

Combination of paths for interactive segmentation

Julien MILLE¹, Sébastien BOUGLEUX² and Laurent D. COHEN³

¹ Université de Lyon, CNRS
Université Lyon 1, LIRIS UMR5202, F-69622, France
<http://liris.cnrs.fr/~jmille>

² Université de Caen-Basse Normandie, CNRS, GREYC, UMR6072
F-14050, Caen, France
<https://bouglex.users.greyc.fr>

³ Université Paris Dauphine, CNRS, CEREMADE, UMR7534
F-75016, Paris, France
<https://www.ceremade.dauphine.fr/~cohen>

Abstract. Active contours and minimal paths have been extensively studied theoretical tools for image segmentation. The recent geodesically linked active contour model, which basically consists in a set of vertices connected by paths of minimal cost, blend the benefits of both concepts. This makes up a closed piecewise-smooth curve, over which an edge or region energy functional can be formulated. As an important shortcoming, the geodesically linked active contour model in its initial formulation does not guarantee the curve to be simple, consistent with respect to the purpose of segmentation. In this paper, we propose to extract a relevant contour from a set of possible paths, such that the resulting structure fits the image data and is simple. Toward this goal, we introduce a novel term to favor the simplicity of the generated contour, as well as a local search method to choose the best combination among possible paths.

1 Introduction

Minimum cost paths for interactive two-phase segmentation provide a solid mathematical background and have proven to find suitable solutions in many practical situations. Provided a set of landmark points typically located along the boundary of interest, paths are computed between these points in order to build a closed contour. Optimality is the main advantage of such approaches, since in most cases, a minimum cost path can be efficiently found as the global solution of the corresponding minimization problem, like Dijkstra's algorithm for shortest paths in directed graphs with no negative cycle. Interactive segmentation methods based on a graph modeling of the image, making them discrete by nature, include the intelligent scissors (or *live-wire*) [18] and their *on-the-fly* extension [11], as well as the riverbed approach [17] based on the image foresting transform [10].

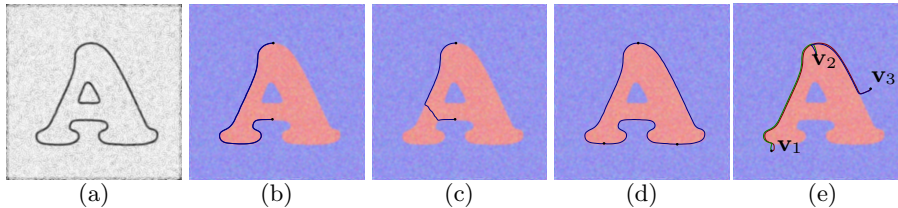


Fig. 1. (a) Potential P (b) Geodesic between two given points. (c) Undesirable shortcut. (d) Piecewise geodesic curve. (e) Undesirable overlapping with unevenly spaced vertices

As regards continuous models, the minimal path method [7] achieved global minimization of the geodesic active contour functional [5], with fixed user-provided endpoints, using the Fast Marching method [20, 21]. An interactive extension was later proposed by [12]. Since the control points are fixed and must be located on the target contour, the minimal path approach does not represent a curve which deforms its shape. Moreover, due to the restricted number of these points, the geodesic may fail to capture a relevant contour if the image is too noisy, not contrasted enough, or if the target contour is too lengthy. While several methods concentrate on avoiding this second drawback [2, 4, 13], the geodesically linked active contour model of [16] allows to overcome the first one, as it can be evolved thanks to a local search method. This latter model combines the advantages of geodesics with the ones of region-based energies, such as the minimal variance term proposed by [6]. Despite its robustness to local minima, this model can fail to construct a simple curve, i.e. without double point, from the initialization step to the end of the evolution.

To overcome this drawback, we propose to generate several relevant paths between landmark points and to select the combination of paths generating the best closed contour. In this extent, we introduce an energy functional, combining contour and region terms with a novel term favoring the simplicity of the curve, penalizing self-overlapping and self-intersections. Our three contributions, i.e. the construction of admissible paths, the simplicity energy as well as the selection of the optimal combination of paths, are described in details. The concepts on which relies the proposed approach are recalled in the following section.

2 Background: geodesic and piecewise-geodesic curves

To extract structures in a given image $\mathbf{I} : \mathcal{D} \rightarrow \mathbb{R}^d$, [7] proposed to find curves of minimal length according to an heterogeneous isotropic metric defined from a potential $P : \mathcal{D} \rightarrow \mathbb{R}^{*+}$. In the context of contour extraction, curves should be located along edges. The potential is thus defined as $P = g + w$, where $g : \mathcal{D} \rightarrow \mathbb{R}^+$ is a decreasing function of the gradient magnitude of \mathbf{I} at some scale s :

$$g(\mathbf{x}) = \frac{1}{1 + \|\nabla G_s * \mathbf{I}(\mathbf{x})\|} \quad (1)$$

where G_s is a Gaussian kernel and $w \in \mathbb{R}^{*+}$ is a regularizing constant. Given two points \mathbf{a} and \mathbf{b} , the curve of minimal length, or *geodesic path* between them,

$$\gamma_{\mathbf{a},\mathbf{b}} = \operatorname{argmin}_{\mathcal{C} \in \mathcal{D}} \left\{ L[\mathcal{C}] = \int_0^1 P(\mathcal{C}(u)) \|\mathcal{C}'(u)\| du \right\} \quad \text{s.t.} \quad \begin{cases} \mathcal{C}(0) = \mathbf{a} \\ \mathcal{C}(1) = \mathbf{b} \end{cases} \quad (2)$$

can be obtained by considering the geodesic distance map, also referred to as the *minimal action map*, $U_{\mathbf{a}}: \mathcal{D} \rightarrow \mathbb{R}^+$ which assigns, to each point $\mathbf{x} \in \mathcal{D}$, the length of the minimal path connecting \mathbf{a} to \mathbf{x} . This map is the unique viscosity solution of the Eikonal equation, which makes it efficiently computable by the Fast Marching (FM) method [20, 21] in $O(N \log N)$ operations, where N is the number of grid points. Once the distance map has been numerically computed, the minimal path between \mathbf{a} and \mathbf{b} can be extracted by a gradient descent on $U_{\mathbf{a}}$, starting from \mathbf{b} until \mathbf{a} is reached. The minimal path approach can fail to extract the desired curve, as depicted in Fig. 1(c). This happens for instance when P is too noisy or not contrasted enough, when the length of the target curve is too important, or when the regularization constant w is too high.

To address this issue, several approaches aim to find a *piecewise-geodesic*, i.e. a set of geodesics connecting pairs of successive landmark points or *vertices* (see Fig. 1(d)). Among them, the geodesically linked active contour (GLAC) model [16] is generated by joining end-to-end geodesic paths built from a set of vertices $\mathcal{V} = \{\mathbf{v}_i\}_{1 \leq i \leq n}$. Given a curve concatenation operator \mathbb{U} ,

$$(\mathcal{C}_1 \mathbb{U} \mathcal{C}_2)(u) = \begin{cases} \mathcal{C}_1(2u) & \text{if } 0 \leq u \leq 1/2 \\ \mathcal{C}_2(2u - 1) & \text{if } 1/2 < u \leq 1 \end{cases}, \quad (3)$$

which is valid only if $\mathcal{C}_1(1) = \mathcal{C}_2(0)$, the closed curve resulting from the assembly of successive geodesics is $\Gamma = \gamma_{\mathbf{v}_1, \mathbf{v}_2} \mathbb{U} \gamma_{\mathbf{v}_2, \mathbf{v}_3} \mathbb{U} \dots \mathbb{U} \gamma_{\mathbf{v}_{n-1}, \mathbf{v}_n} \mathbb{U} \gamma_{\mathbf{v}_n, \mathbf{v}_1}$. It is important to keep in mind that a concatenation of geodesics is not a geodesic itself, thus it is relevant to refer to it as *piecewise-geodesic*. The GLAC approach consists in finding the sequence of n vertices which generates a piecewise-geodesic curve minimizing energy functional E :

$$E[\Gamma] = \omega_{\text{edge}} E_{\text{edge}}[\Gamma] + \omega_{\text{region}} E_{\text{region}}[\Gamma], \quad (4)$$

where ω_{edge} and ω_{region} are user-defined weights. The edge-based energy integrates the edge indicator function g of Eq. (1) along the geodesics, normalized by euclidean length $|\Gamma|$ in order not to penalize lengthy contours. This ensures that short curves, which could be undesirable shortcuts, are not preferred over longer ones.

$$\begin{aligned} E_{\text{edge}}[\Gamma] &= \frac{1}{|\Gamma|} \int_0^1 g(\Gamma(u)) \|\Gamma'(u)\| du \\ &= \frac{1}{|\Gamma|} \sum_{i=1}^n \int_0^1 g(\gamma_{\mathbf{v}_i, \mathbf{v}_{i+1}}(u)) \left\| \gamma'_{\mathbf{v}_i, \mathbf{v}_{i+1}}(u) \right\| du. \end{aligned} \quad (5)$$

In addition, following the piecewise constant model of [6], the region term is expressed as the sum of image variances inside and outside Γ . Since it is not used in the subsequently proposed approach, its mathematical expression is not

detailed. Numerical minimization of E with respect to every \mathbf{v}_i is performed thanks to a greedy algorithm similar in principle to the one proposed in [22]. Considering a local discrete window \mathcal{W}_i centered at each vertex \mathbf{v}_i , greedy evolution is performed, in several passes, by moving vertex \mathbf{v}_i to the position in the window leading to the GLAC having the smallest energy:

$$\mathbf{v}_i^{(t+1)} = \underset{\tilde{\mathbf{v}}_i \in \mathcal{W}_i}{\operatorname{argmin}} E [\gamma_1 \cup \dots \cup \gamma_{i-2} \cup \tilde{\gamma}_{i-1} \cup \tilde{\gamma}_i \cup \gamma_{i+1} \cup \dots \cup \gamma_n],$$

where $\tilde{\gamma}$ are the geodesics, according to Eq. (2), between the current position in the window and the adjacent vertices of \mathbf{v}_i , and γ_i is a shortened form for $\gamma_{\mathbf{v}_i, \mathbf{v}_{i+1}}$. While this model allows to blend the benefits of minimal paths and region-based terms, it turns out to have a significant drawback, as its initial state is not necessarily a simple closed curve. As depicted in Fig. 1(e), this can occur when the initial vertices are unevenly distributed around the target boundary. In this case, geodesics gather on particular sides of the target boundary, as $\gamma_{\mathbf{v}_3, \mathbf{v}_1}$ overlaps $\gamma_{\mathbf{v}_1, \mathbf{v}_2}$ and $\gamma_{\mathbf{v}_2, \mathbf{v}_3}$. The reason is that each geodesic is generated independently from the others, such that the obtained piecewise-geodesic curve does not depend on the visiting order of pairs of adjacent vertices. This undesirable phenomenon may occur either as soon as the geodesically linked contour is initialized, or after several evolution steps on a previously well initialized contour. One could consider imposing hard constraints on the overlapping between paths, or penalizing paths enclosing a region with excessively small area, but the independent construction of paths, which allows parallel implementation, prevents such constraints to be implemented. We address this shortcoming in what follows.

3 Combination of admissible paths

We study a more relevant contour construction preserving the advantages of piecewise geodesic curves. Assuming that a set of several possible relevant paths is available for each pair of successive vertices, we may select a single path from each set and combine these paths in order to build the best boundary curve. The relevancy of the generated contour is measured by an energy functional, combining existing contour and region terms with a novel term penalizing self-tangencies and self-intersections, ensuring the contour to be simple.

Sets of admissible paths Given an ordered pair of successive vertices $(\mathbf{v}_i, \mathbf{v}_{i+1})$, let \mathcal{A}_i be a set of K_i *admissible paths* linking the two vertices,

$$\mathcal{A}_i = \{\gamma_{i,j}\}_{1 \leq j \leq K_i},$$

which we refer to as *admissible set*. To generate these admissible paths, one could first think about solving the K *shortest paths* problem [9, 23]. However, it is unlikely to generate interesting contours in the context of real images, as the $K - 1$ next shortest paths might be identical to the shortest one, up to some very small localized offset, bypassing the shortest path at one place. Moreover, the K *shortest paths* principle does not fit well into the continuous framework

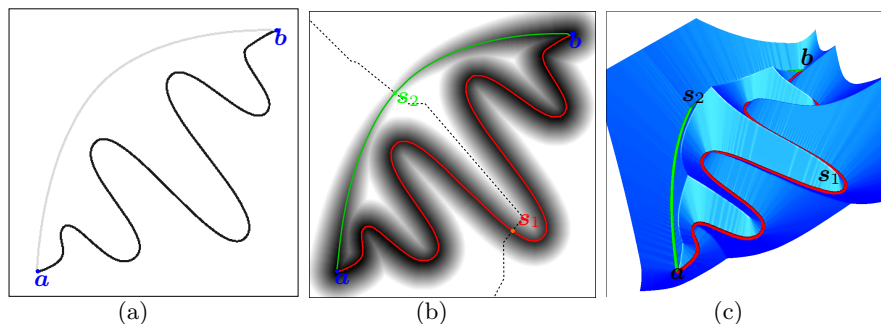


Fig. 2. Saddle points for determining relevant admissible paths between \mathbf{a} and \mathbf{b} : (a) Potential highlights two possible distinctive paths, (b) Action map with two admissible paths with their respective saddle points located halfway and (c) corresponding 3D plot

of minimal paths. A second intuitive technique consists in computing K disjoint paths, as in [15], i.e. preventing each path to pass through points already visited by a previous path. This is achieved by setting $P(\mathbf{x})$ to $+\infty$ for all \mathbf{x} belonging to each newly generated path. However, it would arbitrarily remove sections of contour that could be taken by other relevant paths. Moreover, in case of unsharp contours, low potential areas may get undesirably thick, so that two paths \mathcal{C}_1 and \mathcal{C}_2 may take the same contour section, \mathcal{C}_2 being only an offset curve of \mathcal{C}_1 .

To generate supplemental paths, we propose an approach based on the extraction of *saddle points*. Such critical points were already addressed in the case of minimal action maps, like in [8]. As explained in Section 2, the geodesic linking two points \mathbf{a} and \mathbf{b} is extracted by propagating the minimal action from \mathbf{a} , stop when \mathbf{b} is reached, and perform gradient descent from \mathbf{b} . The opposite may also be done, provided that the path is reversed afterwards. A third way to compute $\gamma_{\mathbf{a},\mathbf{b}}$ consists in propagating **simultaneously** from \mathbf{a} and \mathbf{b} , stopping at the first location where the two fronts meet, performing two gradient descents both sides apart from the meeting location, and assembling the two obtained paths adequately. This principle serves a basis for the generation of multiple paths. When propagation is performed from two source points, yielding the *combined action map* $U_{\mathbf{a},\mathbf{b}} = \min(U_{\mathbf{a}}, U_{\mathbf{b}})$, the two propagation fronts meet at the **saddle points** of $U_{\mathbf{a},\mathbf{b}}$ (see Fig 2(a) and 2(b)). If one intuitively imagine the action as the height in a mountainous area, the saddle points are the mountain passes on the different roads travelling from \mathbf{a} to \mathbf{b} , as depicted in Fig. 2(c). These roads can share common sections, but each road lies in the bottom of a particular valley.

Let $\mathbf{m}_{\mathbf{a},\mathbf{b}} : [0,1] \rightarrow \mathcal{D}$ be the *medial curve*, assumed to be of class C^2 , sweeping along the points geodesically equidistant to \mathbf{a} and \mathbf{b} ,

i.e. $\{\mathbf{x} \mid U_{\mathbf{a}}(\mathbf{x}) = U_{\mathbf{b}}(\mathbf{x})\}$. It forms a crest on the combined action map and contains only critical points, put another way $\nabla U_{\mathbf{a},\mathbf{b}}$ is not defined over $\mathbf{m}_{\mathbf{a},\mathbf{b}}$. However, $U_{\mathbf{a},\mathbf{b}}$ may be differentiated **along** the medial curve, which allows to define the saddle points as the local minima of $U_{\mathbf{a},\mathbf{b}}$ along $\mathbf{m}_{\mathbf{a},\mathbf{b}}$.

In order to extract robust local minima in spite of the effect of discretization, values of the action map are smoothed along $\mathbf{m}_{\mathbf{a},\mathbf{b}}$. Not all saddle points are kept as starting points for path construction. In practice, we limit the number of admissible paths per set below a fixed threshold K_{\max} . As can be seen in Fig. 3(b), many saddle points are intentionally ignored because of their action too high or their proximity to lower saddle points.

The contour resulting from the concatenation of selected paths should exhibit desirable properties in the image. One of these properties is that the generated contour should be simple, i.e. with no multiple point. Instead of imposing simplicity as a hard constraint, which might exclude relevant contours, it is encouraged by an additional energy. Indeed, it is reasonable to allow a certain degree of non-simplicity, e.g. when vertices are located far from the target boundaries, which might cause several admissible paths to have common sections before splitting up, like in Fig. 1(e). The proposed term is based on a distinction between self-tangency and self-intersection, which we present hereafter.

Self-overlap and twisting Consider a curve \mathcal{C} closed and regular, i.e. non-singular in the parametric sense, meaning that its velocity vector \mathcal{C}' vanishes nowhere. If \mathcal{C} is non-simple, it has a number of *multiple points* that should be studied. We only consider *double points*¹, which may be of two kinds. Let $(u, v) \in [0, 1]^2$ s.t. $u \neq v$ be the unordered pair of curve positions identifying a double point: $\mathcal{C}(u) = \mathcal{C}(v)$. On one hand, if (u, v) corresponds to a point of self-tangency, velocity vectors $\mathcal{C}'(u)$ and $\mathcal{C}'(v)$ are colinear. On the other hand, if (u, v) corresponds to a self-intersection - also known as an *ordinary double point* or *crunode* - velocity vectors point towards different directions, making the curve cross itself. This distinction allows to address separately two different defects on curves, which are not necessarily related. A curve with points of self-tangencies will exhibit self-overlapping segments, as depicted in Fig. 4(b), whereas a curve with self-intersections shown in Fig. 4(c) will split the image domain into more than two connected regions.

As regards the first kind of double points, the amount of overlapping is quantified by measuring the length of overlapping curve segments. Considering function ϕ measuring the Euclidean distance between two points on the curve,

$$\phi_{\mathcal{C}}(u, v) = \|\mathcal{C}(u) - \mathcal{C}(v)\|,$$

the zero level set of $\phi_{\mathcal{C}}$, $\mathcal{Z}_{\mathcal{C}} = \{(u, v) \mid \mathcal{C}(u) = \mathcal{C}(v)\}$ is the set of pairs of positions giving equal points (trivially, this set is never empty, since it contains at least

¹ In our framework, curves with points of multiplicity > 2 are detected and excluded from the search

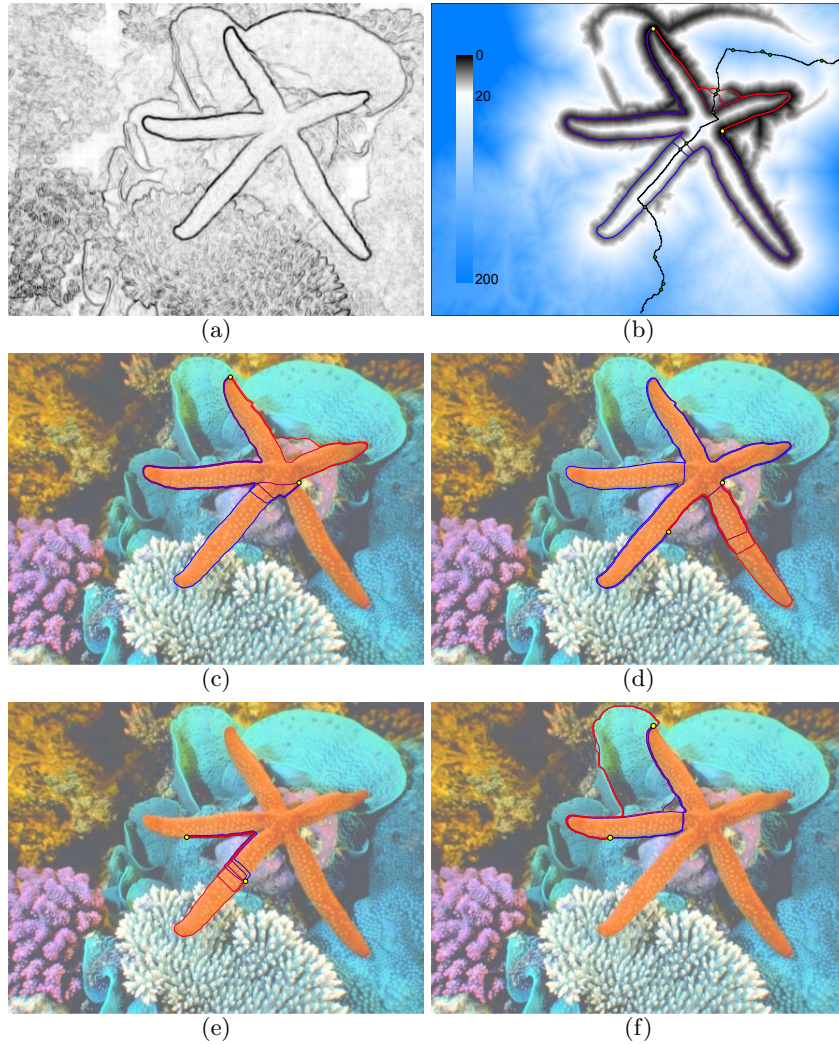


Fig. 3. (a) Potential P (b) Medial curve (black) and saddle points (green) with corresponding paths drawn over combined action map (c-f) Admissible set for each pair of successive vertices with $n = 4$. Paths are sorted (blue to red) according to their exteriority

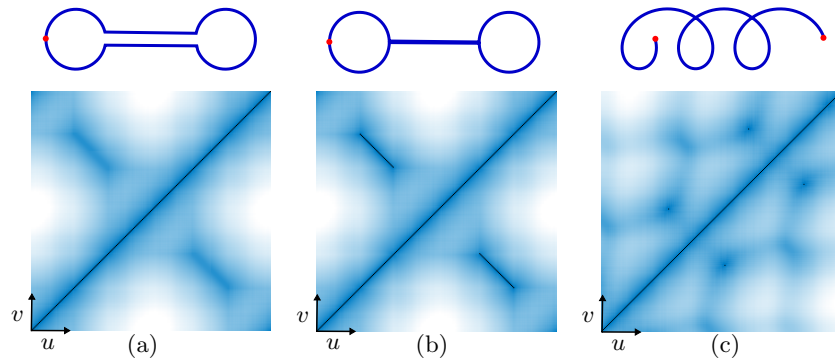


Fig. 4. Distances ϕ_C plotted in (u, v) -space for several types of curve: (a) On a simple closed curve, ϕ_C vanishes only the graph diagonal. (b) On a curve with a section of self-tangency, ϕ_C additionally exhibits two symmetrical zero lines. (c) On a curve with self-intersections, ϕ_C additionally exhibits isolated zeros

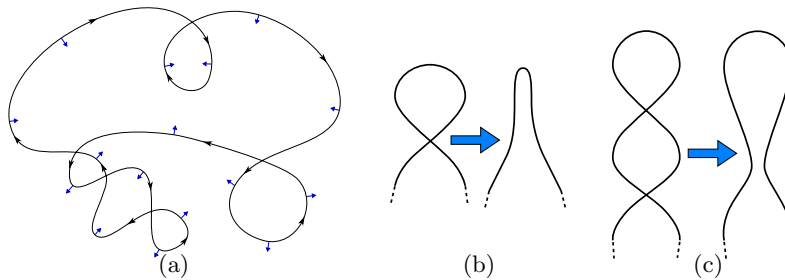


Fig. 5. Self-intersecting curve and inverted segments: (a) A nonsimple curve with only ordinary double points which could be untwisted by (b) Reidemeister move of type I and (c) Reidemeister move of type II

all pairs (u, u) . The length of \mathcal{Z}_C in the (u, v) -space corresponds to the total length of the overlapping segments:

$$|\mathcal{Z}_C| = \int_0^1 \int_0^1 \|\nabla H(\phi_C(u, v))\| \, dudv = \int_0^1 \int_0^1 \delta(\phi_C(u, v)) \|\nabla \phi_C(u, v)\| \, dudv \quad (6)$$

where ∇ is the gradient operator in the (u, v) -space, H is the Heaviside step function and δ is the Dirac delta distribution. It can be proven that for any simple regular curve \mathcal{C} , $|\mathcal{Z}_C| = \sqrt{2}$ ², which can be intuitively understood by observing the plots of ϕ in the curve parameter space in Fig. 4, where dark lines correspond to self-overlapping segments. If the curve is non-simple, $|\mathcal{Z}_C| \geq \sqrt{2}$. Moreover, it is advantageously intrinsic, i.e. independent from parameterization, and invariant to scaling. Note that intersection points may be viewed as overlapping segments of length zero, and thus have no contribution in $|\mathcal{Z}_C|$. Discretizing \mathcal{C} as a 4-connected digital curve, made up of M points in \mathbb{Z}^2 , the length of overlapping segments in Eq. (6) can be computed in $O(M)$ operations.

While tangent double points are used to measure overlapping, ordinary double points will serve as a basis for measuring the amount of *twisting*. Consider a regular closed curve \mathcal{C} oriented clockwise with ordinary double points making self-intersections as shown in Fig. 5(a). It splits the image domain into disjoint subdomains, some of which are demarcated by *inverted segments*, i.e. portions of curves along which the normal vector \mathcal{C}'^\perp points outward (on a simple clockwise-oriented curve, \mathcal{C}'^\perp points inward). We propose to quantify the twisting of \mathcal{C} as the proportion of area demarcated by inverted segments to the total area of \mathcal{C} . This proportion is related to the energy that one would exert to untwist the curve. A self-intersecting curve can be thought of as a planar projection of a knot, i.e. an embedding of the unit circle into Euclidean \mathbb{R}^3 space. Following the theory of knots [1, p. 13] [3], we consider twisted configurations of curves which corresponding knots could be unknotted by Reidemeister moves of type I and II (see Fig. 5(b) and 5(c), respectively). The simple loop, described by a single connected inverted segment and involving a single intersection, is related to the Reidemeister move of type I, whereas the double loop, consisting of two inverted segments and involving two intersections, is related to the Reidemeister move of type II. Let $\text{SL}(\mathcal{C}) \subset [0, 1]^2$ be the set of ordered pairs of curve positions (u, v) s.t. $u < v$ describing intersections involved in single loops, and $\text{DL}(\mathcal{C}) \subset [0, 1]^2 \times [0, 1]^2$ be the set of double ordered pairs $((u_1, v_1), (u_2, v_2))$ s.t. $u_1 < u_2$ and $v_1 > v_2$ describing the couples of intersections involved in double loops. The total inverted area $\mathcal{I}[\mathcal{C}]$ is computed in linear time with respect

² See detailed proof in appendix A

to $|\mathcal{C}|$ using Green's theorem:

$$\begin{aligned} \mathcal{I}[\mathcal{C}] = & - \sum_{(s,t) \in \text{SL}(\mathcal{C})} \min \left(0, \int_s^t \frac{\mathcal{C}^\perp \cdot \mathcal{C}'}{2} du \right) \\ & - \sum_{((s_1,t_1),(s_2,t_2)) \in \text{DL}(\mathcal{C})} \min \left(0, \int_{s_1}^{s_2} \frac{\mathcal{C}^\perp \cdot \mathcal{C}'}{2} du + \int_{t_2}^{t_1} \frac{\mathcal{C}^\perp \cdot \mathcal{C}'}{2} du \right) \end{aligned} \quad (7)$$

where the line integrals involved will be negative for inverted segments.

Optimization The computation of an admissible closed contour can be formulated as determining the sequence of labels $\{x_1, x_2, \dots, x_n\}$ minimizing $E[\gamma_{1,x_1} \uplus \gamma_{2,x_2} \uplus \dots \uplus \gamma_{n-1,x_{n-1}} \uplus \gamma_{n,x_n}]$, where label x_i corresponds to the chosen path in set \mathcal{A}_i . Energy E extends Eq. (4) involved in the GLAC. It is designed to penalize contours exhibiting strongly overlapping sections or self-intersections, poorly fitting to edges, or splitting regions with indistinct color statistics:

$$E[\Gamma] = E_{\text{simplicity}}[\Gamma] + \omega_{\text{edge}} E_{\text{edge}}[\Gamma] + \omega_{\text{region}} E_{\text{region}}[\Gamma]. \quad (8)$$

Weights ω_{edge} and ω_{region} are user-defined parameters controlling the relative significance of the edge and region terms over the simplicity term. This last one involves the normalized self-overlapping and twisting measures defined in Eqs. (6) and (7),

$$E_{\text{simplicity}}[\Gamma] = \frac{|\mathcal{Z}_\Gamma| - \sqrt{2}}{\sqrt{2}} + \frac{1}{|\Omega_{\text{in}}(\Gamma)|} \mathcal{I}[\Gamma]$$

whereas the edge term is defined as in Eq. (5). For the region term, instead of using the piecewise-constant model as in the GLAC, which limits the segmentation to relatively homogeneous objects and backgrounds, we use the Bhattacharyya coefficient between the color probability distributions inside and outside Γ , following [14]:

$$E_{\text{region}}[\Gamma] = \int_{\mathcal{C}} \sqrt{p_{\text{in}}(\Gamma, \boldsymbol{\alpha}) p_{\text{out}}(\Gamma, \boldsymbol{\alpha})} d\boldsymbol{\alpha} \quad (9)$$

Probability distribution functions (PDF), for a given color $\boldsymbol{\alpha}$, are estimated using a Gaussian kernel-based histogram:

$$\begin{aligned} p_{\text{in}}(\Gamma, \boldsymbol{\alpha}) &= \frac{1}{|\Omega_{\text{in}}(\Gamma)|} \int_{\Omega_{\text{in}}(\Gamma)} G_\sigma(\boldsymbol{\alpha} - \mathbf{I}(\mathbf{x})) d\mathbf{x} \\ p_{\text{out}}(\Gamma, \boldsymbol{\alpha}) &= \frac{1}{|\Omega_{\text{out}}(\Gamma)|} \int_{\Omega_{\text{out}}(\Gamma)} G_\sigma(\boldsymbol{\alpha} - \mathbf{I}(\mathbf{x})) d\mathbf{x} \end{aligned}$$

where Ω_{in} and Ω_{out} are the regions inside and outside Γ , respectively. To determine the best sequence of labels $\{x_1, \dots, x_n\}$, a brute-force search would yield an exponential complexity of $O(K_{\text{max}}^n)$. To avoid testing all possible configurations, we propose a greedy search in $O(n^2 K_{\text{max}})$ relying on a specific ordering of paths. In each admissible set \mathcal{A}_i , paths are sorted according to increasing *exteriority* \mathcal{X} , i.e. the signed area, calculated with Green's theorem, formed by a given path \mathcal{C}

and the line segment from $\mathcal{C}(1)$ returning to $\mathcal{C}(0)$ ³:

$$\mathcal{X}[\mathcal{C}] = \frac{1}{2} \int_0^1 \mathcal{C}^\perp \cdot \mathcal{C}' du + \frac{1}{2} \mathcal{C}(1)^\perp \cdot \mathcal{C}(0)$$

If the straight line from $\mathcal{C}(0)$ to $\mathcal{C}(1)$ is taken as a reference horizontal axis, the exteriority is negative (resp. positive) if \mathcal{C} is predominantly below (resp. above) the axis. The vertices being located clockwise, admissible paths are sorted from the innermost to the outermost (see Fig. 3(c-f)). Starting from an initial labelling corresponding to the most interior configuration, i.e. $\{1, \dots, 1\}$, labels are changed according to a local search, by iteratively testing candidate labellings. At each iteration, given a current sequence of labels \mathcal{S} , candidate sequences are tested that differ from a single label from \mathcal{S} , by increasing labels only. For instance, if $\{2, 3, 1\}$ is the current sequence, candidate sequences will be $\{3, 3, 1\}$, $\{2, 4, 1\}$ and $\{2, 3, 2\}$. The candidate sequence leading to the smallest energy, regardless of the current energy, is chosen as the base sequence for the next iteration, while the globally minimal sequence is kept along the iterations. Generating candidate sequences by solely increasing labels makes the contour grow monotonically.

4 Experiments and discussion

We demonstrate the ability of the model to recover closed boundaries of objects in natural color images, given a variable number of user-provided points along the target boundary. Experiments were carried out on the Grabcut dataset [19] in comparison with the original GLAC method [16] (without deformation, i.e. the piecewise-geodesic curve only), in order to show the benefits brought by the use of admissible paths and additional energy terms. To keep a critical look at our contribution, Fig. 6 depicts typical cases where the current approach provides strong global improvements over the GLAC, as well as slight localized improvements, for the same given sets of vertices. As regards the selection of parameters, both methods were evaluated in the most favorable configuration, i.e. parameters such as the regularization weight w - for both methods - or the energy weights ω_{edge} and ω_{region} - specifically for the proposed algorithm -, were tuned separately each time, in order to achieve the most relevant segmentation. The appropriate color space was also chosen for each image, i.e. RGB or the more perceptually uniform Lab, which affects potential P for both methods and the color PDFs involved in the region energy (9).

The BANANA image (row 1) depicts a situation where our approach does not improve segmentation over the GLAC, for this particular configuration of initial points. Despite the complex background and object containing many inner edges, boundaries are well defined and vertices are evenly distributed along the boundary so that the original GLAC manages to extract the object. The FLOWER (row 2) and DOLL (row 3) images are cases where the GLAC exhibits strong

³ See detailed proof in appendix B



Fig. 6. Results on the Grabcut dataset. For each image, left: potential, center: GLAC without deformation (piecewise-geodesic curve), right: combination of paths

overlapping between geodesics when few vertices are provided, although the vertices are reasonably well distributed along the object boundary. Since it does not have any non-overlapping constraint, the GLAC systematically favors portions of contours with the lowest potential. Hence, smooth boundary segments are ignored, not because of their length, but because they may contain sparsely weak edges making the potential increase in small parts of the contour. Conversely, these boundary segments, despite from not being part of the minimal path, make valleys in the distance map and are very likely to be considered as parts of admissible paths by our approach. This proves the proposed method to be inherently less sensitive to weak edges. The CERAMIC (row 4) and TEDDY (row 5) images depicts situations where the GLAC makes *shortcuts* through the object, due to the presence of inner edges stronger than the actual boundaries. Both simplicity and region energies contribute to solve this issue in our algorithm. The former prevents overlapping while the latter favors high discrepancy between inner and outer color distributions, hence avoiding to select the undesirable shortcuts, which would yield to less distinct color histograms than the actual boundaries would. Finally, the MUSHROOM (row 6) combines the issues of inner shortcut and strong self-overlapping.

5 Conclusion

By searching the best paths configuration among sets of admissible paths, given an energy functional combining edge and region data terms with a novel term favoring the simplicity of the curve, we aimed at solving some important issues arising in geodesic-based segmentation, namely shortcuts and self-overlapping or intersecting contours. Comparison against the geodesically linked active contour model, which has similar input and purpose, demonstrated the advantages of the approach. Due to the limited space, an overview of the experiments was given. A thorough evaluation needs to be performed in an extended version of the current paper, including quantification of segmentation accuracy with respect to ground truth segmentations, and evaluation of the sensitivity to several parameters (distribution of initial vertices along the contour, regularization weight, etc.), as well as comparison with related approaches, such as [17].

Appendix

A Self-overlap term

Let \mathcal{C} be a regular curve parameterized over $[0, L]$. Let ϕ be a C^1 function defined over $[0, L]^2$ representing the distance between two positions on the curve:

$$\phi_{\mathcal{C}}(u, v) = \|\mathcal{C}(u) - \mathcal{C}(v)\|^p$$

where p is an arbitrary positive real exponent. The length of the zero level set of $\phi_{\mathcal{C}}$,

$$|\mathcal{Z}_{\mathcal{C}}| = \int_0^L \int_0^L \delta(\phi(u, v)) \|\nabla\phi(u, v)\| \, du dv, \quad (10)$$

quantifies the self-overlap of \mathcal{C} .

Proposition:

If \mathcal{C} is simple, i.e. without self-intersection and self-tangency, then $|\mathcal{Z}_{\mathcal{C}}| = L\sqrt{2}$

Proof:

As a preliminary calculation, let us express the gradient of ϕ (partial derivatives are written using the indexed notation):

$$\begin{aligned} \nabla\phi(u, v) &= [\phi_u(u, v) \quad \phi_v(u, v)]^T \\ &= p \|\mathcal{C}(u) - \mathcal{C}(v)\|^{p-2} \begin{bmatrix} \mathcal{C}'(u) \cdot (\mathcal{C}(u) - \mathcal{C}(v)) \\ -\mathcal{C}'(v) \cdot (\mathcal{C}(u) - \mathcal{C}(v)) \end{bmatrix} \end{aligned}$$

If \mathcal{C} is regular and simple, varying with respect to u in range $[0, L]$, $\phi(u, v)$ is nowhere zero except when $u = v$. Hence, for a fixed v , we have:

$$\delta(\phi(u, v)) = \frac{\delta(u - v)}{|\phi_u(v, v)|} \quad (11)$$

Integrating (11) into (10) and applying the definition of measure δ :

$$\begin{aligned} |\mathcal{Z}_{\mathcal{C}}| &= \int_0^L \int_0^L \delta(\phi(u, v)) \|\nabla\phi(u, v)\| \, du \, dv \\ &= \int_0^L \int_0^L \frac{\delta(u - v)}{|\phi_u(v, v)|} \|\nabla\phi(u, v)\| \, du \, dv \\ &= \int_0^L \frac{\|\nabla\phi(v, v)\|}{|\phi_u(v, v)|} \, dv \end{aligned}$$

Trivially, $\phi(v, v) = 0$. However expanding the gradient gives:

$$\begin{aligned} |\mathcal{Z}_{\mathcal{C}}| &= \int_0^L \frac{p \|\mathcal{C}(v) - \mathcal{C}(v)\|^{p-2} \sqrt{2(\mathcal{C}'(v) \cdot (\mathcal{C}(v) - \mathcal{C}(v)))^2}}{p \|\mathcal{C}(v) - \mathcal{C}(v)\|^{p-2} |\mathcal{C}'(v) \cdot (\mathcal{C}(v) - \mathcal{C}(v))|} \, dv \\ &= \int_0^L \sqrt{2} \, dv \\ &= L\sqrt{2} \end{aligned}$$

B Exteriority term

Let \mathcal{C} be a piecewise-smooth regular curve parameterized over $[0, 1]$,

$$\mathcal{C} : u \mapsto \mathcal{C}(u) = [x(u) \ y(u)]^T.$$

If it is simple and positively oriented such that normal vector \mathcal{C}'^\perp points inward, its inner area may be expressed using Green's theorem:

$$|\Omega_{\text{in}}(\mathcal{C})| = \frac{1}{2} \int_0^1 \mathcal{C}^\perp(u) \cdot \mathcal{C}'(u) \, du = \frac{1}{2} \int_0^1 x(u)y'(u) - x'(u)y(u) \, du$$

When one calculates the previous expression on a non-simple closed curve, one gets the signed area, in which positively and negatively oriented connected components have positive and negative contributions, respectively.

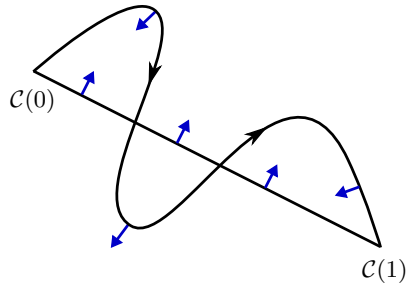


Fig. 7. The exteriority of an open curve is measured as the signed area of the multiple connected region that it forms with the line segment joining its two endpoints

Proposition:

The signed area formed by an open curve \mathcal{C} over $[0, 1]$ and the line segment from $\mathcal{C}(1)$ returning to $\mathcal{C}(0)$ (see Fig. 7), which we use to as the *exteriority* measure in the paper, may be expressed as:

$$\mathcal{X}[\mathcal{C}] = \frac{1}{2} \int_0^1 \mathcal{C}^\perp \cdot \mathcal{C}' du + \frac{1}{2} \mathcal{C}^\perp(1) \cdot \mathcal{C}(0)$$

Proof:

Let S be the parametrization of the line segment joining $\mathcal{C}(1)$ and $\mathcal{C}(0)$, over $[0, 1]$:

$$S(u) = (1 - u)\mathcal{C}(1) + u\mathcal{C}(0)$$

The signed area is then obtained by applying Green's theorem on a piecewise basis:

$$\begin{aligned}
\mathcal{X}[\mathcal{C}] &= \frac{1}{2} \int_0^1 \mathcal{C}^\perp \cdot \mathcal{C}' du + \frac{1}{2} \int_0^1 S^\perp \cdot S' du \\
&= \frac{1}{2} \int_0^1 \mathcal{C}^\perp \cdot \mathcal{C}' du + \frac{1}{2} \int_0^1 ((1-u)\mathcal{C}^\perp(1) + u\mathcal{C}^\perp(0)) \cdot (\mathcal{C}(0) - \mathcal{C}(1)) du \\
&= \frac{1}{2} \int_0^1 \mathcal{C}^\perp \cdot \mathcal{C}' du + \frac{1}{2} \int_0^1 (1-u)\mathcal{C}^\perp(1) \cdot \mathcal{C}(0) - u\mathcal{C}^\perp(0) \cdot \mathcal{C}(1) du \\
&= \frac{1}{2} \int_0^1 \mathcal{C}^\perp \cdot \mathcal{C}' du + \frac{1}{2} \int_0^1 (1-u)\mathcal{C}^\perp(1) \cdot \mathcal{C}(0) + u\mathcal{C}^\perp(1) \cdot \mathcal{C}(0) du \\
&= \frac{1}{2} \int_0^1 \mathcal{C}^\perp \cdot \mathcal{C}' du + \frac{1}{2} \int_0^1 \mathcal{C}^\perp(1) \cdot \mathcal{C}(0) du \\
&= \frac{1}{2} \int_0^1 \mathcal{C}^\perp \cdot \mathcal{C}' du + \frac{1}{2} \mathcal{C}^\perp(1) \cdot \mathcal{C}(0)
\end{aligned}$$

Bibliography

- [1] C.C. Adams. *The Knot Book: An elementary introduction to the mathematical theory of knots*. American Mathematical Society, 2004.
- [2] V.V. Appia and A.J. Yezzi. Active geodesics: Region-based active contour segmentation with a global edge-based constraint. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1975–1980, Barcelona, Spain, 2011.
- [3] V.I. Arnold. Plane curves, their invariants, perestroikas and classification. *Advances in Soviet Mathematics*, 21:33–91, 1994.
- [4] F. Benmansour and L. Cohen. Fast object segmentation by growing minimal paths from a single point on 2D or 3D images. *Journal of Mathematical Imaging and Vision*, 33(2):209–221, 2009.
- [5] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.
- [6] T. Chan and L. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.
- [7] L. Cohen and R. Kimmel. Global minimum for active contour models: a minimal path approach. *International Journal of Computer Vision*, 24(1):57–78, 1997.
- [8] T. Deschamps and L. Cohen. Fast extraction of minimal paths in 3D images and applications to virtual endoscopy. *Medical Image Analysis*, 5(4):281–299, 2001.
- [9] D. Eppstein. Finding the k shortest paths. *SIAM Journal of Computing*, 28(2):652–673, 1998.
- [10] A. Falcão, J. Stolfi, and R. Lotufo. The Image Foresting Transform: Theory, algorithms, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):19–29, 2004.
- [11] A. Falcão, J. Udupa, and F. Miyazawa. An ultra-fast user-steered image segmentation paradigm: Live Wire on the Fly. *IEEE Transactions on Medical Imaging*, 19(1):55–62, 2000.
- [12] O. Gérard, T. Deschamps, M. Greff, and L.D. Cohen. Real-time interactive path extraction with on-the-fly adaptation of the external forces. In *European Conference on Computer Vision (ECCV)*, volume 3, pages 807–821, Copenhagen, Denmark, 2002.
- [13] V. Kaul, A. Yezzi, and Y. Tsai. Detection of curves with unknown endpoints and arbitrary topology using minimal paths. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(10):1952–1965, 2012.
- [14] O. Michailovich, Y. Rathi, and A. Tannenbaum. Image segmentation using active contours driven by the Bhattacharyya gradient flow. *IEEE Transactions on Image Processing*, 16(11):2787–2801, 2007.
- [15] J. Mille, S. Bougleux, and L. Cohen. Minimally overlapping paths sets for closed contour extraction. In *International Conference on Computer Vision Theory and Applications (VISAPP)*, Rome, Italy, 2012.

- [16] J. Mille and L. Cohen. Geodesically linked active contours: evolution strategy based on minimal paths. In *2nd International Conference on Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 5567 of *LNCIS*, pages 163–174, Voss, Norway, 2009. Springer.
- [17] P. Miranda, A. Falcão, and T. Spina. Riverbed: A novel user-steered image segmentation method based on optimum boundary tracking. *IEEE Transactions on Image Processing*, 21(6):3042–3052, 2012.
- [18] E. Mortensen and W. Barrett. Interactive segmentation with intelligent scissors. *Graphical Models and Image Processing*, 60(5):349–384, 1998.
- [19] C. Rother, V. Kolmogorov, and A. Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, 23(3):309–314, 2004.
- [20] J.A. Sethian. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Science*, 93(4):1591–1595, 1996.
- [21] J.N. Tsitsiklis. Efficient algorithms for globally optimal trajectories. *IEEE Transactions on Automatic Control*, 40(9):1528–1538, 1995.
- [22] D.J. Williams and M. Shah. A fast algorithm for active contours and curvature estimation. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 55(1):14–26, 1992.
- [23] J.Y. Yen. Finding the K shortest loopless paths in a network. *Management Science*, 17(11):712–716, 1971.