P_{xx}(\mathbf{x}) - P_{yy}(\mathbf{x}))^2 + 4(P_{xy}(\mathbf{x})^2}\right)/4, \quad (16b)$$

which leads directly to our final convexity condition: the energy function of VP is strictly convex for every active contour in a subset $\mathcal{R}' \subseteq \mathcal{R}$ if

$$A(\mathcal{R}') + 2K_0N^2(1-\cos(\pi/N)) + 4K_1N^4(1-\cos(\pi/N))^2 > 0\,, \quad (17)$$

where

$$A(\mathcal{R}') = \min_{\mathbf{x}\in\mathcal{R}'}\{h_1(\mathbf{x}), h_2(\mathbf{x})\}\,. \quad (18)$$

The use of a subset $\mathcal{R}'$ of $\mathcal{R}$ allows for the possibility that we may only be interested in convexity on a smaller set than the entire domain of definition of the potential. The remaining issue is how to determine $A(\mathcal{R}')$.

D. Approximations

From (17), (18), and (16) it is clear that convexity of $\tilde{\mathcal{E}}(\mathbf{d})$ depends on the regularization coefficients K_0 and K_1 and on the potential $P(\mathbf{x})$. The potential, in turn, depends on the underlying image. To know how to choose K_0 and K_1 to maintain convexity, it is possible, in principle, to calculate $A(\mathcal{R}')$ for a given image or region within an image. In this and the following sections, however, we seek a more general understanding of the relationships between convexity of $\tilde{\mathcal{E}}(\mathbf{d})$ and object shape through a series of approximations.

In order to simplify our analysis, we first observe that as $N \longrightarrow \infty$, (17) becomes independent from N. In particular, using a Taylor series expansion it is readily shown that

$$\lim_{N\to\infty} 2N^2(1-\cos(\pi/N)) = \pi^2\,.$$

This sequence converges fairly fast, so that for any practical value for N ($N > 30$), it is a very good approximation to use the limit in place of the expression. Using this approximation, (17) becomes

$$A(\mathcal{R}') + K_0\pi^2 + K_1\pi^4 > 0\,. \quad (19)$$

As shown in Fig. 1, this condition gives a region in K_0-K_1 space which guarantees convexity of $\tilde{\mathcal{E}}(\mathbf{d})$.

We now seek to find simplified expressions for $h_1(\cdot)$ and $h_2(\cdot)$, which determine $A(\mathcal{R}')$, by expressing these functions in a local coordinate system described next. Consider the curve of constant potential, the *isopotential curve*, passing through a point $\mathbf{x} \in \mathcal{R}$, and assume (at first) that its curvature at $\mathbf{x}$ is nonzero. We can then define a local polar coordinate system whose origin coincides with the center of curvature of this curve, as shown in Fig. 2. We now view the potential as a function of r and ϕ, the radius and angle in the local coordinate system, and seek expressions for P_{xx}, P_{xy}, and P_{yy} at $\mathbf{x}$.

First we note that because of the way we defined the local coordinate system it can be shown that [4] $P_\phi = P_{\phi\phi} = 0$. Using this fact and the chain rule, we can express the derivatives of the potential P in the polar coordinate system. After some algebra we arrive at the following expressions for h_1 and h_2:

$$h_1(\mathbf{x}) = \left(P_{rr}(\mathbf{x}) + P_r(\mathbf{x})/r(\mathbf{x})\right)/4 + \left(\sqrt{\left(P_{rr}(\mathbf{x}) - P_r(\mathbf{x})/r(\mathbf{x})\right)^2 + 4P_{r\phi}^2(\mathbf{x})/r^2(\mathbf{x})}\right)/4, \quad (20a)$$

$$h_2(\mathbf{x}) = \left(P_{rr}(\mathbf{x}) + P_r(\mathbf{x})/r(\mathbf{x})\right)/4 - \left(\sqrt{\left(P_{rr}(\mathbf{x}) - P_r(\mathbf{x})/r(\mathbf{x})\right)^2 + 4P_{r\phi}^2(\mathbf{x})/r^2(\mathbf{x})}\right)/4, \quad (20b)$$

where $r(\mathbf{x})$ is the radius of curvature of the isopotential curve at $\mathbf{x}$ (see Fig. 2). Note that if we let $r \longrightarrow \infty$ we obtain expressions for h_1 and h_2 for the case where the isopotential curve has zero curvature.

Now we are in position for our second approximation. We note that potentials used for locating boundaries have isopotential curves that are nearly parallel to each other, and that are nearly parallel to the boundary[1]. This is especially true when the boundaries are smooth or the point $\mathbf{x}$ is very near the desired boundary. In this case the term $P_{r\phi}(\mathbf{x})$ in (20) is negligible and we can assume that $P_{r\phi}(\mathbf{x}) \approx 0$, which introduced into (20) yields

$$h_1(\mathbf{x}) \approx \frac{1}{4}\left(P_{rr}(\mathbf{x}) + \frac{P_r(\mathbf{x})}{r(\mathbf{x})} + \left|\left(P_{rr}(\mathbf{x}) - \frac{P_r(\mathbf{x})}{r(\mathbf{x})}\right)\right|\right),$$

$$h_2(\mathbf{x}) \approx \frac{1}{4}\left(P_{rr}(\mathbf{x}) + \frac{P_r(\mathbf{x})}{r(\mathbf{x})} - \left|\left(P_{rr}(\mathbf{x}) - \frac{P_r(\mathbf{x})}{r(\mathbf{x})}\right)\right|\right).$$

By defining $e_1(\mathbf{x}) = P_{rr}(\mathbf{x})/2$ and $e_2(\mathbf{x}) = P_r(\mathbf{x})/2r(\mathbf{x})$ it is straightforward to show that

$$\min_{\mathbf{x}\in\mathcal{R}'}\{h_1(\mathbf{x}), h_2(\mathbf{x})\} = \min_{\mathbf{x}\in\mathcal{R}'}\{e_1(\mathbf{x}), e_2(\mathbf{x})\}\,.$$

Therefore, to good approximation $A(\mathcal{R}')$ is given by

$$A(\mathcal{R}') \approx \min_{\mathbf{x}\in\mathcal{R}'}\{e_1(\mathbf{x}), e_2(\mathbf{x})\}\,. \quad (21)$$

We have now arrived to the main conclusion of this section: Under the approximations described above, the

[1]The value of the potential along each isopotential curve increases with the distance from the underlying boundary.

analysis of the convexity of VP reduces to the study of the functions $e_1(\cdot)$ and $e_2(\cdot)$. We use this result in the following section, in which we study the convexity of four particular active contour potentials, and we show that under certain assumptions, $e_1(\cdot)$ and $e_2(\cdot)$ are functions of only one variable, the radius of curvature r of the isopotential curve at $\mathbf{x}$, which greatly simplifies the convexity analysis.

IV. Specific Examples of the Potential

In this section we apply the convexity analysis of Section III to four particular active contour potentials, which are designed to find and parameterize boundaries. We derive explicit expressions for $A(\cdot)$ in (21) for all four potentials and, based on these expressions, we analyze the condition in (19). This analysis reveals important correlations between the uniqueness of the solution of VP, the selection of its regularization parameters, the initialization of the active contour relative to the boundary, and the geometric characteristics of the boundary.

A. Ribbons

In order to analyze all four potentials in a unified framework, we use an idealized model for object boundaries: a *ribbon* of thickness w (see Fig. 3). Convolutional and morphological edge detection operations often produce edge maps resembling thin ribbons, whose thickness is directly related to the size of the kernel of the edge-detection operator. We denote $\mathcal{C}$ the collection of points belonging to a ribbon, and we define the mass of the ribbon to be equal to $1/w$ if $\mathbf{x} \in \mathcal{C}$ and zero otherwise. The constant $1/w$ is a normalization factor that makes the total mass of the ribbon equal to its length L. We note that, as w approaches zero, m approaches a Dirac function at $\mathcal{C}$, which defines an ideal boundary. We also denote $\boldsymbol{\alpha}(s)$, $s \in [0, 1]$, a parameterization of the spine of a ribbon; this is the desired solution to VP. Finally, we denote $\boldsymbol{\alpha}_1(s)$ and $\boldsymbol{\alpha}_2(s)$ the two outer curves defining the ribbon (Fig. 3).

B. Ribbon Potentials

In order to use the variational framework of VP to find $\boldsymbol{\alpha}(s)$, we consider four different potentials, which are all based on the geometry of ribbons and are designed to take their minimum value in the vicinity of the spine of a ribbon. The first two, denoted P_1 and P_2, were introduced in [2]. The last two, denoted P_3 and P_4, are variants of the potentials commonly used in snake models [1, 3]. In order to define P_1 and P_2, we first introduce the *center of mass function* $\mathbf{c}(\mathbf{x})$, which is the center of the mass of the ribbon included in a circular neighborhood $\mathcal{N}(\mathbf{x})$ centered at $\mathbf{x}$ (see Fig. 3). We notice that, for sufficiently small neighborhood radius ρ, for any point $\mathbf{x}$ on the spine of the ribbon,

$$\mathbf{c}(\mathbf{x}) \approx \mathbf{x}\,. \tag{22}$$

Accordingly, we define P_1 and P_2 to be minimal on points which coincide with the center of the mass of the ribbon present in their neighborhood:

$$P_1(\mathbf{x}) = \|\mathbf{x} - \mathbf{c}(\mathbf{x})\|^2\,, \tag{23}$$

$$P_2(\mathbf{x}) = 2\int_{\mathcal{L}} \|\mathbf{g} - \mathbf{c}(\mathbf{g})\|\,dl\,, \tag{24}$$

where $\mathcal{L}$ is the line of flow[2] connecting an image point $\mathbf{x}$ with the the minimum potential curve. The minimum potential curve for P_1 and P_2 is the collection of points satisfying the following condition:

$$\mathbf{x} = \mathbf{c}(\mathbf{x})\,. \tag{25}$$

We note that if the diameter of the disk $\mathcal{N}(\cdot)$ is much larger than the width of the ribbon, then (22) is not necessarily satisfied on the spine of a ribbon. Therefore, in the implementation of an active contour algorithm using P_1 or P_2, the neighborhood size should gradually decrease as the active contour approaches the spine [4]. For the purposes of the convexity analysis, we will assume that ρ is sufficiently small so that (22) is satisfied on the spine.

From (23) and (24) we easily see that the potentials P_1 and P_2 are minimal and equal to zero close to the spine of the ribbon, and precisely on the points satisfying (25), and gradually increase with increasing distance from it. Accordingly, the external force field induced by both P_1 or P_2 is a restoring force field tending to deform an active contour toward the spine.

In addition to P_1 and P_2, we also consider the following two potentials:

$$P_3(\mathbf{x}) = -\|m(\mathbf{x}) * * g(\mathbf{x})\|\,, \tag{26}$$

$$P_4(\mathbf{x}) = -\|m(\mathbf{x}) * * g(\mathbf{x})\|^2\,. \tag{27}$$

Here, $g(\mathbf{x})$ is the impulse response of a low-pass filter and is defined to be unity if $\mathbf{x} \in \mathcal{N}(\mathbf{0})$ and zero otherwise. These potentials are similar to the usual snake potentials [1, 3], but differ in the following respects:

1. P_3 and P_4 use the binary mass function $m(\cdot)$ instead of the magnitude of the image intensity gradient which is customary in snake models. A binary mass can be readily obtained by thresholding the edge map obtained through the image gradient.

2. P_3 and P_4 use the "disk" function $g(\cdot)$, instead of a Gaussian smoothing function.

These modifications are introduced here in order to simplify the convexity analysis and allow direct comparisons between the four models.

We will assume throughout our development that P_1, P_2, P_3, and P_4 are defined at $\mathbf{x}$, if the disk $\mathcal{N}(\mathbf{x})$ intersects both $\boldsymbol{\alpha}_1(s)$ and $\boldsymbol{\alpha}_2(s)$ (see Fig. 3). Accordingly, the domain of definition of these potentials is

$$\mathcal{R} = \left\{\mathbf{x} \in \mathbb{R}^2 : \{\mathcal{N}(\mathbf{x}) \cap \boldsymbol{\alpha}_1(s) \neq \emptyset\} \text{ and } \{\mathcal{N}(\mathbf{x}) \cap \boldsymbol{\alpha}_2(s) \neq \emptyset\}\right\}. \tag{28}$$

We note here that outside $\mathcal{R}$, the energy function $\mathcal{E}$ is equal to $\mathcal{E}_I$, and therefore the active contour contracts under the influence of its internal elastic forces, until it reaches the vicinity of the ribbon. Since in that region $\mathcal{E}$ is obviously convex, we will restrict our convexity analysis to $\mathcal{R}$.

C. Convexity of Ribbon Models

Having defined the four potentials, we now turn our attention to $A(\cdot)$ and the convexity condition in (19). In order to obtain explicit expressions for $A(\cdot)$, we need to

[2]The line of flow is the trajectory of a particle under the influence of a potential gradient.

make certain assumptions about the shape of the underlying boundary. In particular, we approximate the intersections of the ribbon with neighborhoods of active contour points by arcs of radius R and thickness w, as shown in Fig. 4. Under this approximation, $P_{r\phi}$ is exactly zero, and $A(\cdot)$ is exactly given by (21). Moreover, $e_1(\cdot)$ and $e_2(\cdot)$ are functions only of the radial coordinate r. Therefore, the convexity analysis reduces to the study of the one-dimensional functions $e_1(r)$ and $e_2(r)$.

In [4] expressions for $e_1(r)$ and $e_2(r)$ were derived for all four potentials P_1, P_2, P_3, and P_4. These expressions depend on the curvature of the ribbon, the ribbon width, and on the neighborhood radius, which makes the display and interpretation of them a fairly difficult task. One way of displaying this information is shown in Figs. 5 and 6, which plot $e_1(r)$ and $e_2(r)$ against $(r - R)/\rho$, for different ratios R/ρ. Several conclusions can be drawn from these figures and the convexity condition in (19):

1. The function $e_1(r)$ (Fig. 5) is positive for $r \in [R+w/2-\rho, R-w/2+\rho]$. Therefore, it is positive everywhere in the domain $\mathcal{R}$ and its minimum value in $\mathcal{R}$ is positive. Given (19) and (21), this implies that $e_1(r)$ imposes no restriction on selecting a pair (K_0, K_1) which yields a convex function.

2. The function $e_2(r)$ is always positive for $r \geq R$ (positive horizontal axis in Fig. 6). Therefore, both $e_1(r)$ and $e_2(r)$ are non-negative at all points in $\mathcal{R}$ for which

$$r(x, y) \geq R(x, y), \tag{29}$$

where $R(x, y)$ is the radius of curvature of the spine in $\mathcal{N}(x, y)$, and $r(x, y)$ is the radius of curvature of the isopotential curve passing through (x, y). Let $\mathcal{R}^+$ be the collection of points satisfying (29), which can be shown [4] to be a connected subdomain of $\mathcal{R}$. Then the convexity condition in (19) is satisfied everywhere in $\mathcal{R}^+$ for all (K_0, K_1) such that $K_0 > 0, K_1 > 0$. This region of convexity is shown schematically in Fig. 7.

The implications of this result are important. In particular, an optimization method seeking a solution to the Euler equations of VP within the subdomain of convexity $\mathcal{R}^+$ is guaranteed to converge to the global minimum in $\mathcal{R}^+$. A similar result was used in [4] as the basis for an adaptive active contour algorithm which restricted the active contour in $\mathcal{R}^+$, avoiding, therefore, local minima.

3. For $r < R$ (negative hotizontal axis in Fig. 6), e_2 is negative for all potentials, and it is considerably smaller for P_3 and P_4 than for P_1 and P_2. Therefore, the minimum value of e_2 in $\mathcal{R}$ for the P_3 and P_4 formulations is smaller than that of the formulations using P_1 and P_2. Considering (19) and (21) this implies that K_0 and K_1 should be larger for the models using P_3 and P_4. Consequently, the elasticity forces in the P_1 and P_2 formulations can be smaller while still maintaining the convexity of VP. This is important since large elasticity forces often cause oversmoothing of boundaries [2].

4. The plots of Fig. 6 show that $e_2(r)$ decreases in absolute value with the distance from the spine. This implies that as the active contour gets closer to the spine, smaller values of K_0 and K_1 can be used. Similarly, if an active contour is initialized close to the spine, much smaller K_0 and K_1 are required to maintain convexity. A similar result motivated the development of a time-adaptive active contour algorithm in [5], which was shown to improve the ability of the active contour to bypass local minima considerably, by gradually reducing the elasticity parameters.

5. From Fig. 6 we deduce that the absolute value of $e_2(r)$ increases rapidly with increasing curvature of the spine $(1/R)$. This implies that for convoluted ribbons, larger elasticity parameters are required to maintain convexity.

V. Summary

This paper studied a general active contour formulation (VP), and derived conditions for the convexity of its energy function. The derived condition quantitatively expressed the relation between the convexity of the energy function and the selection of the regularization parameters. The results were then applied to four special cases of VP designed to find parametrizations of boundaries, which are modeled as ribbons.

The first main conclusion drawn from this analysis is that in order for the energy function of VP to be convex, K_0 and K_1 should be selected from a specific domain in the K_0–K_1 plane. This region is determined by the potential P of the active contour formulation. The application of this result to four particular potentials in Section IV lead to the second and most important conclusion of our analysis. Specifically, it was shown that, under certain assumptions, there exists a connected domain where the energy function is convex for any pair of positive regularization constants, for all four potentials for boundary parameterization. This implies that an active contour algorithm can avoid local minima by searching for a solution within the domain of convexity. The development of such domain adaptive active contour algorithms [4] is a direction of further research originating from the convexity analysis presented in this paper.

Acknowledgements

The authors would like thank Nick Bryan for his support to this work. This research was partly supported by the NSF grant MIP-9350336 and the NIH grant NICHD P50HD 25806-03.

References

[1] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1:321–331, 1988.

[2] C.A. Davatzikos and J.L. Prince. An active contour model for mapping the cortex. *IEEE Trans. on Medical Imaging*, 14:65–80, 1995.

[3] L.D. Cohen and I. Cohen. Finite-element methods for active contour models and baloons for 2D and 3D images. *IEEE Trans. Pattern Analysis and Machince Intelligence*, 15(11):1131–1147, 1993.

[4] C. Davatzikos. *Model-Based Boundary Mapping with Applications to Medical Imaging*. PhD thesis, Johns Hopkins University, 1994.

[5] C.A. Davatzikos and J.L. Prince. Adaptive active contour algorithms for extracting and mapping thick curves. *Proc. of the IEEE Conf. Comp. Vision and Patt. Recognition*, pages 524–529, 1993.

[6] W. Lawton. Mathematical methods for active geometry. *Annals of Numerical Mathematics, eds. C. Micchelli and H. Said, (CAGD conference, Penang, 1994)*, 1, 1995.

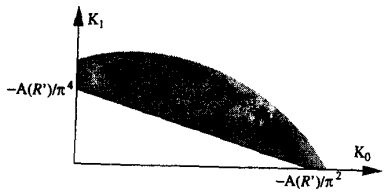

Figure 1. The region of convexity of the active contour problem in the K_0-K_1 plane.

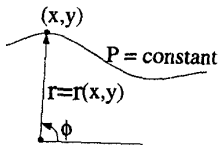

Figure 2. The local polar coordinate system centered at the center of curvature of the isopotential curve through (x, y).

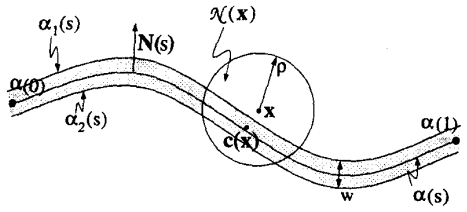

Figure 3. The geometry of a thin ribbon with spine $\alpha(s)$.

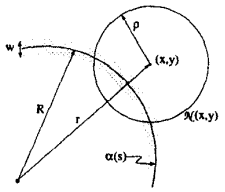

Figure 4. An intersection of $\mathcal{N}(x, y)$ with a ribbon of width w.

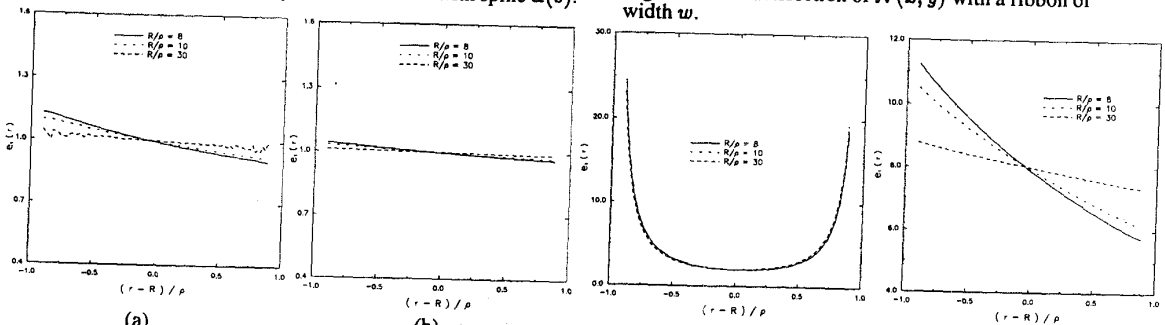

Figure 5. Plots of $e_1(r)$ for different ratios R/ρ and for a ribbon of thickness $w/\rho = 0.01$. (a) P_1, (b) P_2, (c) S_1, and (d) S_2.

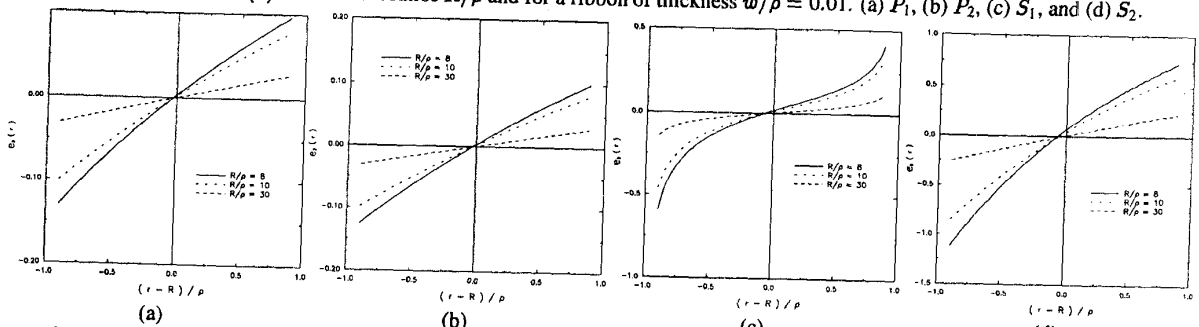

Figure 6. Plots of $e_2(r)$ for different ratios R/ρ and for a ribbon of thickness $w/\rho = 0.01$. (a) P_1, (b) P_2, (c) S_1, and (d) S_2.

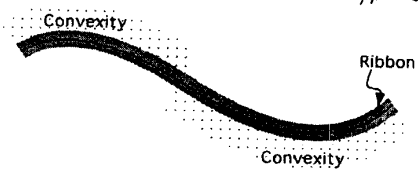

Figure 7 The region of convexity of the thin ribbon active contour problems.