Self-repelling snakes for topology-preserving segmentation models

Carole Le Guyader and Luminita Vese

UCLA C.A.M. Report 07-20, 2007

IRMAR, UMR CNRS 6625
Institut National des Sciences Appliquées de Rennes
20 Avenue des Buttes de Coesmes CS 14315 35043 RENNES Cedex, France
cleguyad@insa-rennes.fr

UCLA Department of Mathematics
405 Hilgard Avenue
Los Angeles, CA 90095-1555, U.S.A.
lvese@math.ucla.edu

Abstract

The implicit framework of the level set method has several advantages when tracking propagating fronts. Indeed, the evolving contour is embedded in a higher-dimensional level-set function and its evolution can be phrased in terms of an Eulerian formulation. The ability of this modeling to handle topological changes (merging and breaking), its being parameter-free and intrinsic make it useful in a wide range of fields (fluid mechanics, computer vision) and particularly in image segmentation, the main subject of this paper. Nevertheless, in some applications, this topological flexibility turns out to be undesirable, for instance, when the shape to be detected has a known topology, when the resulting shape must be homeomorphic to the initial one. The necessity of designing topology-preserving processes arises in medical imaging. An example arises in the human cortex reconstruction. It is known that the human cortex has a spherical topology so throughout the reconstruction process this topological feature must be preserved. We thus propose in this paper a segmentation model based on an implicit formulation and on the geodesic active contours, in which a topological constraint is enforced.

Keywords: geodesic active contour, level set method, topology preservation, AOS scheme.

1 Introduction

Segmentation processing consists mainly in detecting and visualizing the boundaries of the objects contained in the image. One of the well-known tools to perform this partition of the image is the active contour model (also called snake model, or deformable model), as proposed in the work by Kass, Witkin, Terzopoulos [17]. Starting with a contour that encloses the object to be detected, deformable models aim at evolving this contour so that it matches the edge of the considered object.
The shape taken by the contour throughout the process is related to an energy-minimization problem. The energy functional which is minimized is a weighted combination of internal forces (describing properties of elasticity and rigidity of the contour) and external forces (that guide the curve towards the boundary of the object).

In this parametric model [17], the functional introduced is non-intrinsic since it depends on the parameterization and is not related to the geometry of the object to be detected. Besides, in this framework, topology is maintained by the Lagrangian formulation.

Extensive work has been done in a similar way by other contributors such as Cohen et al. [9], Cohen [8] or McInerney et al. [20], [4] to improve the overall accuracy of the modeling.

In Caselles et al. [5], Kichenassamy et al. [29], the authors focus on a particular active contour model (called ‘Geodesic Active Contours’) and prove that it is equivalent to finding a geodesic curve in a Riemann space whose metrics is linked to the image content.

To establish the evolution equation of the contour, a family of curves depending on a parameter \( t \geq 0 \) representing time is defined. The derivative of the functional with respect to \( t \) is computed and one seeks the direction for which the energy decreases the most rapidly.

The level set formulation can be derived straightforwardly using the theory of curve evolution (see [23], [21], [34] for instance). The implicit representation of the curve via a level set function avoids parameterization issues. The discretization of the problem is made on a fixed rectangular grid. The model is intrinsic (if we define a new parameterization, the energy is unchanged), allows for corners and the initial condition does not have necessarily the same topology as the resulting contour. The geometric properties of the contour (curvature, unit normal vector) are easily deduced from the level set function.

The topological flexibility is a great benefit since it allows for the simultaneous detection of several objects in the image, which was not possible in the case of parametric deformable models without additional modifications or additional work. However, in some applications, this flexibility is undesirable: for instance, when the final shape should be homeomorphic to the initial one, when the topology of the final contour should be consistent with the known topology of the object. In brain imaging, for example, when located near a lobe, parts of the contour may fold up and finally have a contact point (without merging), which is not consistent with the spherical topology of the brain. The contrast of the image can also be impaired and thus can increase the number of components of the object.

These remarks motivated the work presented next: to keep the benefits of the level set framework (parameter-free, intrinsic, easily-deduced geometrical properties of the front, etc.), we propose a new non-parametric topology-constrained segmentation model based on geodesic active contours [5], [29].

1.1 Prior related work

The work of Han, Xu, and Prince Our work is much motivated by the work [16], [14], [15] where Han, Xu and Prince propose a first approach to preserve the topology of the implicit contour while the embedding level set function is evolving. The key idea of the model lies in the concept of simple points from digital topology. The authors assume that the topology of the zero level set is equivalent to the topology of the boundary of the digital object it defines. The topology-preservation problem is therefore simplified in the following way: the topology of the implicit contour can change only if the level set function changes sign at a grid point. This is only a necessary condition: not every change of sign of the level set function implies a topology change of the zero level set and consequently of the digital object boundary. One
needs to introduce the concept of simple point to identify grid points for which, when removed or added, topology is no longer preserved: a grid point is said to be simple if it can be added or removed without changing the topology of the object. The computation of two topological numbers is sufficient to characterize simple points. Thus Han et al. introduce an algorithm that at each iteration monitors the change of signs of the level set function and prevents the level set function from changing sign on grid points which are not simple. The procedure is therefore pixel-based. The authors also stress the arbitrariness of the result produced by the algorithm depending on the order in which points are treated in the narrow band. Their proposed method has been applied to the geodesic active contour model [5]. Here, we propose to define, in the continuum case, a geodesic-active-contour-based segmentation model that globally integrates a topological constraint emerging from a geometrical remark.

The work of Alexandrov and Santosa Our work is also much motivated by the interesting work [1], which is a curve evolution method based on level sets for shape optimization problems arising in material science. The signed distance function is used, and their proposed method avoids overlaps of the narrow band of the evolving contour. Their method, like ours, is a variational one, but it has not been proposed for image segmentation and only artificial tests in shape optimization have been illustrated in [1]. In a shape optimization problem, their model minimizes $F_\mu(\Phi) = F(\Phi) + \mu H(\Phi)$, with $\mu << 1$, where $F(\Phi)$ is a general shape optimization functional and $H(\Phi)$ is the topological constraint term, defined by

$$H(\Phi) = -\int_{\partial D} \log \left[ \Phi(x + d\nabla \Phi(x)) \right] ds - \int_{\partial D} \log \left[ -\Phi(x - l\nabla \Phi(x)) \right] ds,$$

where $D = \{ x : \Phi(x) > 0 \}$ and $d > 0$, $l > 0$ are given parameters. The unknown level set function $\Phi$ must be a signed distance function.

The work of Sundaramoorthi and Yezzi Another related and very interesting approach based on the knot energy is presented in [36], [37]. This also includes a penalty in the segmentation functional in the form of a double integral, that could prevent a curve $C$ from changing topology. The relevant term that is minimized inspired from the knot energy is

$$E(C) = \frac{1}{2} \int_{C \times C} \frac{dp dp'}{\|C(p) - C(p')\|^\gamma},$$

where $\| \cdot \|$ is the usual Euclidean norm and $\gamma > 0$. The topology depends on the choice of the parameter $\gamma$. This term is also coupled with the method from [16], [15] to insure that topology is always preserved. Experimental results similar to ours are presented in [36] and [37].

The work of Rochery, Jermyn and Zerubia This related work [25], [26], [27], although devoted to a different problem, uses an idea similar to ours, idea which consists in preventing pieces of the same curve from coming in contact, merging or breaking. Their method is devoted to motion and evolution of thin long objects, with applications to the automatic extraction of road networks in remote sensing images. The authors in [25], [26], [27] have proposed interesting regularization terms on a curve $C(p)$, $p \in [0,1]$ that are not local. In [25], these are of the form $E(C) = -\int_0^1 \int_0^1 \tilde{t}(p) \cdot \tilde{t}(p') \Psi(||C(p) - C(p')||) dp dp'$, where $\tilde{t}(p)$, $\tilde{t}(p')$ are the tangents to the curve at points $C(p), C(p')$, and $||C(p) - C(p')||$ is again the
Euclidean distance between curve points $C(p), C(p')$. The function $\Psi$ is chosen to be $\Psi(l) = \sinh^{-1}(1/l) + l - \sqrt{1+l^2}$, thus decreasing on $[0, \infty)$. Other non-local forms are considered as well, and geometric motions of thin long objects are obtained. As in our case, the implicit representation by level sets is used for the implementation in [25], [27]. In [26], the authors continue their ideas, but this time in the phase-field approach for the representation of the motion. The work [25] is continued with more details in [27]. However, although related, our proposed approach is different in several aspects, as it will be seen later.

Work of Cecil et al. In a preliminary work, the PhD Thesis of Cecil [6] (in collaboration with S. Osher and the second author), is also devoted to this problem, with satisfactory partial results based on the active contour model without edges [7]. The penalty is defined in a completely different way, but for more details we refer the reader to [6].

Finally, for completeness, we also refer to other segmentation methods by variational approaches and active contours using coupled curve evolution equations, even if these are not directly related to the main ingredients of the proposed approach. These include the work of Zhao et al. [43], Samson et al. [30], [31], of Chan and the second author [40], of Paragios-Deriche [24], of Yezzi, Tsai and Willsky [42], [39], of Dufour et al. [13], among others.

The outline of the paper is as follows: Section 2 is devoted to the depiction of the model. Section 3 presents the associated evolution problem (with its derivation details given in the Appendix A, and its numerical discretization in the Appendix B). We conclude the paper with experimental results on various synthetic and real images, in 2 or 3 dimensions, in Section 4.

## 2 Depiction of the model

Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$, $\partial \Omega$ its boundary and let $I$ be a given bounded image function defined by $I : \Omega \to \mathbb{R}$. For the purpose of illustration we consider $n = 2$.

**Review of the geodesic active contour model** We first recall one of the main ingredients of the proposed model, the geodesic active contour model [5], [29] for boundary detection using the gradient of the image data as an edge detector. Let us denote by $C$ the set of closed parameterized curves such that: $C = \{C : [0, 1] \to \mathbb{R}^2 \text{ piecewise } C^1, C(0) = C(1)\}$ (these curves are parameterized by $q \in [0, 1]$). In [5], [29] the authors introduce the following minimization problem:

$$
\min_{C \in \mathcal{C}} J(C) = \min_{C \in \mathcal{C}} \int_0^1 g(|\nabla I(C(q))|)|C'(q)|dq,
$$

where $g$ is an edge-detector function satisfying: $g : [0, +\infty) \to [0, +\infty], g(0) = 1$, $g$ strictly decreasing and $\lim_{r \to +\infty} g(r) = 0$. This can be minimized using gradient descent by constructing a family of curves $C(t, \cdot)$ for $t \geq 0$, such that $C(0, \cdot)$ is an initial curve, and as $t$ increases, $J(C(t, \cdot))$ decreases in time to a local minimizer.

The modeling is then cast in the level set setting [23], [11], [12]. The evolving contour is embedded in a higher-dimensional Lipschitz continuous function $\Phi$ defined by $\Phi : \Omega \times [0, +\infty] \to \mathbb{R}$, with $(x, t) \mapsto \Phi(x, t)$, such that $C(t, \cdot) = \{x \in \Omega \mid \Phi(x, t) = 0\}$ and

- $\Phi < 0$ on $\omega$ the interior of $\mathcal{C}$,
- $\Phi > 0$ on $\Omega \setminus \overline{\omega}$.
The function $\Phi$ is preferred to be a signed-distance function for the stability of numerical computations (over time, one may periodically reinitialize the level set function $\Phi$ as the signed-distance function to the set $\{ x \in \Omega \mid \Phi(x) = 0 \}$, see [38], [28] for instance).

Denoting by $\delta$ the one-dimensional Dirac measure, the length of the zero level set is given by: $\int_{\Omega} \delta(\Phi(x))|\nabla \Phi(x)|dx$. We could have phrased straightforwardly the segmentation problem (1) in the level set framework by minimizing the functional $F$ defined by:

$$F(\Phi) = \int_{\Omega} g(|\nabla I(x)|)\delta(\Phi(x))|\nabla \Phi(x)|dx.$$  (2)

Computing the Euler-Lagrange equation and using a gradient-descent method (similar to [43], [7]), we would have then obtained the following evolution equation:

$$\frac{\partial \Phi}{\partial t} = \delta_x(\Phi)\text{div}(g(|\nabla I|)\frac{\nabla \Phi}{|\nabla \Phi|}),$$  (3)

$\delta_x$ being a $C^\infty$ regularization of the Dirac measure. Then the scale factor $\delta_x(\Phi)$ is replaced by $|\nabla \Phi|$, to extend the motion uniformly to all level sets of $\Phi$.

**Description of the proposed topological constraint problem**  The segmentation method we propose is based on the geodesic-active-contour model and includes a topological constraint based on a geometrical observation. In the following, we assume that $\Phi$ is a signed-distance function to $C$:

$$\Phi(x) = \begin{cases} 
\text{dist}(x,C) & \text{if } x \in \Omega - \bar{\omega} \\
-\text{dist}(x,C) & \text{if } x \in \omega 
\end{cases}.$$  (4)

In that case, $\Phi$ satisfies $|\nabla \Phi| = 1$. The gradient $\nabla \Phi$ is perpendicular to the isocontours of $\Phi$, points in the direction of increasing $\Phi$ so the unit outward normal vector to the zero level line at point $x$ is merely $\nabla \Phi(x)$ (more generally, the unit outward normal vector to the $k$ level line at point $x$ is $\nabla \Phi(x)$). Also, the set $\{ x \in \Omega \mid \Phi(x) = l \}$ with $l$ a real number is obtained by shifting the points from the zero level line $C$ by a quantity $l$ in the outward normal to $C$, that is

$$\{ x \in \Omega \mid \Phi(x) = l \} = \{ x + l\nabla \Phi(x), x \in C \}.$$  (5)

Let us now consider two points $(x, y) \in \Omega \times \Omega$ belonging to the zero level line of $\Phi$ close enough to each other and let $\nabla \Phi(x)$ and $\nabla \Phi(y)$ be the unit outward normal vectors to the contour at these points. As shown in Figure 1, when the contour is about to merge, split or have a contact point (that is when the topology of the evolving contour is to change), then $\langle \nabla \Phi(x), \nabla \Phi(y) \rangle \simeq -1$.

Instead of working with the only points of the zero level line, we propose to focus on the points contained in a narrow band around the zero level line, more precisely, on the set of points $\{ x \in \Omega \mid -l \leq \Phi(x) \leq l \}$, $l$ being a level parameter, and we plan to minimize the following functional:

$$F(\Phi) + \mu E(\Phi),$$  (5)
Figure 1: Geometrical characterization of points in a zone where the curve is to split, merge or have a contact point.

$F$ being the functional of the geodesic-active-contour model (2) and $E$ the functional defined by:

$$E(\Phi) = -\int_{\Omega} \int_{\Omega} \exp\left(-\frac{||x - y||^2}{d^2}\right) < \nabla \Phi(x), \nabla \Phi(y)> H(\Phi(x) + l)H(l - \Phi(x))$$

$$\cdot H(\Phi(y) + l)H(l - \Phi(y))dxdy$$

$$= \int_{\Omega} \int_{\Omega} G(||x - y||^2) g(\nabla \Phi(x), \nabla \Phi(y)) h(\Phi(x)) h(\Phi(y)) dxdy. \quad (6)$$

The potential $G$ such that $G(||x - y||^2) = \exp\left(-\frac{||x - y||^2}{d^2}\right)$ measures the nearness of the two points $x$ and $y$. Here $H$ denotes the one-dimensional Heaviside function and therefore $h(\Phi(x)) = H(\Phi(x) + l)H(l - \Phi(x)) = \chi_{\{x \in \Omega, -l \leq \Phi(x) \leq l\}}(x)$. In the calculations, the Heaviside function being not differentiable, it is substituted by a smooth version denoted by $H_\epsilon$. At last, setting $g(\nabla \Phi(x), \nabla \Phi(y)) = -\nabla \Phi(x), \nabla \Phi(y) >$, if the unit outward normal vectors to the level lines passing through $x$ and $y$ have opposite directions, then the functional is not minimal. In the following, we will use the notations $g = g(z_1, z_2, z_3, z_4)$ and we will denote by $\nabla_{(z_1, z_2)}g = (g_{z_1}, g_{z_2})^T$ and by $\nabla_{(z_3, z_4)}g = (g_{z_3}, g_{z_4})^T$. Here $\mu > 0$ is a tuning parameter (note that we use the same notation $g(||\nabla I||)$ for the edge function in (1), and also for the function $g(z_1, z_2, z_3, z_4)$ in (6), but we think that there will be no confusion).

3 Evolution problem

We give in Appendix A the details of the derivation of the Euler-Lagrange equation associated with the minimization problem $\min_{\Phi} F(\Phi) + \mu E(\Phi), \mu > 0$. Parameterizing the gradient descent direction by an artificial time $t \geq 0$, the Euler-Lagrange equation in $\Phi(x, t)$ with $\Phi(x, 0) = \Phi_0(x)$ (signed-distance function to the initial contour) is:

$$\frac{\partial \Phi}{\partial t} = \delta_\epsilon(\Phi) \text{div} \left(g(||\nabla I||) \frac{\nabla \Phi}{|\nabla \Phi|}\right) + 4 \frac{\mu}{d^2} H_\epsilon(\Phi(x) + l)H_\epsilon(l - \Phi(x))$$

$$\cdot \int_{\Omega} \left\{ (x_1 - y_1) \frac{\partial \Phi}{\partial y_1}(y) + (x_2 - y_2) \frac{\partial \Phi}{\partial y_2}(y) \right\} e^{-\frac{||x - y||^2}{d^2}} H_\epsilon(\Phi(y) + l)H_\epsilon(l - \Phi(y)) dy. \quad (7)$$

We replace $\delta_\epsilon(\Phi)$ by $|\nabla \Phi|$ (rescaling stage, see Zhao et al. \cite{44} and Alvarez et al. \cite{2}) and consider Neumann homogeneous boundary conditions: $\frac{\partial \Phi}{\partial \tilde{v}} = 0$ on $\partial \Omega$ with $\tilde{v}$ denoting the
unit outward normal to the boundary of \( \Omega \). The speed of convergence can be increased by adding the component \( kg(\|\nabla I\|)\nabla \Phi \) in the evolution equation, \( k \) being a constant. Thus we consider the following evolution problem:

\[
\begin{aligned}
\frac{\partial \Phi}{\partial t} &= |\nabla \Phi| \left[ \text{div}(g(\|\nabla I\|)\frac{\nabla \Phi}{|\nabla \Phi|}) + kg(\|\nabla I\|) \right] \\
+ &4\frac{\mu}{d^2} H_e(\Phi(x) + l) H_e(l - \Phi(x)) \int_{\Omega} \left\{ (x_1 - y_1) \frac{\partial \Phi}{\partial y_1}(y) + (x_2 - y_2) \frac{\partial \Phi}{\partial y_2}(y) \right\} e^{-\frac{|x-y|^2}{d^2}} \\
\cdot H_e(\Phi(y) + l) H_e(l - \Phi(y)) dy.
\end{aligned}
\]

\[\Phi(x, 0) = \Phi_0(x),\]

\[\frac{\partial \Phi}{\partial v} = 0, \text{ on } \partial \Omega.\]

**Remark:**
Let us consider the zero level line \( C \). A calculation gives that \( C \) moves according to:

\[
\frac{\partial C}{\partial t} = g(\|\nabla I\|) \kappa \hat{n} - \nabla g(\|\nabla I\|) \cdot \hat{n} + kg(\|\nabla I\|) \hat{n}
\]

\[+ \frac{\mu'}{|\nabla \Phi|} \left[ \int_{\Omega} < x - y, \nabla \Phi(y) > e^{-\frac{|x-y|^2}{d^2}} H_e(\Phi(y) + l) H_e(l - \Phi(y)) dy \right] \hat{n}, \quad (8)
\]

\( \hat{n} \) being the unit inward normal vector, \( \kappa \) the curvature. Consider now two points \( x \) and \( y \) close enough to each other, \( x \) belonging to \( C \) and \( y \) belonging to the narrow band of width \( 2l \). In Figure 2, the contour \( C \) is about to merge or to have a contact point. If we compute \( < x - y, \nabla \Phi(y) > \), we get a positive quantity.

If we now take two points \( x \) and \( y \) close enough to each other that both belong to \( C \) (Figure 3), the contribution \( < x - y, \nabla \Phi(y) > \) is almost zero. Lastly, in the case of figure 4, \( < x - y_1, \nabla \Phi(y_1) > \) is positive while \( < x - y_2, \nabla \Phi(y_2) > \) is negative. Globally, the contribution is zero in this part of the narrow band. Thus in critical cases (when the curve is about to change topology), we can expect to have a positive quantity of the component

\[\int_{\Omega} < x - y, \nabla \Phi(y) > e^{-\frac{|x-y|^2}{d^2}} H_e(\Phi(y) + l) H_e(l - \Phi(y)) dy \]

and then the term \( \int_{\Omega} < x - y, \nabla \Phi(y) > e^{-\frac{|x-y|^2}{d^2}} H_e(\Phi(y) + l) H_e(l - \Phi(y)) dy \) \( \hat{n} \) can be interpreted as a repelling force which ensures topology preservation.

**Remark:** We continue here the comparison discussion between our proposed approach and the one proposed by Rochery, Jermyn and Zerubia [25], [26], [27]. First of all, the application is different, devoted to detection of long and thin objects, such as roads in satellite images. Also, the function \( \Psi(\|C(p) - C(p')\|) \) in Rochery et al.’s approach is different in our model. Finally, the main difference lies in the formulation and in the implementation. In our approach, the energy is straightforwardly formulated in the level set framework and the gradient descent PDE in the unknown \( \Phi \) is directly obtained from the Euler-Lagrange equation. In Rochery et al.’s approach, the energy is first formulated in terms of curves, then the gradient descent for the curve evolution is obtained with a speed \( F \). Only after this step the model is cast in the level set setting to implement the evolution, but the speed \( F \) has to be accurately computed and extended to all level lines. This is done in a two-step algorithm:
the speed $F$ is first evaluated in a complicated manner in several steps by explicitly extracting the curve points from the zero-level line of $\Phi$. Interpolation is also needed. Finally, the speed $F$ on the contour is extended to all level lines by solving an additional PDE, re-initialization to the distance function being needed as in our approach. Note that in all variational relevant approaches by Alexandrov-Santosa, Rochery-Jermyn-Zerubia and Sundaramoorthi-Yezzi, the penalty term is a line integral and more difficult to handle in practice, unlike our case where we have a region integral.

Details of the numerical implementation are given in Appendix B.
4 Experimental results

We conclude the paper by presenting several numerical results and comparisons on both synthetic and real images in 2 and 3 dimensions. Our experimental tests and images are similar to those from the related prior work by Han et al. [16], [15] and by Sundaramoorthi and Yezzi [36], [37] and thus can be considered as comparisons as well. In addition, we also show comparisons with images used in these prior works, and we can conclude that our proposed method qualitatively performs in a similar way, but by a different approach. The experiments have been performed on a 2.21GHz Athlon with 1.00 Go of RAM. Parameters are given together with a verification of the topology on the final contours (surfaces), as suggested by one of the unknown referees. We also show comparisons with the geodesic active contour model (without topology constraint), and we will see that the use of the penalty term improves the segmentation. Many of our numerical results are applied to medical images; however, we do not really use any anatomical or physiological information. By the illustrated results, we only wish to demonstrate how the method performs in practice; for more accurate segmentations in specialized medical applications, the methods and the parameters have to be modified and adapted to such applications. For any of the tests, the parameter $\mu$ can be simply set to zero, so that the topology is not constrained.

We start with a synthetic image representing two discs in Figure 5 (similar to tests performed in related prior work by Han et al. [16], [15] and by Sundaramoorthi and Yezzi [36], [37]). We aim at segmenting these two discs while maintaining the same topology throughout the process, which means that we expect to get one path-connected component. The initialization is made with a signed-distance function $\Phi_0$ that encircles the two balls and which is the cone defined by: $\Phi_0(x, y) = \sqrt{(x - 48h)^2 + (y - 50h)^2} - 26$. A comparison is made with the results obtained in the case where no topological constraint is enforced in Figure 6.

In Figure 7, we propose the same type of problem on a real image which is a slice of the brain with two pieces well apart (again, this is an artificial experiment only and the main purpose is just the illustration of the method; we do not claim in this test that we solve an anatomical segmentation problem). The initial condition is still a cone such that the zero level line encircles the two elements. It is defined by: $\Phi_0(x, y) = \sqrt{(x - 72h)^2 + (y - 94h)^2} - 63$.

The third example in Figure 8 was taken from [16] and thus we can compare our proposed method with the result on the same image obtained in [16] (and shown in the last image from Figure 8, courtesy of Xiao Han and Jerry Prince). One can observe that for our test, the two middle fingers touch so with the use of classical geodesic active contours, the evolving contour is going to merge and a hole will then appear (see Figure 9), which is undesirable. With our method, the repelling forces prevent the curve from merging, as in [16, 15].

The method has been tested on complex slices of the brain. We can see that the method enables us to get the details of the brain envelope as in Figure 10 without creating contact points. In Figure 11, this is even more complicated since the slice shows two disconnected parts that are very close to each other.

In Figure 12 we show how we can preserve topology in other cases different from those mentioned above. We start with two curves that segment two cells very close to each other without change of topology. This test is similar to tests presented in related prior work by Han et al. [16], [15] and by Sundaramoorthi and Yezzi [36], [37].

We also show in Figures 13, 14 and 15, 16 additional segmentation results together with comparisons with the geodesic active contour model. These visually show that the use of the topology penalty term improves the segmentation results.
Figure 5: Segmentation of the synthetic image with two disks when topological constraints are applied. The parameters are: $h = 1$, $\tau = 0.5$, $l = 1$, size of the window: 5, $k = 0.3$, $\mu = 0.4$. Iterations 0, 50, 100, 150, 180, 210.

Figure 6: Segmentation of the synthetic image with two disks when no topological constraint is enforced with the same parameters as above: the contour has split into two components. Iterations 0, 50, 100, 140, 170, 180.
Figure 7: Result obtained with the following parameters: $h = 1$, $\tau = 0.4$, $l = 1$, size of the window: 5, $k = 0.1$, $\mu = 0.25$. Iterations 0, 200, 600, 1000, 1200, 1430, 1500.

A last and more extreme 2D synthetic example (again included only for the purpose of illustration) is given in Figure 17. The synthetic considered image is a black crown in a white background. The process is initialized with a curve that is inside the crown. The topological constraints are applied so that the contour does not split into two contours.

We conclude the paper by presenting three dimensional experiments. We first consider two black spheres in a white background as depicted in Figure 18, left. The process is initialized with a larger sphere that encloses the two small spheres. We apply our topology-preserving method and get the final result shown in Figure 18 right.

The second 3D image consists of some slices of the brain. We start with an initial condition such that the zero level set encloses the brain. We apply both the topology-preserving model and the classical geodesic active contour model. Owing to data noise/quality impairment of the image there is a kind of handle in the gray matter, which is not consistent with the known spherical topology. When no topological constraint is enforced, the handle appears on the top right corner in Figure 20, which is not the case when the constraint is integrated in the model, Figure 19.

We end this section with a 3D image representing an entire brain, in Figures 21, 22. We expect to get a shape topologically equivalent to the sphere. The details of the lobes are clearly extracted without contact points, there is no handle or anatomically incorrect structures. The spherical topology is kept through the process.
Figure 8: Segmentation of the hand image taken from [16] by the proposed method with topology constraint. The result is obtained with the following parameters: $h = 1$, $\tau = 0.4$, $l = 1$, size of the window: $7$, $k = -0.2$, $\mu = 0.2$. Iterations 0, 50, 150, 200, 300, 500, 600 and result obtained in [16], [14], [15] by Han, Xu and Prince.

Figure 9: Step of the process when no topological constraint is enforced, with the same parameters. Iterations 0, 500, 780, 890.
4.1 Numerical verification and validation of the topology constraint

For our two-dimensional experiments, only by visual inspection we can see that the topology constraint has been correctly imposed. However, for the three-dimensional results, visual inspection only is not sufficient to verify that the correct desired topology has been achieved.

To make sure of the algorithm efficiency, we have computed for each 3D example the Euler characteristic $\chi$ from the obtained closed surface. The Euler characteristic is a topological invariant which, with the knowledge of the orientability of the considered closed surface, enables to (uniquely) topologically classify this surface.
In the case of polyhedra, it is classically defined as follows:

$$\chi = V - E + F,$$

with \( V \) the number of vertices, \( E \) the number of edges and \( F \) the number of faces.

It is known that for any polyhedron homeomorphic to a sphere, the Euler characteristic is equal to 2. Thus, as we aim at getting surfaces topologically equivalent to a sphere, we can expect to get an Euler characteristic equal to 2 in our experiments.

In the example of the two spheres, we have obtained \( \begin{cases} V = 2732 \\ E = 8190 \\ F = 5460 \end{cases} \), which gives \( \chi = 2 \).
Figure 12: Segmentation of two blood cells that are very close to each other. Iterations 0, 50, 100, 150, 200, 300, 360.

In the example related to Figure 14, we have obtained \( \begin{align*}
V &= 6828 \\
E &= 20478 \\
F &= 13652
\end{align*} \), which gives again \( \chi = 2 \).

In comparison, the computation of the Euler characteristic in the example related to Figure 15 gives \( \begin{align*}
V &= 6062 \\
E &= 18186 \\
F &= 12124
\end{align*} \), that is \( \chi = 0 \). This result is consistent with the theory. The Klein bottle and the torus are the only closed surfaces with Euler characteristic 0. Due to the hole in the top right part of the gray matter, the obtained surface is topologically equivalent to a torus.

We conclude with the image of the whole brain corresponding to Figures 21, 22. We have obtained \( \begin{align*}
V &= 63626 \\
E &= 190872 \\
F &= 127248
\end{align*} \), that is \( \chi = 2 \).

4.2 Discussion on the choice and dependence of parameters

The window size in the Gaussian function is small due to computational cost and this is related to the parameter \( l \) (kept 1 in all our calculations). We choose the window size of \( 5 \times 5 \) or \( 7 \times 7 \), while the parameter \( d \) is always defined as \( d = l + 3 \). The values for \( \mu \) are chosen small, inside the interval \([0.2, 0.4]\), so that we do not have the penalty as the dominant term. Still, with such small \( \mu \) values, the topological constraint is satisfied. The value of \( \Delta t = \tau \) is chosen function of the CFL condition of the numerical scheme, and finally the parameter \( k \) has the
same meaning as in other deformable balloon models (to speed up the evolution, and to force the motion inward or outward). The distance between two close curves is thus dependent on the values of $l$, $d$ and of the Gaussian window size. With mesh refinement near the contour, it would be possible to use smaller such values, so that the curves can touch each other but without changing the topology. This may be desirable for more accurate segmentations of the cortical surface.

5 Conclusion

In this paper, we have addressed the issue of topology preservation in segmentation problems. We have proposed a geodesic-active-contour based model that globally integrates a topological constraint. This topological constraint is based on a geometrical observation and can be seen as a repelling force in the evolution problem. In terms of computational cost, the use of an AOS scheme leads to solve two (three) separate tridiagonal linear subsystems. The computation time could be reduced by parallelizing the code even if in most 2D cases, computation time is of the order of a minute. The topological constraint is easy to compute since we have restricted the integration domain to a small window centered on the current point. Regarding the parameters, the CFL conditions enable us to have an estimate of parameter $\tau$, $k$ and $\mu$. The level parameter $l$ has been chosen equal to 1 in all the tests in order not to constrain too much the model and to preserve the fine details of the boundaries. It may be interesting and useful to apply the proposed penalty term to other segmentation approaches, including region based methods.
Figure 14: Segmentation without topology constraint. Iterations 0, 50, 100, 150, 200, 250, 300.

Acknowledgment

The authors would like to thank Jean-Michel Morel for his very useful suggestions on this project. Also, the authors would like to thank Christian Gout for carefully reading the manuscript and for his very helpful comments. Finally, the authors would like to thank Jerry Prince for suggesting the reference by T. Yezzi.

Appendix A: Derivation of the Euler-Lagrange equations

In the sequel, the Euler-Lagrange equation associated to problem (5) is computed. It can be formally obtained as follows. Let us set \( \Upsilon(\nu) = f(\nu) + \mu e(\nu) \), with \( f(\nu) = F(\Phi + \nu \Psi) \) and \( e(\nu) = E(\Phi + \nu \Psi) \), \( \nu \) being a small parameter and \( \Psi \) a function like \( \Phi \). The minimizer \( \Phi \) will satisfy \( \Upsilon'(0) = 0 = f'(0) + \mu e'(0) \) for all functions \( \Psi \). One has:

\[
e(\nu) = \int_{\Omega} \int_{\Omega} G(||x-y||^2) g(\nabla \Phi(x) + \nu \nabla \Psi(x), \nabla \Phi(y) + \nu \nabla \Psi(y)) h(\Phi(x) + \nu \Psi(x)) \cdot h(\Phi(y) + \nu \Psi(y)) dxdy,
\]

17
Figure 15: Segmentation with topology constraint. Iterations 0, 100, 200, 300 and 400.

and a simple calculus gives

\[
e'(\nu) = \int_{\Omega} \int_{\Omega} G(||x - y||^2_2) g(\nabla \Phi(x) + \nu \nabla \Psi(x), \nabla \Phi(y) + \nu \nabla \Psi(y))
\]

\[
\cdot \left[ h'(\Phi(x) + \nu \Psi(x)) h(\Phi(y) + \nu \Psi(y)) \Psi(x) + h(\Phi(x) + \nu \Psi(x)) h'(\Phi(y) + \nu \Psi(y)) \Psi(y) \right] dx dy
\]

\[
+ \int_{\Omega} \int_{\Omega} G(||x - y||^2_2) h(\Phi(x) + \nu \Psi(x)) h(\Phi(y) + \nu \Psi(y))
\]

\[
\left[ < \nabla (z_1, z_2) g(\nabla \Phi(x) + \nu \nabla \Psi(x), \nabla \Phi(y) + \nu \nabla \Psi(y)), \nabla \Psi(x) >
\right. 
\]

\[
+ \left. < \nabla (z_3, z_4) g(\nabla \Phi(x) + \nu \nabla \Psi(x), \nabla \Phi(y) + \nu \nabla \Psi(y)), \nabla \Psi(y) > \right] dx dy.
\]

Then

\[
e'(0) = \int_{\Omega} \int_{\Omega} G(||x - y||^2_2) g(\nabla \Phi(x), \nabla \Phi(y)) \left[ h'(\Phi(x)) h(\Phi(y)) \Psi(x) + h(\Phi(x)) h'(\Phi(y)) \Psi(y) \right] dx dy
\]

\[
+ \int_{\Omega} \int_{\Omega} G(||x - y||^2_2) h(\Phi(x)) h(\Phi(y)) \left[ < \nabla (z_1, z_2) g(\nabla \Phi(x), \nabla \Phi(y)), \nabla \Psi(x) >
\right. 
\]

\[
+ \left. < \nabla (z_3, z_4) g(\nabla \Phi(x), \nabla \Phi(y)), \nabla \Psi(y) > \right] dx dy.
\]
Figure 16: Segmentation without topology constraint. Iterations 0, 100, 200, 300, 400.

Figure 17: Segmentation of a torus with topology constraint (equivalent with the disk) with the following parameters: $h = 1$, $\tau = 0.5$, $l = 1$, size of the window: 5, $k = -0.3$, $\mu = 0.4$. Iterations 0, 50, 150, 300, 400, 500.
Figure 18: Left, the initialization of the process is made with the function defined by 
\[ \Phi_0(x, y, z) = \sqrt{(x - 25h)^2 + (y - 25h)^2 + (z - 25h)^2} - 18; \] in light blue, the zero level set. Right, in light blue, the final result.

Figure 19: Two views of the final result when the topology constraint is enforced.

We can switch \( x \) and \( y \) in the first integral and we get as first component:

\[
\int_{\Omega} \int_{\Omega} G(||x - y||^2_2)g(\nabla \Phi(x), \nabla \Phi(y))h'(\Phi(x))h(\Phi(y))\Psi(x)dxdy
\]

\[
+ \int_{\Omega} \int_{\Omega} G(||x - y||^2_2)g(\nabla \Phi(y), \nabla \Phi(x))h'(\Phi(x))h(\Phi(y))\Psi(x)dydx.
\]

Assuming that we can interchange the order of integration, we get as first component:

\[
\int_{\Omega} \left\{ \int_{\Omega} G(||x - y||^2_2)g(\nabla \Phi(x), \nabla \Phi(y))h'(\Phi(x))h(\Phi(y))dy \right\} \Psi(x)dx
\]

\[
+ \int_{\Omega} \left\{ \int_{\Omega} G(||x - y||^2_2)g(\nabla \Phi(y), \nabla \Phi(x))h'(\Phi(x))h(\Phi(y))dy \right\} \Psi(x)dx.
\]

For the second component the same reasoning is applied. We switch \( x \) and \( y \) in the second part of the component. Assuming once again that we can interchange the order of integration,
Figure 20: Two views of the final result when no topology constraint is applied.

Figure 21: Top view of the final result.

Figure 22: Front view of the final result.
we have:
\[
\begin{aligned}
&\int_\Omega \left\{ \int_\Omega G(||x-y||_2^2)h(\Phi(x))h(\Phi(y)) < \nabla_{(z_1,z_2)}g(\nabla\Phi(x), \nabla\Phi(y)), \nabla\Psi(x) > dx \right\}dy \\
&+ \int_\Omega \left\{ \int_\Omega G(||x-y||_2^2)h(\Phi(x))h(\Phi(y)) < \nabla_{(z_3,z_4)}g(\nabla\Phi(y), \nabla\Phi(x)), \nabla\Psi(x) > dx \right\}dy.
\end{aligned}
\]

Integrating by parts with respect to \(x\), interchanging the order of integration and setting the necessary boundary conditions to zero, we obtain:
\[
\begin{aligned}
e'(0) &= \int_\Omega \left\{ \int_\Omega G(||x-y||_2^2)g(\nabla\Phi(x), \nabla\Phi(y))h'(\Phi(x))h(\Phi(y))dy \right\}\Psi(x)dx \\
&+ \int_\Omega \left\{ \int_\Omega G(||x-y||_2^2)g(\nabla\Phi(y), \nabla\Phi(x))h'(\Phi(x))h(\Phi(y))dy \right\}\Psi(x)dx \\
&- \int_\Omega \left\{ \int_\Omega \text{div}\left(G(||x-y||_2^2)h(\Phi(x))h(\Phi(y))\nabla_{(z_1,z_2)}g(\nabla\Phi(x), \nabla\Phi(y))\right)dy \right\}\Psi(x)dx \\
&- \int_\Omega \left\{ \int_\Omega \text{div}\left(G(||x-y||_2^2)h(\Phi(x))h(\Phi(y))\nabla_{(z_3,z_4)}g(\nabla\Phi(y), \nabla\Phi(x))\right)dy \right\}\Psi(x) dx,
\end{aligned}
\]

for all \(\Psi\).

Using (3), we deduce that the obtained Euler-Lagrange equation formally is:
\[
\begin{aligned}
&- \delta_\epsilon(\Phi) \text{div}(g(|\nabla I|)\frac{\nabla\Phi}{|\nabla\Phi|}) \\
&+ \mu \left\{ \int \Omega \left\{ \left[ g(\nabla\Phi(x), \nabla\Phi(y)) + g(\nabla\Phi(y), \nabla\Phi(x)) \right] h'(\Phi(x))h(\Phi(y))dy \right\} \\
&- \int \text{div}\left(G(||x-y||_2^2)h(\Phi(x))h(\Phi(y))\nabla_{(z_1,z_2)}g(\nabla\Phi(x), \nabla\Phi(y))\right)dy \\
&- \int \text{div}\left(G(||x-y||_2^2)h(\Phi(x))h(\Phi(y))\nabla_{(z_3,z_4)}g(\nabla\Phi(y), \nabla\Phi(x))\right)dy \right\} = 0,
\end{aligned}
\]

with the associated boundary conditions.

After some intermediate computations, we eventually obtain:
\[
\begin{aligned}
2H_\epsilon(\Phi(x) + l)H_\epsilon(l - \Phi(x)) \int - \frac{2\mu}{d^2} \left\{ (x_1 - y_1)\frac{\partial\Phi}{\partial y_1}(y) + (x_2 - y_2)\frac{\partial\Phi}{\partial y_2}(y) \right\} e^{-\frac{|x-y|^2}{\sigma^2}} \\
H_\epsilon(\Phi(y) + l)H_\epsilon(l - \Phi(y))dy - \delta_\epsilon(\Phi) \text{div}(g(|\nabla I|)\frac{\nabla\Phi}{|\nabla\Phi|}) = 0.
\end{aligned}
\]

**Appendix B: Description of the numerical algorithm**

For the discretization stage, we use an Additive Operator Splitting scheme (AOS), and we follow Weickert et al. (see [41] for more details), which allows for the decomposition of the initial two-dimensional problem into two one-dimensional subproblems. This splitting-based scheme is easy to implement and requires a linear computational cost at each step.
To simplify the notation, we use a vectorial representation of the function $\Phi$ via a concatenation of the rows (respectively columns) of the associated matrix. Thus $\Phi \in \mathbb{R}^{N \times M}$ where $N$ is the number of lines and $M$ the number of columns. The center of gravity of a pixel $i$ coincides with the node of the mesh of coordinates $x_i$. Let us denote by $h$ the space step (assumed to be equal to 1) and $\tau$ the time step. Let $\Phi^n$ be an approximation of $\Phi(x_i, t_n) = \Phi(x_i, n\tau)$. The considered discretization of the problem is the following one:

$$
\Phi^{n+1} = \frac{1}{2} \sum_{l \in \{x, y\}} (Id - 2\tau A_l(\Phi^n))^{-1} \left[ \Phi^n + 4\frac{\mu \tau}{d_s^2} H_s(\Phi^n + l)H_s(l - \Phi^n) \right]
$$

$$
\left\{ \int_{\Omega} < x - y, \nabla \Phi^n(y) > e^{-\frac{|x-y|^2}{d_s^2}}H_s(\Phi^n(y) + l)H_s(l - \Phi^n(y))dy \right\} + \tau \sum_{l \in \{x, y\}} (|\nabla I|)|\nabla \Phi^n|,
$$

with $A_l(\Phi^n)$ defined by:

$$
a_{ij} = \begin{cases} 
\frac{2}{\sum_{m \in \Lambda_l(i)} \left( \frac{|\nabla \Phi^n_i|}{g(|\nabla I|)} \right)_m^n + \left( \frac{|\nabla \Phi^n_j|}{g(|\nabla I|)} \right)_j^n}, & j \in \Lambda_l(i), \\
-\frac{2}{\sum_{m \in \Lambda_l(i)} \left( \frac{|\nabla \Phi^n_i|}{g(|\nabla I|)} \right)_m^n + \left( \frac{|\nabla \Phi^n_j|}{g(|\nabla I|)} \right)_j^n}, & i = j, \\
0 & \text{otherwise},
\end{cases}
$$

and where $\Lambda_l(i)$ represents the neighboring pixels of $i$ with respect to direction $l$.

The AOS scheme can be rewritten by means of the matrices $B_l(\Phi^n) = Id - 2\tau A_l(\Phi^n)$, $l \in \{x, y\}$. We obtain two disconnected linear systems that are both tridiagonal if one reorders the nodes of the mesh in a different way for the second system. The matrices $B_l$ are tridiagonal, strictly diagonally dominant (which ensures uniqueness of the solution of both linear systems thanks to Gershgorin’s theorem for instance) and lead to systems that can be solved computing Thomas algorithm with linear complexity.

Weickert et al. [41] prove that all the coefficients of the matrices $B_l^{-1}(\Phi^n)$ are non-negative and that the sum of the elements of each row of $B_l^{-1}(\Phi^n)$ is equal to one. Thus if we neglect the topological constraints and the component $kg(|\nabla I|)|\nabla \Phi|$, it implies that each element of $\Phi^{n+1}$ is computed from a convex combination of components of $\Phi^n$. The discrete minimum-maximum principle applies and stability is therefore guaranteed for any time step. Regarding the constant motion $kg(|\nabla I|)|\nabla \Phi|$, the function $g(|\nabla I|)$ controls this term and annihilates the action of the force when the curve is approaching edge points. A classical discretization of the gradient can lead to the creation of corners. The entropy condition [34] prevents the curve from propagating where it has already been and thus avoids corner formation. The resulting discretization of the gradient comes from the hyperbolic conservation
laws and can be found in [34]:

\[ |\nabla \Phi_i^n|_2^2 \approx \begin{cases} 
max(D^{-x}\Phi_i^n, 0)^2 + \min(D^{+x}\Phi_i^n, 0)^2 + \max(D^{-y}\Phi_i^n, 0)^2 + \min(D^{+y}\Phi_i^n, 0)^2 & \text{if } k < 0, \\
max(D^{+x}\Phi_i^n, 0)^2 + \min(D^{-x}\Phi_i^n, 0)^2 + \max(D^{+y}\Phi_i^n, 0)^2 + \min(D^{-y}\Phi_i^n, 0)^2 & \text{if } k > 0,
\end{cases} \]

with the classical operators \( D^{\pm x}\Phi_i = \Phi_{i+1} - \Phi_i \) and \( D^{\pm y}\Phi_i = \Phi_i - \Phi_{i-1} \) in one dimension.

As for the topological constraint, we have used a \( C^\infty \) regularization of the Heaviside function [7] defined by:

\[ H_\varepsilon(z) = \frac{1}{2} \left( 1 + \frac{2}{\pi} \arctan \left( \frac{z}{\varepsilon} \right) \right). \]

The partial derivatives are approximated using central finite difference schemes. While the topological constraint introduced in the modeling enables us to use the level set framework to track interfaces whose topology must be preserved, the main drawback is computational cost due to the non-local term appearing in the evolution equation. This limitation can be overcome by slightly simplifying the issue: instead of computing the non-local term over the whole grid, we compute it over a square centered on the current point. The window size depends on the level parameter \( l \). This simplification looks realistic since we are mainly interested in the behavior of \( \Phi \) in a neighborhood of the current point.

Lastly, we have assumed that \( \Phi \) is a signed distance function. Thus we need to periodically apply a reinitialization process. We have used the scheme obtained by Russo and Smereka in [28], defined by:

\[
\Phi_{i,j}^{n+1} = \begin{cases} 
\Phi_{i,j}^n - \frac{\Delta t}{\Delta x} \left( sgn(\Phi_{i,j}^0|\Phi_{i,j}^n| - D_{i,j}) \right) & \text{if } \Phi_{i,j}^0 \Phi_{i-1,j}^0 < 0 \text{ or } \Phi_{i,j}^0 \Phi_{i+1,j}^0 < 0 \text{ or } \Phi_{i,j}^0 \Phi_{i,j-1}^0 < 0 \\
\Phi_{i,j}^n - \Delta t \ sgn(\Phi_{i,j}^0) G(\Phi)_{i,j} & \text{otherwise},
\end{cases}
\]

with \( D_{i,j} \), the distance of node \((i, j)\) from the interface and where, if \( \Phi_{i,j}^0 > 0 \),

\[ G(\Phi)_{i,j} = \sqrt{\max \left( \left( \max(D^- x \Phi_{i,j}, 0) \right)^2, \left( \min(D^+ y \Phi_{i,j}, 0) \right)^2 \right) + \max \left( \left( \max(D^- y \Phi_{i,j}, 0) \right)^2, \left( \min(D^+ x \Phi_{i,j}, 0) \right)^2 \right) - 1}, \]

and if \( \Phi_{i,j}^0 < 0 \),

\[ G(\Phi)_{i,j} = \sqrt{\max \left( \left( \min(D^- x \Phi_{i,j}, 0) \right)^2, \left( \max(D^+ y \Phi_{i,j}, 0) \right)^2 \right) + \max \left( \left( \min(D^- y \Phi_{i,j}, 0) \right)^2, \left( \max(D^+ x \Phi_{i,j}, 0) \right)^2 \right) - 1}. \]

In the three-dimensional case, the numerical scheme is defined in a same manner as in the
two dimensional case:

\[ \Phi_{n+1} = \frac{1}{3} \sum_{l \in \{x,y,z\}} (\text{Id} - 3\tau A_l(\Phi^n))^{-1} \left[ \Phi^n + 4 \frac{\mu T}{\Delta t} H_x(\Phi^n + l)H_x(l - \Phi^n) \right. \]

\[ \left. \left\{ \int_\Omega < x - y, \nabla \Phi^n(y) > e^{-\frac{|x-y|^2}{\alpha^2}}H_x(\Phi^n(y) + l)H_x(l - \Phi^n(y)) dy \right\} + \tau kg(\nabla I)|\nabla \Phi^n| \right], \]

with \( A_l(\Phi^n) \) defined as previously done and \(< \cdot, \cdot >\) the Euclidean scalar product in \( \mathbb{R}^3 \).

References


