Global Regularizing Flows With Topology Preservation for Active Contours and Polygons

Ganesh Sundaramoorthi and Anthony Yezzi, *Senior Member, IEEE*

Abstract—**Active contour and active polygon models have been used widely for image segmentation. In some applications, the topology of the object(s) to be detected from an image is known *a priori*, despite a complex unknown geometry, and it is important that the active contour or polygon maintain the desired topology. In this work, we construct a novel geometric flow that can be added to image-based evolutions of active contours and polygons in order to preserve the topology of the initial contour or polygon. We emphasize that, unlike other methods for topology preservation, the proposed geometric flow continually adjusts the geometry of the original evolution in a gradual and graceful manner so as to prevent a topology change long before the curve or polygon becomes *close* to topology change. The flow also serves as a *global regularity* term for the evolving contour, and has smoothness properties similar to curvature flow. These properties of gradually adjusting the original flow and global regularization prevent geometrical inaccuracies common with simple discrete topology preservation schemes. The proposed topology preserving geometric flow is the gradient flow arising from an energy that is based on electrostatic principles. The evolution of a single point on the contour depends on all other points of the contour, which is different from traditional curve evolutions in the computer vision literature.**

Index Terms—**Active contours, global flows, global regularization, polygons, topology preservation, variational methods.**

I. INTRODUCTION

SNAKES, or active contours, were introduced by Kass *et al.* [1]. In this work, the objective is to segment an an image by deforming an initial contour towards the boundary of the object of interest. This is done by deforming an initial contour in such a way that it minimizes an energy functional defined on contours. There are two components in this energy: the potential energy, which is small when the contour is aligned to edges of the image, and the internal deformation energy, which is small when the contour is smooth. Both components are contour integrals with respect to a parameter of the contour.

Caselles *et al.* [2] and Kichenassamy *et al.* [3], building on [4] and [5], have considered a geometric energy, i.e., an energy independent of contour parameterization, related to the energy in [1]. The energy is a contour integral with respect to the arc-length of a potential function that is designed to be low near image edges. The resulting minimizing flow consists of two terms: one that attracts the contour towards image edges, and the other that shrinks and smooths the contour. This latter term is a positive function multiplied by curvature flow. A number of interesting properties have been proven about curvature flow [6], [7], and affine invariant curvature flow [8], [9]. Other edge-based active contour models include [10], [11].

Region-based approaches (e.g., [12]–[16]) to active contours, in contrast to edge-based approaches, define energies on contours that incorporate global statistics of the image, rather than image information local to the contour. These energies are integrals over the region enclosed by a contour, rather than over just the contour. These models generally have wider capture ranges for initial contours than edge-based models. To keep the evolving contour smooth, a length shortening term is added to the image-based term. The common features of all these active contour models, either region-based or edge-based, is a data fidelity term, which matches the contour to the image, and a term that penalizes irregularities of the contour.

One way to implement geometric active contour evolutions is using a level-set method, which was introduced by Osher and Sethian [17]. A function whose domain is the domain of the image is evolved so that its zero level-set motion corresponds to the desired contour evolution. When implemented with a *narrowband* technique, a level set method has nearly the same computational cost as other implementations of contour evolutions. Advantages of this approach are independence to reparameterization of the contour, and splitting and merging, i.e., topology changes, of the evolving contour(s) are handled without additional effort.

Although topological flexibility is an advantage of level-set methods in many applications, in some applications, such topology changes are not desirable [18], [19]. In these applications, the topology of the object of interest is known *a priori*, and must be enforced during the evolution. Han *et al.* [18], [19] examine extracting the brain cortex, which is known to be homeomorphic to a sphere, from an image. The method for topology preservation in [18] and [19] assumes that the active contour evolution is implemented using a level-set method. This method simply stops the evolution of the contour when the digital topology determined by the level set function is one pixel away from changing topology. The condition for detecting topology change at a grid point is based on the level set function in a small neighborhood of the grid point (see Section III-B2). One advantage of this method is that the computational speed is negligible when compared to the original function evolution.

Although the method of Han *et al.* guarantees that the resulting segmentation has the correct topology, there are some undesirable features resulting from the way this method is used in [18], [19]. The most notable disadvantage of Han *et al.*'s

method as presented there is that topology preservation is an abrupt, discontinuous motion that is unnatural. Often times, as we will see in Section IV, one obtains an unnatural segmentation, plagued by *geometrical inaccuracies* and is one pixel away from self intersection. The undesirable features resulting from the method of Han *et al.* stem from the fact that the underlying partial differential equation (PDE) guiding the evolving contour does not preserve topology. Thus, the abrupt stopping of the contour in the method of Han *et al.* does not simulate the actual behavior of the underlying PDE. It would be justified to use the method of Han *et al.*, as a numerical scheme, in conjunction with a PDE that actually preserves topology much like the use of numerical schemes that are designed to simulate certain features of PDEs (for example, conservation laws). Thus, in the case that limitations of the discretized PDE prevent topology preservation in the simulation of the PDE, the method of Han *et al.* can be used to prevent a discrete topology change (see Section III-B2).

In the proposed work, we construct a PDE that actually preserves topology. We construct a geometric term that can be added to an existing image-based contour evolution. This novel geometric flow preserves the topology of the contour by continually adjusting the original flow so that the contour gradually moves away from changing topology well before the contour becomes close to changing topology. We will show that the flow we construct *globally regularizes* the evolving contour and may be used as a contour regularizer, which among other properties, ensures smoothness of the contour. Therefore, the usual curvature smoothness term that is used in most active contour works can be replaced with the proposed geometric flow when topology preservation is needed. Geometric inconsistencies in the resulting segmentation that typically arise from using the method of Han *et al.* on a PDE that does not preserve topology are not likely with our new flow because of the global regularizing properties and gradual topology preservation of our flow. Because we use a geometric flow, our method is not restricted to level set methods. We also construct a global regularizing and topology preserving term for the case of evolving polygons [20], [21].

In order to derive a geometric flow for topology preservation in the presence of image-based forces, we observe a necessary property of the flow. In order for a point on the contour to move away from self intersection (i.e., topology change), it must have some "knowledge" of the location of the other points on the contour. Thus, the evolution of a point must depend globally on other points of the contour, as opposed to a local dependence as a curvature regularizer. A contour evolution that depends on each point of the contour was considered by Rochery *et al.* [22], [23]. In that work, an energy that is a double integral around a contour of some function was introduced. The resulting contour evolution depends on a contour integral. A similar use of energies that are double integrals is also used in the work of Kim *et al.* [24]. In this case, however, the energy is a double integral over the region enclosed by the contour rather than the contour. In the proposed work, we introduce an energy that depends on a double integral over the contour, and has natural connections to electrostatics. Recent works that use electrostatics principles for image segmentation are [25] and [26]; however, these works cannot be used for topology preservation.

Recent works that incorporate prior topological knowledge into image segmentation are [27] and [28]. In [27], an energy is constructed to favor nonintersection of distinct connected components of the contour; however, there is nothing in the energy that prevents a single connected component from splitting. In [28], a level set scheme is formulated for preserving topology. This method is different than ours in that it is based on the level set representation of the contour. The key difference is that their level set evolution does not provide global regularization of the contour, which is often necessary in applications, and a major contribution of our work.

A preliminary version of this work appeared in [29], and a full version of this paper with full technical details is [30].

II. VARIATIONAL APPROACH TO TOPOLOGY PRESERVATION

First, we give a description of the active polygons and active contours considered in this paper. An active polygon is a polygon with vertices $v_0, \ldots, v_{n-1}$, where $n \in \mathbb{N} \setminus \{1, 2\}$, in $\mathbb{R}^2$ that moves at each instant of time according to the set of ordinary differential equations

$$v'_k(t) = I_k(t) + \alpha R_k(t), \text{ where } k = 0, 1, \ldots, n-1 \quad (1)$$

where $I_k, R_k : \mathbb{R}^+ \to \mathbb{R}^2$; I_k is the force derived from the image we wish to segment, R_k is a regularizing and topology preserving force that is derived only from the geometry of the polygon, and $\alpha > 0$ is a constant. We assume uniform boundedness of I_k, that is, $\sup_{t \in \mathbb{R}^+, k=0,1,\ldots,n-1} \|I_k(t)\| < +\infty$, where $\| \cdot \|$ is the Euclidean norm of $\cdot$. An active contour C is a twice differentiable curve in $\mathbb{R}^2$ that moves at each instant of time according to the partial differential equation

$$\partial_t C(p, t) = i(p, t)\mathcal{N}(p, t) + \alpha \mathcal{R}(p, t) \quad (2)$$

where $p \in S^1$ denotes a parametrization of the curve C, $S^1 = [0, 1]/\{0, 1\}$ is the unit interval with endpoints identified, $i : S^1 \times \mathbb{R}^+ \to \mathbb{R}$, and $\mathcal{R} : S^1 \times \mathbb{R}^+ \to \mathbb{R}^2$. The image-based force is the term $i\mathcal{N}$, where $\mathcal{N}$ is the unit normal vector to the curve C. The regularizing and topology preserving term of the active contour is $\mathcal{R}$. We make the assumption that i is uniformly bounded, i.e., $\sup_{t \in \mathbb{R}^+, p \in S^1} |i(p, t)| < +\infty$. Our assumptions of uniform boundedness are not too restrictive and are typically true in practical applications. For example, the data terms of Mumford–Shah [12], Chan–Vese [15], and other region-based curve evolutions satisfy these conditions when the image is bounded, which is true for digital images.

Our approach to derive R_k and $\mathcal{R}$ so that (1) and (2) preserve the topology of the initial configurations is to minimize energies. The gradient descent flows will correspond to the terms R_k and $\mathcal{R}$. These flows approach infinity as the contour approaches self intersection, and becomes large when the contour is "irregular." The intuition for our approach is from electrostatics. We imagine that the contour has uniformly distributed charge along its perimeter, and that the contour moves in response to the charge as well as its original image-based force. We expect

that as the contour becomes close to self intersection, a repulsive force will arise due to its charge distribution and prevent self intersection. With this intuition, we propose to find the flow that minimizes the following energy functional:

$$E_\gamma(C) = \frac{1}{2} \iint_{C \times C} \frac{\mathrm{d}\hat{s}\,\mathrm{d}s}{\|C(\hat{s}) - C(s)\|^\gamma} \qquad (3)$$

where C denotes a curve, $\mathrm{d}s$ and $\mathrm{d}\hat{s}$ are arc-length measures, $\|\cdot\|$ is the usual Euclidean norm, and $\gamma > 0$. Each pair of points on the curve contributes an amount inversely proportional to its distance to the total energy. In the case when $\gamma = 1$, E_γ is the electric potential energy of the charge configuration assuming a 3-D flux. Note that, for $\gamma \geq 1$, $E_\gamma = +\infty$ for all contours. In this case, the inner integral diverges for every $C(s)$. In the case when $\gamma < 1$, the energy is finite; however, one can show that, as a point on the contour approaches another point, the gradient of the energy is finite. That is, for $\gamma < 1$, a topology change can occur if there is an image-based term present.

Energies similar to E_γ have been considered in mathematical literature on knot energies [31]–[34]. A knot is an embedding of the circle S^1 into $\mathbb{R}^3$. Two knots are said to be equivalent (or have the same knot type) if there exists an orientation preserving homeomorphism of $\mathbb{R}^3$ that maps one knot onto the other. It is of interest in the mathematical literature to introduce invariants of knots. One approach by O'Hara [31], [33] proposes an energy that becomes infinite as the knot changes its knot type, i.e., it begins to self intersect. An energy proposed is

$$E_\gamma(C) = \frac{1}{2} \iint_{C \times C} \left(\frac{1}{\|C(\hat{s}) - C(s)\|^\gamma} - \frac{1}{d_C(\hat{s}, s)^\gamma} \right) \mathrm{d}\hat{s}\,\mathrm{d}s \qquad (4)$$

where $1 < \gamma < 3$, and $d_C(s, \hat{s})$ is the geodesic distance (the shortest distance along C) from $C(s)$ to $C(\hat{s})$. The second term of (4) is used to "subtract out" the infinity from the first term, and, thus, the result is that the energy is finite. However, the energy has the property that it becomes infinite as the knot changes knot type. Our approach will be to consider (4) with $\gamma = 1$ for active contours, which has not been considered for use as a knot invariant because it does not diverge as the contour changes knot-type. We choose the energy with $\gamma = 1$ since, unlike the energy for $\gamma \geq 2$, it roughly favors contours with shorter length. Therefore, it seemed the corresponding gradient flow for $\gamma = 1$ would have a length shrinking effect, that would regularize the contour as in curvature flow; indeed, this is the case. Moreover, based on experimental and analytical results, we believe the gradient flow still preserves topology.

A. Polygon Case

In this section, we define an energy for the case of polygons, and then derive its gradient flow. We define an n-polygon, where $n \geq 3$, to be a closed contour, embedded in the plane, consisting of a n straight line segments. First, we introduce some notation: $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$, C is the set of all points in the polygon, v_i is the ith vertex of C where $i \in \mathbb{Z}_n$, and C_i is the edge of C connecting v_i to v_{i+1}.

1) Energy: We define an energy on n-polygons, $E_p : \mathbb{R}^{2 \times n} \to \mathbb{R}$, as

$$E_p(v_0, \dots, v_{n-1}) = \sum_{i \in \mathbb{Z}_n} (|C_i| \ln |C_i| - |C_i|)$$
$$+ \frac{1}{2} \sum_{(i,j) \in \mathbb{Z}_n \times \mathbb{Z}_n, i \neq j} \iint_{C_i \times C_j} \frac{\mathrm{d}\hat{s}\,\mathrm{d}s}{\|C_i(s) - C_j(\hat{s})\|} \qquad (5)$$

where $|C_i|$ denotes the length of C_i. The terms where the integral in (3) diverges, that is, the self-energies, $\iint_{C_i \times C_i} \mathrm{d}\hat{s}\,\mathrm{d}s / \|C(\hat{s}) - C(s)\|$, are replaced by the first term of (5), which arises from taking the "finite part" of the self-energies and discarding the "infinite component." The second term in (5) is the energy in (3) with $\gamma = 1$ excluding the self-energies.

A polygon regularizer using similar ideas to our approach is considered by Unal *et al.* [21]. Unal *et al.* assume that the polygon consists of a uniform charge distributed along the polygon's perimeter. However, an energy is not minimized; instead, the electrostatic force at each vertex is computed numerically, and the vertex is moved in the direction of that force. The electrostatic force at each vertex is infinite due to the near neighbor effects of the adjacent segments to the vertex. To deal with this problem, the near neighbor effects are simply ignored. While this method may prevent some topology changes, it will not prevent two adjacent sides from touching. Moreover, as the number of vertices increases and the length of the segments decreases, thus approaching the continuum, the flow appears to be unstable. The proposed minimization of the energy in (5) does not have the problems discussed above.

2) Vertex Evolution: We now derive a motion of the vertices that minimizes the energy given in (5). The usual gradient of E_p, is the vector $\nabla E_p(v_0, \dots, v_{n-1}) = (\partial E_p / \partial v_k(v_0, \dots, v_{n-1}))_{k=0}^{n-1}$. It can be shown that

$$F_k = F_k^{\text{self},k-1} + F_k^{\text{self},k} + \sum_{i \in \mathbb{Z}_n \setminus \{k-1\}} F_k^{k-1,i} + \sum_{i \in \mathbb{Z}_n \setminus \{k\}} F_k^{k,i} \qquad (6)$$

where $F_k := -\partial E_p / \partial v_k$

$$F_i^{j,k} = -\frac{1}{2} \frac{\partial}{\partial v_i} \iint_{C_j \times C_k} \frac{\mathrm{d}\hat{s}\,\mathrm{d}s}{\|C(\hat{s}) - C(s)\|}$$

and

$$F_i^{\text{self},j} = -\ln |C_j| \frac{v_i - v_j^*}{|C_j|} \quad \text{for } j = i-1, i. \qquad (7)$$

Note that $v_j^* = v_j$ when $j = i - 1$, and $v_j^* = v_{j+1}$ when $j = i$. The F's in the above expressions represent "forces"; they get their name from the fact that they are the negative gradient of some term of the energy, E_p. It is important to note, however, that F_k is not the same as the electrostatic force, which is the negative gradient of the *potential integral*, i.e., a single integral. Moreover, F_k is finite for embedded polygons, but the electrostatic force is infinite along the polygon. We have derived simplifications for F_k, which are found in [30]. The motion of the vertices to minimize E_p is

$$v_k'(t) = F_k(t) = -\partial_{v_k} E_p(t) \quad \text{for } k \in \mathbb{Z}_n \qquad (8)$$
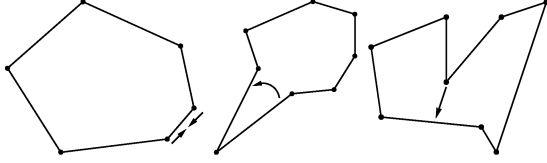
where t denotes an artificial time variable.

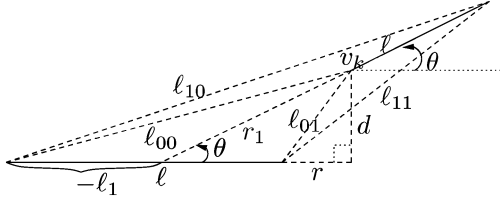Fig. 1. Illustration of three different cases of polygon topology change.

Fig. 2. Definition of quantities used in justifying Proposition 1.

3) Embeddedness of Polygon: We justify that an active polygon with the regularizer $R_k = F_k$ does not change topology, i.e., stays embedded. It can be shown that E_p does not become infinite, in general, as a polygon approaches a topology change; therefore, our approach is to show that R_k becomes infinite and moves the polygon away from changing topology. The system of ODE (8) is nonlinear and autonomous, and existence/uniqueness of a solution is of interest; however, this is beyond the scope of this paper.

Proposition 1 (Embeddedness of Polygon): Suppose $\{v_0^0, v_1^0, \ldots, v_{n-1}^0\}$ defines an embedded n-polygon, I_k is uniformly bounded, and a continuous solution of (1) exists and is unique. Then, the n-polygon evolving in time according to (1) with initial conditions $v_k(0) = v_k^0$ for $k \in \mathbb{Z}_n$ maintains its initial topology for all $t \in \mathbb{R}^+$ and any $\alpha > 0$.

Proof: We assume that the n-polygon defined according to (1) changes topology in finite time and derive a contradiction. The contradiction will arise from the fact that $\|R_k\| = \|F_k\| \to +\infty$ for an appropriate $k \in \mathbb{Z}_n$ (with I_k bounded by assumption), and that the direction of this R_k is opposite to the direction of topology change. There are three types of topology changes as illustrated in Fig. 1. In the interest of space, we analyze only the third case, and refer the reader to [30] for other cases and more details.

Assume a vertex, v_k, of the n-polygon approaches a segment, C_l, which is not adjacent to the segment containing v_k and does not contain v_k. We do not lose generality by assuming that $\|v_k - v_l\| \to 0$ where v_l is a vertex of C_l, and assuming that $|C_l| = |C_k| = \ell$ (see [30]). The diagram in Fig. 2 illustrates v_k and C_l, and labels some quantities that will be used. We can assume $\theta \in [0, \pi)$. For the moment we assume that $\theta \neq 0$ for all time up to and including the time of topology change. After an integration, we see that

$$E_p^{k,l} := \iint\limits_{C_l \times C_k} \frac{d\hat{s}ds}{\|C_l(\hat{s}) - C_k(s)\|}$$

$$= \sum_{i,j=0}^{1} (-1)^{j+1} \ell_{ij}^* \ln\left(\frac{\ell_{jj} + \ell_{\overline{i \oplus j}, i \oplus j} + \ell}{\ell_{jj} + \ell_{\overline{i \oplus j}, i \oplus j} - \ell}\right) \quad (9)$$

where $\oplus$ denotes the xor operation, $\overline{\cdot \oplus \cdot}$ denotes not xor, the quantities on the right hand side are defined in Fig. 2, $\ell_{00}^* = \ell_1$, $\ell_{01}^* = \ell + \ell_1$, $\ell_{10}^* = r_1$, and $\ell_{11}^* = \ell + r_1$. Note that the independent variables are d, r and θ. Differentiating (9) with respect to d we find

$$\frac{\partial E_p^{k,l}}{\partial d} = \sum_{i,j=0}^{1} (-1)^{j+1} \left[\frac{\partial \ell_{ij}^*}{\partial d} \ln\left(\frac{\ell_{jj} + \ell_{\overline{i \oplus j}, i \oplus j} + \ell}{\ell_{jj} + \ell_{\overline{i \oplus j}, i \oplus j} - \ell}\right) \right.$$
$$\left. - \frac{2\ell \ell_{ij}^*}{(\ell_{jj} + \ell_{\overline{i \oplus j}, i \oplus j})^2 - \ell^2} \frac{\partial}{\partial d}(\ell_{jj} + \ell_{\overline{i \oplus j}, i \oplus j}) \right]. \quad (10)$$

It is clear that the terms of the sum in (10) for $(i, j) \neq (1, 0)$ have finite limit as $(d, r) \to (0, 0)$. Thus, to show $\partial E_p^{k,l}/\partial d \to -\infty$, we only need to show the term with $(i, j) = (1, 0)$ diverges. To show this, first note that $\ell_{00} = \sqrt{(r+\ell)^2 + d^2}$, $\ell_{01} = \sqrt{r^2 + d^2}$, and $\ell_{10}^* = d/\sin\theta$. This gives that $(\ell_{00})_d = d/\ell_{00}$, $(\ell_{01})_d = d/\ell_{01}$ and $(\ell_{10})_d^* = 1/\sin\theta$ where the subscript d denotes partial differentiation. With some manipulation, we find that the second term in (10) has limit

$$\lim_{(r,d) \to (0,0)} \frac{\ell}{\ell_{00}\sin\theta} \frac{d^2(\ell_{01} + \ell_{00})}{\ell_{01}(r^2 + r\ell + d^2) + (r^2 + d^2)\ell_{00}}$$
$$= \frac{\ell}{\sin\theta} \lim_{(r,d) \to (0,0)} \frac{d^2}{\ell_{01}(r^2 + r\ell + d^2) + (r^2 + d^2)\ell_{00}}. \quad (11)$$

We claim that (11) has finite limit along any path. To show this, note that locally around $(r, d) = (0, 0)$, a path may be represented as a function. Further, from continuity of the flow, the function is continuous. There are two cases. First, $r = r(d)$ and $r'(0)$ exists in that case, $r = \mathcal{O}(d)$, where $\mathcal{O}$ denotes big-oh. The second case is when $d = d(r)$ and $d'(0)$ exists in that case, $d = \mathcal{O}(r)$. It is easy to verify that (11) is finite in both cases. Therefore, we see that

$$\lim_{(r,d) \to (0,0)} \frac{\partial E_p^{k,l}}{\partial d} = c - \lim_{(r,d) \to (0,0)} \frac{1}{\sin\theta} \ln\left(\frac{\ell_{00} + \ell_{01} + \ell}{\ell_{00} + \ell_{01} - \ell}\right).$$

Hence, it is clear that $\|F_k\| \to +\infty$, and the direction of F_k is in a direction away from $d \to 0$. This contradicts the assumption that $d \to 0$ (see [30] in the case when $\theta = 0$). $\square$

B. Differentiable Contour Case

1) Energy Functional: Let $\mathcal{C}$ denote the set of all twice differentiable embedded curves in $\mathbb{R}^2$. We define an energy $E_c : \mathcal{C} \to \mathbb{R}^+$ as follows:

$$E_c(C) = \frac{1}{2} \iint\limits_{C \times C} \left(\frac{1}{\|C(s) - C(\hat{s})\|} - \frac{1}{d_C(s, \hat{s})}\right) d\hat{s}ds. \quad (12)$$

We set $d_C(s, \hat{s}) = +\infty$ if there is no path along C connecting $C(s)$ to $C(\hat{s})$, which is possible if the curve consists of two or more embedded curves. Note that the energy in (12) is the same energy as in (4) with $\gamma = 1$, and, therefore, it is not a knot energy, i.e., it does not, in general, diverge as a curve approaches self-intersection. Nevertheless, we shall give analytical justification in Section II-B4 that the gradient flow will preserve the embeddedness of a curve. The regularization term, i.e., the second

term, in (12) only "cancels out" the infinity of the electrostatic term, i.e., first term, due to points on the curve that are close together in the geodesic sense. The regularization does not affect the asymptotic behavior of the electrostatic term as a curve approaches self-intersection. This is because points on the curve that touch during self-intersection are far away in the geodesic sense.

2) Curve Evolution: For each curve $C \in \mathcal{C}$, we find the perturbation of C that decreases the energy E_c the fastest with respect to the usual geometric L^2 inner product. This perturbation is the negative gradient (if it exists), $-\nabla E_c(C)$, which is defined by the relation

$$\mathrm{d}E_c(C) \cdot \tilde{C} = \left\langle \tilde{C}, \nabla E_c(C) \right\rangle_{C, L^2} \tag{13}$$

where $\mathrm{d}E_c(C) \cdot \tilde{C}$ denotes the variation of E_c at C in the direction of a perturbation $\tilde{C}$.

Proposition 2 (Gradient of E_c): The gradient of E_c is

$$-\nabla E_c(C) = \lim_{\epsilon \to 0^+} \left[\mathcal{E}_\epsilon + \mathcal{P}_\epsilon \kappa - \ln\left(\frac{L}{2\epsilon}\right)\kappa \right] \mathcal{N} \tag{14}$$

where

$$\mathcal{E}_\epsilon(s) = \int_{B_C(\epsilon, s)} \frac{C(s) - C(\hat{s})}{\|C(s) - C(\hat{s})\|^3} \cdot \mathcal{N}(s)\mathrm{d}\hat{s} \tag{15}$$

$$\mathcal{P}_\epsilon(s) = \int_{B_C(\epsilon, s)} \frac{\mathrm{d}\hat{s}}{\|C(s) - C(\hat{s})\|} \tag{16}$$

κ denotes the curvature of C, $\mathcal{N}$ is the unit inward normal to C, L denotes the length of C, and $B_C(\epsilon, s) = \{C(\hat{s}) : d_C(\hat{s}, s) > \epsilon\}$.

Proof: We refer the reader to [30] for details.

The curve evolution given by the PDE

$$C_t = -\nabla E_c(C) =: \mathcal{R} \tag{17}$$

minimizes the energy E_c optimally, in the L^2 sense. Notice that this flow at a particular point depends on all other points on the curve in the form of integrals around the curve.

The term $\mathcal{E}_\epsilon + \mathcal{P}_\epsilon \kappa$ arises from the electrostatic term, or the first term, of E_c, and the term $\ln(L/2\epsilon)\kappa$ arises from the regularization term, or second term, of E_c. Although $\mathcal{R}$ appears to have a term that is negative curvature, when the limit is evaluated, there is only a positive curvature component, and, thus, the flow is stable. Intuitively in terms of electrostatics, $\mathcal{E}_\epsilon$ is the projection of a regularized electric field vector of the charge distribution along C onto the inward normal of C. Repulsion between nearby points on the curve arises from this term. The term, $\mathcal{P}_\epsilon$, is a regularization of the electrostatic potential, and $\kappa\mathcal{P}_\epsilon\mathcal{N}$ *regularizes* and shrinks the curve. Note that shrinking the curve reduces charge (because of our assumption of a uniform distribution), and, therefore, this motion is consistent with a reduction of E_c. The purpose of the $\ln(L/2\epsilon)\kappa$ term is to cancel out the infinity arising from $\mathcal{E}_\epsilon + \mathcal{P}_\epsilon \kappa$ as $\epsilon \to 0$.

3) Behavior of the Circle: We give a simple example of computing E_c and $\mathcal{R}$ for a circle of radius R centered at the origin. Note that by symmetry of the circle, it is clear that $\mathcal{E}_\epsilon(s)$,

$\mathcal{P}_\epsilon(s)$, and the inner integral of E_c do not depend on s. Therefore, we choose for simplicity to evaluate the previous quantities at $s = 0$, i.e., the point $(R, 0)$. We use the parameterization $C(\theta) = R(\cos\theta, \sin\theta)$ where $\theta \in [0, 2\pi]$ and $\hat{s} = R\theta$. A computation gives that $\|C(0) - C(\hat{s})\| = 2R\sin(\theta/2)$, and $d_C(0, \hat{s}) = R\theta$ for $\theta \in [0, \pi]$. Substituting the previous expressions into (12) and simplifying, we find that

$$E_c(C) = 2\pi R \lim_{\epsilon \to 0^+} \ln\left(\frac{\frac{\epsilon}{R}\sin\epsilon/2R}{\pi(1 - \cos\epsilon/2R)}\right) = 2\pi R \ln\frac{4}{\pi}.$$

Noting that the inward normal is given by $\mathcal{N}(0) = (-1, 0)$, that $(C(0) - C(\hat{s})) \cdot \mathcal{N}(0) = -R(1 - \cos\theta) = -2R\sin^2\theta/2$, and simplifying, we see that

$$\mathcal{E}_\epsilon = -\frac{1}{2R}\ln\left(\frac{1 + \cos\epsilon/2R}{1 - \cos\epsilon/2R}\right), \mathcal{P}_\epsilon = 2\ln\left(\frac{\sin\epsilon/2R}{1 - \cos\epsilon/2R}\right).$$

Clearly, $\kappa = 1/R$. Hence, we see

$$\begin{aligned}
\mathcal{R} = \lim_{\epsilon \to 0^+} &\left[-\frac{1}{2R}\ln\left(\frac{1 + \cos\epsilon/2R}{1 - \cos\epsilon/2R}\right) \right. \\
&\left. + \frac{2}{R}\ln\left(\frac{\sin\epsilon/2R}{1 - \cos\epsilon/2R}\right) - \frac{1}{R}\ln\left(\frac{\pi R}{\epsilon}\right)\right] \\
= &\frac{1}{R}\ln\frac{4}{\pi}.
\end{aligned} \tag{18}$$

First, we observe that the energy computed above clearly reconciles with the flow $C_t(s) = \ln(4/\pi)/R\mathcal{N}(s)$ computed in (18). Second, we see from the flow that the circle shrinks to a point in finite time. This property is also seen in curvature flow. This suggests that the flow $C_t = \mathcal{R}$ may shrink arbitrary curves to points in finite time.

4) Embeddedness of Curve: We justify that an active contour with (17) and a bounded force stays embedded. We give a nonrigorous argument that $\mathcal{R}$ becomes infinite and is in a direction opposite to self intersection as the curve approaches self-intersection. We assume that a solution of (2) exists and κ stays well-defined, although we cannot prove this assumption.

Conjecture 1 (Embeddedness of Curve): Suppose $C_0 \in C^2(S^1, \mathbb{R})$ defines an embedded curve and $i \in C^2(S^1 \times \mathbb{R}^+, \mathbb{R})$ is uniformly bounded. Then the curve evolving in time according to (2) with initial condition $C(\cdot, 0) = C_0$ stays embedded for all $t \in \mathbb{R}^+$ for any $\alpha > 0$.

Proof: We summarize the argument (see [30] for details). Let us first define for convenience

$$\begin{aligned}
I(p, q, t) := \left(\frac{\kappa(p, t)}{\|C(p, t) - C(q, t)\|} + \frac{C(p, t) - C(q, t)}{\|C(p, t) - C(q, t)\|^3} \right. \\
\left. \cdot \mathcal{N}(p, t)\right) \|C_p(q, t)\| \mathcal{N}(p, t). \tag{19}
\end{aligned}$$

Step 1) We assume for the sake of contradiction that $C(\cdot, t)$ changes topology, for the first time, in finite time $t_i \in \mathbb{R}^+$. That is, $\exists \delta_0 > 0$ and $p_1, p_2 \in C([0, t_i]; S^1)$ such that $\lim_{t \to t_i}\|C(p_1(t), t) - C(p_2(t), t)\| = 0$ and $d_{C(\cdot, t)}(p_1(t), p_2(t)) > \delta_0 \,\forall t \in [0, t_i)$.

Step 2) By translating the coordinate axis so that $C(p_1(t), t)$ coincides with the origin, we may assume that $\mathrm{d}/\mathrm{d}t C(p_1(t), t) = 0$. This, along with the conclusion in Step 1), implies that $\exists t_0 \in (0, t_i)$
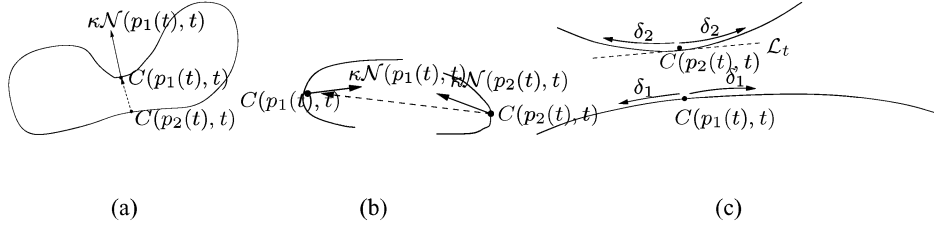
Fig. 3. (a) Curve moving toward topology change with quantities labeled as in the justification of Conjecture 1. A topology change cannot occur, that is $C(p_2(t), t) \nrightarrow C(p_1(t), t)$, if locally, the curve and directions are in (b) for all time close to t_i. (c) Shows quantities defined in Step 3) of the proof.

such that $\mathrm{d}/\mathrm{d}t C(p_2(t), t) \neq 0 \; \forall t \in [t_0, t_i)$. We may now assume without loss of generality that $(\mathrm{d}/\mathrm{d}t C(p_2(t), t)) \cdot \kappa(p_1(t), t)\mathcal{N}(p_1(t), t) \geq 0$ $\forall t \in [t_0, t_i)$. This statement says that $C(p_2(t), t)$, after some time, moves with a component in the same direction as $\kappa(p_1(t), t)\mathcal{N}(p_1(t), t)$. An intuitive argument for this statement is in Fig. 3(b). If the opposite of the statement is true, then for time near enough to t_i, the curve locally looks as shown in Fig. 3(b), in which a topology change cannot occur.

Step 3) We assume that κ stays well-defined (does not become infinite) from $t \in (0, t_i)$, and then break the domain of the curve into three pieces: **1.** $A_\epsilon(t) := \{p \in S^1 : \epsilon < d_{C(\cdot, t)}(p_1(t), t) < \delta_1\}$ where $\delta_1 > 0$ does not depend on time and is chosen so that

$$\lim_{\epsilon \to 0^+} \left[\int_{A_\epsilon(t)} I(p_1(t), p, t) \, \mathrm{d}p - \ln \frac{L(t)}{2\epsilon} \kappa(p_1(t), t)\mathcal{N}(p_1(t), t) \right]$$

can shown to be uniformly bounded for $t \in (t_0, t_i)$ through a Taylor series. **2.** $D_t := \{p \in S^1 : d_{C(\cdot, t)}(p_2(t), p) < \delta_2\}$ where δ_2, independent of time, is chosen so that $C(D_t, t)$ may be approximated by a line segment, $\mathcal{L}_t$, of constant length ℓ that passes through $C(p_2(t), t)$. This line segment approximation is made possible through the assumed smoothness of the curve. **3.** $E_t := S^1 \setminus (A_\epsilon(t) \cup D_t)$ where it is possible to show that $\int_{B_t} I(p_1(t), p, t) \mathrm{d}p$ is uniformly bounded for $t \in (t_0, t_i)$ since $C(E_t, t)$ stays sufficiently far from $C(p_1(t), t)$.

Step 4) The infinite behavior of $\mathcal{R}$ can only be from $C(D_t, t)$, which is approximated by $\mathcal{L}_t$. So

$$\int_{D_t} I(p_1(t), p, t) \, \mathrm{d}p$$

$$\approx \int_{\mathcal{L}_t} \frac{\kappa(p_1(t), t)\mathcal{N}(p_1(t), t) \, \mathrm{d}s}{\|C(p_1(t), t) - \mathcal{L}_t(s)\|}$$

$$+ \int_{\mathcal{L}_t} \frac{C(p_1(t), t) - \mathcal{L}_t(s)}{\|C(p_1(t), t) - \mathcal{L}_t(s)\|^3} \cdot \mathcal{N}(p_1(t), t) \, \mathrm{d}s \mathcal{N}(p_1(t), t)$$

$$= \ln \left(\frac{r_1(t) + r_2(t) + \ell}{r_1(t) + r_2(t) - \ell} \right) \kappa(p_1(t), t)\mathcal{N}(p_1(t), t)$$

$$+ \frac{2\mathcal{N}(p_1(t), t)}{(r_1(t) + r_2(t))^2 - \ell^2} \left(\frac{\vec{r_1}(t)}{r_1(t)} + \frac{\vec{r_2}(t)}{r_2(t)} \right)$$

$$\cdot \mathcal{N}(p_1(t), t) \tag{20}$$

where $\vec{r_1}(t)$ and $\vec{r_2}(t)$ are the vectors from the endpoints of $\mathcal{L}_t$ to $C(p_1(t), t)$, $r_1(t) = \|\vec{r_1}(t)\|$ and $r_2(t) = \|\vec{r_2}(t)\|$. By Step 1), we have that $\lim_{t \to t_i} \|C(p_1(t), t) - C(p_2(t), t)\| = 0$, and since $C(p_2(t), t) \in \mathcal{L}_t$ for $t \in (t_0, t_i)$, we conclude that $r_1(t) + r_2(t) \to \ell$ as $t \to t_i$. It is clear that for t near t_i, the second term of (20) points in the same direction as $\kappa(p_1(t), t)\mathcal{N}(p_1(t), t)$. Moreover, the norm of both terms of (20) approach $+\infty$ as $t \to t_i$. Since we have chosen the coordinate system so that $C(p_1(t), t)$ is fixed, it follows that for t close to t_i, we have that $(\mathrm{d}/\mathrm{d}t C(p_2(t), t)) \cdot \kappa(p_1(t), t)\mathcal{N}(p_1(t), t) < 0$. This contradicts the conclusion in Step 2). $\qquad\square$

5) Remarks on Related Work: The work of Rochery *et al.* [22] makes use of an energy that is defined as a double integral around contours rather than a single integral, which is traditionally used. The goal of [22] is to segment line structures in satellite images, and, thus, a *prior* shape energy is constructed to favor thin elongated structures. The energy is

$$E(C) = - \int_{S^1} \int_{S^1} g \left(\|C(p) - C(\hat{p})\| \right) C'(p) \cdot C'(\hat{p}) \mathrm{d}p \mathrm{d}\hat{p} \tag{21}$$

where p and $\hat{p}$ are parameters denoting a parametrization of C, and $g : \mathbb{R} \to \mathbb{R}$ is heuristically chosen so that E will be minimized by thin elongated structures.

We now illustrate the differences between the model of Rochery *et al.* and our model. First, we note that the energy in (21) contains the term $C'(p) \cdot C'(\hat{p})$ in the integrand. This is different in our model, in which the contribution a pair of points makes to the energy depends *only* on the distance between them in the Euclidean and geodesic sense. We learned from an anonymous reviewer that this tangent term leads to producing "arms" of the contour, which is not desirable for the purpose of this paper. A major difference is the lack of a regularization term arising from the gradient flow that minimizes (21); therefore, their gradient flow does not have a *global regularizing* property as the flow we consider. For regularity, the authors of [22] add a separate curvature term to their flow. Finally, (21) does not allow for a term inside the integrand depending on the geodesic distance. In [23], the authors mention replacing $C'(p) \cdot C'(\hat{p})$ with $\|C'(p)\|\|C'(\hat{p})\|$ as in the energy we consider, but do not pursue it.

III. NUMERICAL IMPLEMENTATION

A. Polygons

We implement the polygon evolution using a standard marker particle method. That is, the vertices of the evolving polygon are

stored in a $N \times 2$ array where N is the number of vertices of the polygon. The vertices are assumed to be ordered so that v_k is adjacent to v_{k+1} for all $k \in \mathbb{Z}_n$. Discretizing the continuous evolution given in (8) yields: $v_k^{n+1} = v_k^n + \Delta t F_k^n$, where v_k^n denotes the kth vertex at time n, similarly for the force F_k^n, and Δt is the step size, which is a small positive number. There are no restrictions on the step size to guarantee stability. Typically, we use $\Delta t = 1$. However, due to the fact that finite time steps are used, it may be necessary to use an additional discrete topology preservation step to guarantee that the topology is preserved numerically (see Section III-B2).

B. Level Set Method

1) Approximating the Flow: We consider the flow

$$C_t(s) = \left(\mathcal{E}_\epsilon(s) + \kappa(s) \tilde{\mathcal{P}}_\epsilon(s) \right) \mathcal{N}(s) \quad (22)$$

where $\tilde{\mathcal{P}}_\epsilon = \mathcal{P}_\epsilon - \ln(L/2\epsilon)$ and $\epsilon > 0$ chosen small, as an approximation of implementing (17). We perform the numerical implementation of the contour evolution given in (22) using level set methods [17]. To implement the flow in (17), the evolving curve is embedded as the zero level set of a scalar function $\Psi : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{R}$ that evolves in time. The evolution of Ψ is

$$\Psi_t(x,t) = -\nabla \Psi(x,t) \cdot C_t(x) \quad \text{for } x \in C(t)$$

where Ψ_t denotes the derivative with respect to the second argument, and $\nabla \Psi$ denotes the gradient with respect to the first argument. Noting that $\mathcal{N} = -\nabla \Psi / \|\nabla \Psi\|$ (if Ψ is arranged to be positive outside the curve, and negative inside the curve) the curve evolution in (17) becomes

$$\Psi_t = -\nabla \Psi \cdot \vec{E} + \kappa P \|\nabla \Psi\| \quad (23)$$

where

$$\vec{E}(s) = \int_{B_C(\epsilon,s)} \frac{C(s) - C(\hat{s})}{\|C(s) - C(\hat{s})\|^3} \mathrm{d}\hat{s}, \text{ and } P = \tilde{\mathcal{P}}_\epsilon. \quad (24)$$

We implement (23) on a finite-square grid, so a discretization of the equation must be performed. Careful attention is paid to estimating the derivatives of Ψ, which are performed using appropriate differencing schemes. The discrete approximation to (23) using a forward Euler scheme is

$$\Psi_{n+1}(i,j) = \Psi_n(i,j) \\ + \Delta t \left(-\nabla \Psi_n(i,j) \cdot \vec{E} + \kappa P \|\nabla \Psi\|_n(i,j) \right) \quad (25)$$

where $\vec{E}$ and P denote the discrete approximation. The time step, Δt, is chosen to ensure stability.

To efficiently implement (25), we use a narrowband technique. We can now define $\vec{E}$ and P in the narrowband by simply defining them to be the same value as at its closest point to the contour, i.e., the zero-level set. Since the narrowband will be thin, this definition makes sense, that is, there is a unique closest point to the contour. A high-level description of the algorithm to implement the level set evolution is given as follows.

1) Find polygonal estimate of zero-level set of Ψ.
2) Compute $\vec{E}$ and P at each vertex of polygonal estimate.
3) Extend $\vec{E}$ and P to narrow band of zero-level set of Ψ.

4) Update Ψ based on (25).
5) Repeat above steps until Ψ converges.

An alternative *ad-hoc* method based on using the FFT for implementing (25) can be found in [30]. The advantage of this method is that it does not require the polygon extraction step.

2) Ensuring Discrete Topology Preservation: Because of the finite grid size and a discrete time step in (25), it is possible that when the time step is not chosen sufficiently small, there may be a topology change in the discretization of the proposed PDE when there are large image-based forces. Instead of trying to derive a bound on the step size to guarantee that the topology is preserved, which may lead to excessively small time steps, and, therefore, a time consuming implementation, we *pair* the discrete topology preserving scheme of Han *et al.* [18], [19] (any other discrete topology scheme will suffice) with the constructed geometric flow. One may wonder what the necessity of using the proposed flow is when the method of Han *et al.* already preserves discrete topology. Our objective is not just topology preservation, but also producing a geometrically accurate segmentation. The addition of the proposed term makes it much less likely that the geometric pathologies of using the method of Han *et al.* alone, which occur because of using the method on an inappropriate PDE as mentioned in the Introduction, will occur (see experiments in Section IV).

Using a discrete topology preserving step is also a better simulation of the underlying topology preserving PDE than not using such a step considering the discrete time step and the finite resolution grid. This is because without such a discrete step, a topology change may occur, and such a change is inconsistent with the topology preserving property of the PDE. Also, as the grid resolution becomes finer and the step size becomes smaller, the proposed method will not need to invoke the discrete topology preserving step to guarantee topology preservation, and the evolution with the discrete topology step approaches the true solution of the PDE. Therefore, the use of the discrete step leads to a good approximation of the true solution of the PDE, which is also consistent with the topology preserving property of the PDE.

The key to the method of Han *et al.* is detecting *simple points*. A point is said to be *simple* if its addition to or removal from a digital object does not change the object topology. A local condition to determine if a point is simple is proved in [35]. A point x is simple iff the number of connected components of the foreground in a 3×3 neighborhood of x is equal to the number of connected components in the background, and both numbers are one. The algorithm of Han *et al.* updates each grid point of the level set function successively and checks each grid point to see whether the point is simple; if the point is simple, then the level set function at the point is updated ordinarily; if not and the point would change sign, then the point is updated as far as possible without changing sign, thus preventing a discrete topology change.

IV. SIMULATIONS

A. Geometric Properties of Flow

The top row of Fig. 4 shows the evolution of a polygonal spiral evolving solely under the flow (8). Notice that the spiral unravels from the inside by shrinking its inner segments and pushing segments into a small area. The polygon then becomes
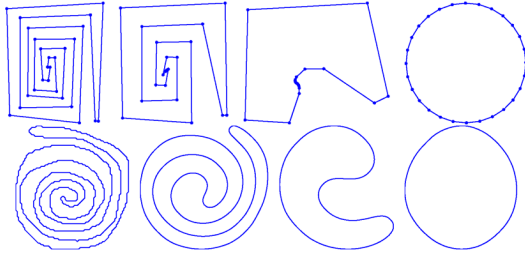
Fig. 4. (Top) Snapshots of an evolution of a polygonal spiral under (8) and (bottom) evolution of a spiral under (22). For visibility, the spatial scale is shrunk as the evolution progresses.

Fig. 6. Snapshots of an evolution of active polygon using Chan–Vese flow with the proposed topology preserving flow.
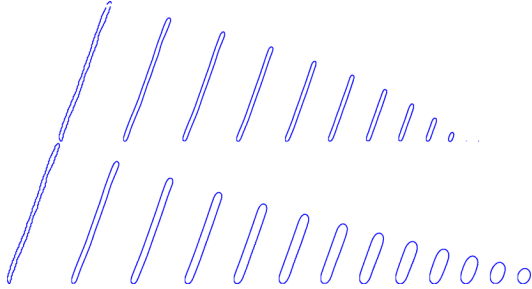

Fig. 5. (Top) Illustration of a difference between curvature flow and (bottom) proposed flow (22).

convex, and finally converges to a regular polygon. This suggests that any polygon converges to a regular polygon under this flow, but this has not been proven. Therefore, this flow can also be used as a global regularity term in addition to preserving topology of active polygons.

The bottom row of Fig. 4 shows the evolution of a spiral evolving solely under the flow (22). The spiral eventually becomes convex, becomes circular, and shrinks to a configuration that is not representable with the given resolution of the level set function. The behavior of this flow, as witnessed through this experiment, shows some similarities to curvature flow. In particular, the topology preserving flow has smoothness properties like the curve shortening flow. It also unwinds the spiral without crossing itself like curvature flow. It was shown in [34] that the energy E_c is minimized by a circle among contours having a constant perimeter. Fig. 5 shows an obvious difference between (22), which is shown on the bottom, and the curvature flow, which is shown on the top. Notice in the bottom of Fig. 5, that the curve becomes thicker as parallel sides are pushed apart. This is due to the term $\mathcal{E}_\epsilon \mathcal{N}$ in (22). Since these parallel sides have nearly zero curvature, curvature flow cannot not push the sides apart. For topology preservation, this repulsion is desirable.

B. Image Segmentations

In all of the following image segmentations, we have used the flow considered in [15] as the image-based term. The flow, in summary, moves a contour or polygon to separate an image into two regions that are piecewise constant. Note that any weight on the proposed regularizing/ topology preserving flow can be used to preserve topology. However, the lower the weight on the topology preserving term, the greater the possibility for dependence on the method of Han *et al.*
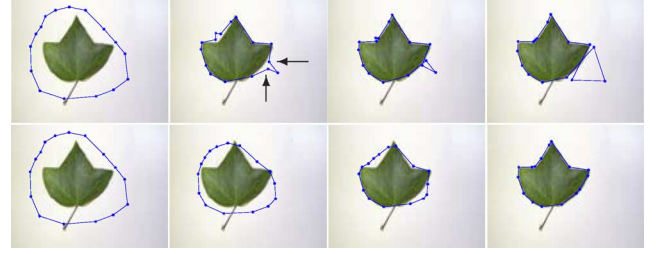
Results of segmentation of a leaf image with an active polygon are shown in Fig. 6. The result of running the Chan–Vese [15] flow without the proposed topology preserving flow is shown on the top row of Fig. 6. From the nature of the Chan–Vese flow, the vertices indicated by the arrows in Fig. 6 will move the polygon towards self intersection. This is because the vertices will move in the direction that reduces the white area as fast as possible; this direction moves the polygon towards self intersection. After self intersection, the normal vectors become flipped at the indicated vertices, and the flow moves in the wrong direction. Since there is no way to consistently define a normal vector field when the polygon is no longer embedded, there is no simple way to correct this problem. The bottom row of Fig. 6 shows snapshots of the evolution of Chan–Vese flow weighted 98% and the topology preserving flow weighted 2%. Notice the continuous motion of the polygon to keep the polygon away from self intersection. The polygon is more "regular" during the evolution than the polygon that results without using the proposed flow, and this verifies that the flow serves to globally regularize the polygon.

The next simulations are using active contours for image segmentation. Fig. 7 shows the results of segmentation of an image with two closely spaced bones with no topology preservation, topology preservation of Han *et al.* [19] alone, and the proposed method. The top row is the result with a curvature regularization term weighted 50% to compensate for the noise. The two contours merge across the thin gap between the two bones. This merging is due to the nature of the image and Chan–Vese flow, and not a step size issue. The middle row shows the result of the topology preserving method of Han *et al.* with curvature weighted 50%. The final row shows the result with our proposed flow weighted 27%. A zoom of the region separating the two bones is shown in Fig. 8. Notice that the contour is stopped abruptly in an arbitrary location relative to the image features in the result using the method of Han *et al.* alone. These kind of artifacts are typical of a forced topology constraint that does not model the actual behavior of the underlying PDE. In contrast, using the proposed flow globally regularizes the evolving contour and thereby pushes the contour to attract relevant image features. Since the proposed flow gradually adjusts the flow away from topology change, the contour is never one pixel away from self-intersecting, and, thus, the discrete topology preservation step never needs to be invoked in this example.

Fig. 9 shows the evolution of an active contour segmenting a simple image of two circles connected by a thin strip. The top row shows the result of the topology preserving level set
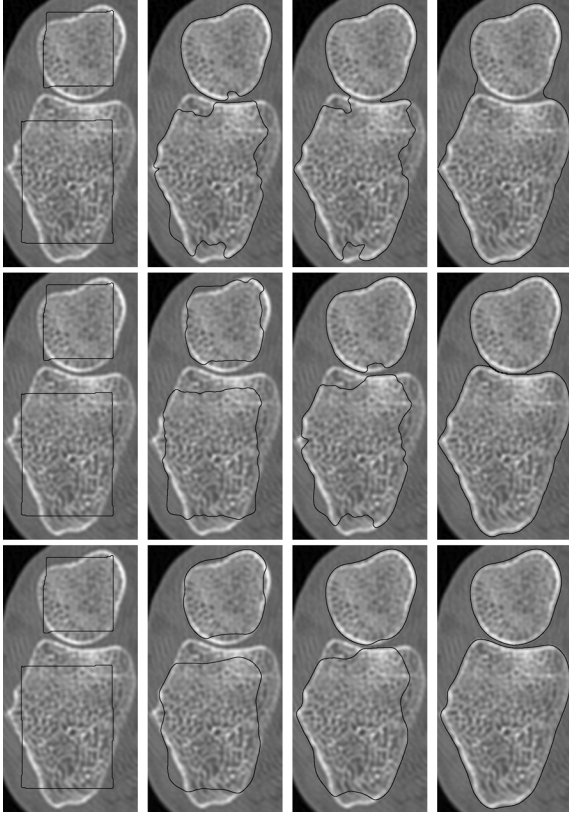
Fig. 7. Segmentation of bones CT image with (top row) no topology preservation, (middle) topology preservation of Han *et al.* alone, and (bottom) proposed topology preservation.
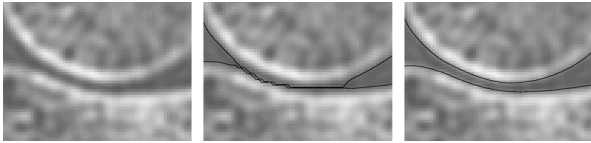

Fig. 8. Gap between bones with contours overlaid. Results of no topology preservation, topology preservation of Han alone, and proposed topology preservation.

method of Han *et al.* alone, and the bottom row shows the results of the proposed method. Note that we have used curvature regularization (weighted 10%) in the method of Han to keep the curve smooth. Although the topology is correct, the segmentation is geometrically inaccurate. One may argue that weighting the curvature high enough in the method of Han *et al.* alone will correct this problem. However, weighting the curvature just high enough to overcome the pathology, produces a flow that looks like curvature flow, and fails to even capture the circles in the image. The bottom row of Fig. 9 shows the result of segmentation with Chan–Vese flow weighed 96% and the proposed flow weighted 4%. The curve repels from itself in a continuous manner as it becomes close to self intersection. The curve is then pushed in the right direction to capture the thin strip as the thin strip slowly becomes detected by the image-based term. This experiment serves to illustrate geometrical pathologies that generally occur when a discrete topology constraint is artificially forced on a PDE that exhibits no topology preserving properties. The experiment also illustrates that a true topology preserving
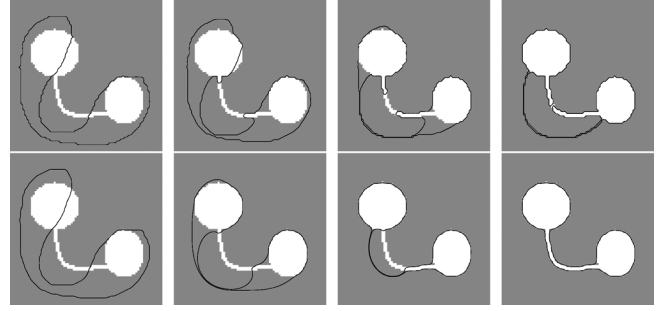

Fig. 9. Geometrical pathology typical of forced topology preservation constraint of (top) Han *et al.* alone, and the correct segmentation by the proposed geometric regularization flow.
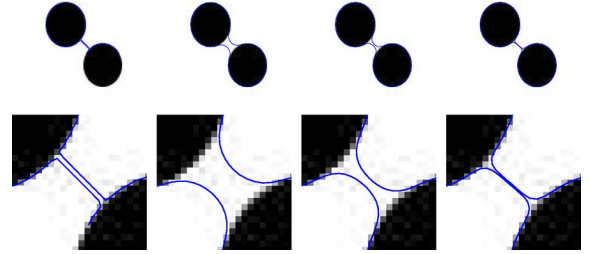

Fig. 10. Segmentations of a simple image to illustrate the effect of changing the weight on the proposed flow. Top row (left to right): Result of Han *et al.* (curvature regularizer weighted 30%), proposed flow weighted 33%, 20%, and 7%, respectively. Bottom row: Enlargements of center area between circles of images in top row.

PDE with global geometric regularization reduces the possibilities of these geometric pathologies.

A simulation illustrating the effect of the weighting factor of the proposed flow on the final segmentation is shown in Fig. 10. The initial contour is a circle that encloses both black circles. In the simulations showing the proposed flow, the portion of the contour between the circles stops when the Chan–Vese force balances the proposed force. Weighting the proposed term lower gives a result similar to using Han *et al.*'s method alone. Increasing the weight of the proposed term ensures the final segmentation is increasingly further from topology change, and is more globally regular. The proposed flow affects the original image-based evolution significantly when the curve is close to self intersection, and is highly irregular.

V. CONCLUSION

In summary, we have presented a novel method for introducing global regularity into contour and polygon evolutions and simultaneously enforcing a prior assumption on the topology of objects to be detected from an image. This was done by constructing a geometric flow that can be added to existing image-based evolutions of active polygons or contours. We have demonstrated that the constructed flow through its global regularization and gradual topology preservation reduces the possibility of geometric pathologies that plague discrete topology constraints forced on PDEs having no tendency to preserve topology. In particular, we noted that the justified way for using a discrete topology constraint is in conjunction with a PDE that preserves topology as an appropriate numerical technique. We have given analytical justification that the

constructed geometric flow preserves topology in the presence of arbitrary bounded image-based forces in continuous time. Experiments verified the topology preserving properties of the proposed flow, and showed that global regularization and gradually adjusting topology produce geometrically accurate segmentations that are often needed in applications.

REFERENCES

[1] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *Int. J. Comput. Vis.*, vol. 1, pp. 321–331, 1987.

[2] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *Proc. IEEE Int. Conf. Computer Vision*, Cambridge, MA, Jun. 1995, pp. 694–699.

[3] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Gradient flows and geometric active contour models," in *Proc. IEEE Int. Conf. Computer Vision*, 1995, pp. 810–815.

[4] V. Caselles, F. Catte, T. Coll, and F. Dibos, "A geometric model for active contours in image processing," Tech. Rep. 9210, Université Paris Dauphine, Ceremade, France, 1992.

[5] R. Malladi, J. Sethian, and B. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 2, pp. 158–175, Feb. 1995.

[6] M. Gage and R. S. Hamilton, "The heat equation shrinking convex plane curves," *J. Diff. Geom.*, no. 23, pp. 69–96, 1986.

[7] M. Grayson, "The heat equation shrinks embedded planes curves to round points," *J. Diff. Geom.*, vol. 26, pp. 285–314, 1987.

[8] G. Sapiro and A. Tannenbaum, "Affine invariant scale space," *Int. J. Comput. Vis.*, vol. 9, no. 26, pp. 25–44, 1993.

[9] ——, "On affine plane curve evolution," *J. Funct. Anal.*, vol. 1, no. 119, pp. 79–120, 1994.

[10] K. Siddiqi, Y. B. Lauzière, A. Tannenbaum, and S. Zucker, "Area and length minimizing flows for shape segmentation," *IEEE Trans. Image Process.*, vol. 3, no. 7, pp. 433–443, Jul. 1998.

[11] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 359–369, Mar. 1998.

[12] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.

[13] R. Ronfard, "Region based strategies for active contour models," *Int. J. Comput. Vis.*, vol. 13, no. 2, pp. 229–251, Oct. 1994 [Online]. Available: http://perception.inrialpes.fr/Publications/1994/Ron94

[14] A. Yezzi, A. Tsai, and A. Willsky, "A statistical approach to snakes for bimodal and trimodal imagery," in *Proc. Int. Conf. Computer Vision*, Oct. 1999, pp. 898–903.

[15] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[16] N. Paragios and R. Deriche, "Geodesic active regions: a new paradigm to deal with frame partition problems in computer vision," *Int. J. Vis. Commun. Image Represent.*, vol. 13, no. 2, pp. 249–268, Jun. 2002.

[17] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on the Hamilton–Jacobi equations," *J. Comput. Phys.*, vol. 79, pp. 12–49, 1988.

[18] X. Han, C. Xu, and J. Prince, "Topology preserving deformable model using level sets," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Dec. 2001, pp. 765–770.

[19] ——, "A topology preserving level set method for geometric deformable models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 6, pp. 755–768, Jun. 2003.

[20] A. M. Bruckstein, G. Sapiro, and D. Shaked, "Evolutions of planar polygons," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 9, pp. 991–1014, 1995.

[21] G. Unal, A. Yezzi, and H. Krim, "Unsupervised texture segmentation by information-theoretic active polygons," *Int. J. Comput. Vis.*, vol. 62, no. 3, pp. 199–220, 2005.

[22] M. Rochery, I. Jermyn, and J. Zerubia, "Higher order active contours and their application to the detection of line networks in satellite imagery," presented at the IEEE Workshop VLSM, Oct. 2003.

[23] ——, "Higher order active contours," Res. Rep. 5656 INRIA, France, Aug. 2005 [Online]. Available: http://www.inria.fr/rrrt/rr-5656.html

[24] J. Kim, J. Fisher, A. Yezzi, M. Cetin, and A. Willsky, "Nonparametric methods for image processing using information theory and curve evolution," in *Proc. IEEE Int. Conf. Image Processing*, 2002, vol. 3, pp. 797–800.

[25] H. Chang and D. Valentino, "Medical image segmentation using a simulated charged fluid," *Proc. SPIE*, vol. 5370, pp. 494–505, 2004.

[26] A. Jalba, M. Wilkinson, and J. Roerdink, "Cpm: a deformable mode for shape recovery and segmentation based on charged particles," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 10, pp. 1320–1335, Oct. 2004.

[27] Y. Shi and W. C. Karl, "Differentiable minimin shape distance for incorporating topological priors in biomedical imaging," in *IEEE Int. Symp. Biomed. Imag.*, 2004, pp. 1247–1250.

[28] O. Alexandrov and F. Santosa, "A topology-preserving level set method for shape optimization," *J. Comput. Phys.*, vol. 204, pp. 121–130, 2005.

[29] G. Sundaramoorthi and A. J. Yezzi, "More-than-topology-preserving flows for active contours and polygons," in *Proc. ICCV*, 2005, pp. 1276–1283.

[30] G. Sundaramoorthi and A. Yezzi, "Global regularizing flows with topology preservation for active contours and polygons," Tech. Rep. Georgia Inst. Technol., Atlanta, 2005 [Online]. Available: http://users.ece.gatech.edu/~ganeshs

[31] J. O'Hara, "Energy of a knot," *Topology*, vol. 30, no. 2, pp. 241–247, 1991.

[32] M. H. Freedman, Z. X. He, and Z. Wang, "Mobius energy of knots and unknots," *Ann. Math.*, vol. 139, no. 1, pp. 1–50, 1994.

[33] J. O'Hara, "Energy of knots," in *Ideal Knots*, A. Stasiak, V. Katritch, and L. Kauffman, Eds. Singapore: World Scientific, 1998, pp. 288–314.

[34] A. Abrams, J. Cantarella, J. Fu, M. Ghomi, and R. Howard, "Circles minimize most knot energies," *Topology*, vol. 42, pp. 381–394, 2003.

[35] G. Bertrand, "Simple points, topological numbers and geodesic neighborhoods in cubic grids," *Pattern Recognit. Lett.*, vol. 15, pp. 1003–1011, 1994.

Ganesh Sundaramoorthi, photograph and biography not available at the time of publication.

Anthony Yezzi, photograph and biography not available at the time of publication.