Image Segmentation by Image Foresting Transform with Non-smooth Connectivity Functions

Lucy A. C. Mansilla*, Paulo A. V. Miranda
Department of Computer Science,
University of S˜ ao Paulo (USP),
05508-090, S˜ ao Paulo, SP, Brazil.
Email: lucyacm@vision.ime.usp.br, pmiranda@vision.ime.usp.br

Abstract—Image segmentation, such as to extract an object from a background, is very useful for medical and biological image analysis. In this work, we propose new methods for interactive segmentation of multidimensional images, based on the Image Foresting Transform (IFT), by exploiting for the first time non-smooth connectivity functions (NSCF) with a strong theoretical background. The new algorithms provide global optimum solutions according to an energy function of graph cut, subject to high-level boundary constraints (polarity and shape), or consist in a sequence of paths’ optimization in residual graphs.

Our experimental results indicate substantial improvements in accuracy in relation to other state-of-the-art methods, by allowing the customization of the segmentation to a given target object.

Keywords—graph cut; image foresting transform; oriented image foresting transform; non-smooth connectivity function; geodesic star convexity;

I. INTRODUCTION

Image segmentation is one of the most fundamental and challenging problems in image processing and computer vision [2]. In medical image analysis, accurate segmentation results commonly require the user intervention because of the presence of structures with ill-defined borders, intensity non-standardness among images, field inhomogeneity, noise, artifacts, partial volume effects, and their interplay [3]. The high-level, application-domain-specific knowledge of the user is also often required in the digital matting of natural scenes, because of their heterogeneous nature [4]. These problems motivated the development of several methods for semi-automatic segmentation [5], [6], [7], aiming to minimize the user involvement and time required without compromising accuracy and precision.

One important class of interactive image segmentation comprises seed-based methods, which have been developed based on different theories, supposedly not related, leading to different frameworks, such as watershed from markers [7], random walks [8], fuzzy connectedness [9], graph cuts [6], distance cut [4], image foresting transform (IFT) [10], and grow cut [11]. The study of the relations among different frameworks, including theoretical and empirical comparisons, has a vast literature [12], [13], [14]. However, these methods in most studies are restricted to undirected graphs, and the most time-efficient methods, including the IFT, present a lack of boundary regularization constraints. Moreover, the quality of their segmentation results with minimal user intervention, strongly depends on an adequate estimate of the weights assigned to the graph’s arcs [15].

The main contribution of this work is a theoretical development to support the usage of non-smooth connectivity functions (NSCF) in the IFT, opening new perspectives in the research of image processing using graphs, since NSCF were, until now, avoided in the literature. More specifically, we prove that some NSCF can lead to optimum results according to a graph-cut measure on a digraph [16], [17] or consist in a sequence of paths’ optimization in residual graphs. We have as main results:

1) The customization of the segmentation by IFT to match the global and local features of a target object: (a) The design of more adaptive and flexible connectivity functions, which allow better handling of images with strong inhomogeneity by using dynamic weights. (b) The orientation of the object’s intensity transitions, i.e., bright to dark or the opposite (boundary polarity), and (c) shape constraints to regularize the segmentation boundary (geodesic star convexity constraint).

2) The development of an interactive segmentation tool within the software, called Brain Image Analyzer (BIA), to support research in neurology involving volumetric magnetic resonance images of a 3T scanner from the FAPESP-CInApCe.

3) Four conference papers were published in international events of high regard [17], [18], [19], [20], and one journal paper was published in the IEEE Transactions on Image Processing (impact factor: 3.111) [16].

For the sake of completeness in presentation, Section II includes the relevant previous work of image segmentation by IFT. In Sections III [5] V and VI we present the main contributions covered in the master’s dissertation [1]: The classification of NSCF, the use of adaptive weights via NSCF, the boundary polarity through digraphs, and the elimination of false delineations by shape constraints. Our conclusions are stated in Section VII.

II. IMAGE FORESTING TRANSFORM (IFT)

An image 2D/3D can be interpreted as a weighted digraph $G = (\mathcal{V}, \mathcal{E}, \omega)$ whose nodes $\mathcal{V}$ are the image pixels in its...
image domain \( \mathcal{I} \subset \mathbb{Z}^N \), and whose arcs are the ordered pixel pairs \((s,t) \in \xi\) (e.g., 4-neighborhood, or 8-neighborhood, in case of 2D images, and 6-neighbors in 3D). The digraph \( G \) is symmetric if for any of its arcs \((s,t)\), the pair \((t,s)\) is also an arc of \( G \). We have an undirected graph when \( \omega(s,t) = \omega(t,s) \) in a symmetric graph \( G \). We use \((s,t) \in \xi \) or \( t \in \xi(s) \) to indicate that \( t \) is adjacent to \( s \). Each arc \((s,t) \in \xi \) may have a weight \( \omega(s,t) \geq 0 \), such as a dissimilarity measure between pixels \( s \) and \( t \) (e.g., \( \omega(s,t) = |I(t) - I(s)| \)) for a single channel image with values given by \( I(t) \).

For a given image graph \( G \), a path \( \pi_t = \langle t_1, t_2, \ldots, t_n = t \rangle \) is a sequence of adjacent pixels with terminus at a pixel \( t \). A path is trivial when \( \pi_t = \{ t \} \). A path \( \pi_t = \pi_s \gg (s,t) \) indicates the extension of a path \( \pi_s \) by an arc \((s,t)\). The notation \( \pi_{s \rightarrow t} = \langle t_1 = s, t_2, \ldots, t_n = t \rangle \) may also be used, where \( s \) stands for the origin and \( t \) for the destination node. A predecessor map is a function \( P \) that assigns to each pixel \( t \) in \( \mathcal{V} \) either another adjacent pixel in \( \mathcal{V} \), or a distinctive marker nil not in \( \mathcal{V} \) — in which case \( t \) is said to be a root of the map. A spanning forest is a predecessor map which contains no cycles — i.e., one which takes every pixel to an origin and \( \mathcal{V} \) is a disjoint union of such sets.

A path-value function computes a value \( f(\pi_t) \) for any path \( \pi_t \), usually based on arc weights. A path \( \pi_t \) is optimum if \( f(\pi_t) \leq f(\pi_{t'}) \) for any other path \( \pi_{t'} \) in \( G \). By taking each pixel \( t \in \mathcal{V} \) as a terminus, we obtain the optimum-path value \( V(t) \), which is uniquely defined by \( V(t) = \min_{\pi_t \in G} \{ f(\pi_t) \} \). A path \( \pi_{t_0} = \{ t_1, t_2, \ldots, t_n \} \) is complete optimum if all paths \( \pi_{t_i} = \{ t_1, t_2, \ldots, t_i \}, i = 1, 2, \ldots, n \) are optimum paths. The IFT takes an image graph \( G \), and a path-cost function \( f \); and assigns one optimum path \( \pi_t \) to every pixel \( t \in \mathcal{V} \) such that an optimum-path forest \( P \) is obtained — i.e., a spanning forest where all paths are optimum. However, \( f \) must be smooth (Definition 1), otherwise, the paths may not be optimum 10.

Definition 1 (Smooth path-cost function). A path-cost function \( f \) is smooth if for any pixel \( t \in \mathcal{I} \), there is an optimum path \( \pi_t \), which either is trivial, or has the form \( \pi_s \gg (s,t) \) where

\[
\begin{align*}
\text{(C1)} & \quad f(\pi_s) \leq f(\pi_t), \\
\text{(C2)} & \quad \pi_s \text{ is optimum, and} \\
\text{(C3)} & \quad \text{for any optimum path } \pi_{t'} \text{ ending at } s, f(\pi_{t'}) \gg (s,t) = f(\pi_t).
\end{align*}
\]

We consider image segmentation from two seed sets, \( \mathcal{S}_o \) and \( \mathcal{S}_b \) (\( \mathcal{S}_o \cap \mathcal{S}_b = \emptyset \)), containing pixels selected inside and outside the object, respectively. The search for optimum paths is constrained to start in \( \mathcal{S} = \mathcal{S}_o \cup \mathcal{S}_b \). The image is partitioned into two optimum-path forests — one rooted at the internal seeds, defining the object, and the other rooted at the external seeds, representing the background. A label, \( L(t) = 1 \) for all \( t \in \mathcal{S}_o \) and \( L(t) = 0 \) for all \( t \in \mathcal{S}_b \), is propagated to all unlabeled pixels during the computation 10.

III. IFT with Non-Smooth Connectivity Functions

Clearly, from Definition 1 we have that a connectivity function is not smooth if it doesn’t satisfy at least one of the conditions C1, C2 or C3. For example \( f_{\Sigma} \) violates C2 and C3:

\[
f_{\Sigma}(\langle t \rangle) = \begin{cases} 0, & \text{if } t \in \mathcal{S}, \\ +\infty, & \text{otherwise.} \end{cases}
\]

In 20, we formally classified several non-smooth functions (Figure 1) according to the conditions C1, C2 and C3 (Definition 1), and C4 (Definition 2).

Definition 2 (Condition C4). A path-value function \( f \) satisfies the condition C4, if for any node \( s \in \mathcal{I} \) the following condition is verified \( \forall t \in \xi(s)\):

- For any paths \( \pi_s \) and \( \pi_{s,t} \) ending at \( s \), if \( f(\pi_s) = f(\pi_{s,t}) \), then we have \( f(\pi_s \gg (s,t)) = f(\pi_{s,t}) \).

Fig. 1. Schematic representation of the relations between smooth and non-smooth connectivity functions: \( C_1 \), \( C_2 \), \( C_3 \), and \( C_4 \) are sets of connectivity functions that satisfy these respective conditions for a general graph.

IV. Adaptive Weights via NSCF

Methods based on IFT 10 have been successfully used in the segmentation of 1.5 Tesla MR datasets 21, 22. However, inhomogeneity effects are stronger in higher magnetic fields, and it is extremely important to define the optimal solution for these images. NSCFs are more adaptive to cope with the problems of field inhomogeneity, which are common in MR images of 3 Tesla 23.

In order to give a theoretical foundation to support the usage of NSCF, we theoretically proved that the IFT with any non-smooth function \( f \in (C_1 \cap C_4) \setminus C_2 \) is, in fact, equivalent to the result of a sequence of optimizations, each of them involving a maximal set of elements, in a well-structured way 20. This proof was supported by the following proposition:

Proposition 1. Consider a function \( f \in (C_1 \cap C_4) \setminus C_2 \). For a given image graph \( G = (\mathcal{V}, \xi, \omega) \), and set of seeds \( \mathcal{S} \), let \( O \) be the set of all pixels \( t \in \mathcal{V} \), such that there exists a complete-optimum path \( \pi_t \) for \( f \). In any spanning forest \( P \) computed in \( G \) by the IFT algorithm for \( f \), all the paths \( \pi_t \) with \( t \in O \) are optimum paths.
In our experiments, we used 10 T1-weighted 3D images of male and female adults with normal brains. The image dataset included the head and, at least, a small portion of the neck. Our experimental result, using a robot user for segmenting the brain dataset, indicates that substantial improvements can be obtained by NSCFs in the 3D segmentation of MR images of 3 Tesla, with strong inhomogeneity effects, when compared to smooth connectivity functions. That is because NSCFs permit a more adaptive configuration of the arc weights.

Figure 2 shows the experimental curves, where IRFC [9] and $PW_{q=2}$ [14] represent different algorithms related to the smooth function $f_{\text{smooth}}$, and we used $\omega(s, t) = G(s) + G(t)$, where $G(s)$ is the magnitude of Sobel gradient at a voxel $s$. Clearly, $f_{\text{lex}}^{E_{\text{direct}}}$ presented the best accuracy. Figure 3 shows one example for user-selected markers. These results emphasize the importance of non-smooth connectivity functions. The non-smooth connectivity function $f_{\text{lex}}^{E_{\text{direct}}}$ is a variation of $f_{\text{lex}}^{E_{\text{direct}}}$ (Eq. 1), in order to guarantee that $f_{\text{lex}}^{E_{\text{direct}}} \in (C_1 \cap C_2) \setminus C_4$. The function $f_{\text{lex}}^{E_{\text{direct}}}$ gives pairs of values that should be compared according to the lexicographical order. The first component is the non-smooth function $f_{\text{lex}}^{E_{\text{direct}}}$ (Eq. 1), and the second is the priority level of the seed/root for that path. The lower its value the higher is its priority. In interactive segmentation, we give lower priority for new inserted seeds, since they are used mainly for corrective actions, so that we can keep their effects more locally. The same process was done for $f_{\text{max}}^{E_{\text{direct}}}$ in $(C_1 \cap C_4) \setminus C_2$ and $f_{\text{lex}}^{E_{\text{direct}}} \in (C_1 \cap C_4) \setminus C_2$, in relation to $f_{\text{max}}^{E_{\text{direct}}}$ and $f_{\text{lex}}$,

respectively [20].

V. BOUNDARY POLARITY VIA NSCF

In order to resolve between very similar nearby boundary segments, in [16], [17] we successfully incorporated the boundary polarity constraint in the IFT using NSCF in digraphs, resulting in a novel method called Oriented Image Foresting Transform (OIFT).

In the case of digraphs, there are two different types of cut for each object boundary: an inner-cut boundary composed by arcs that point toward object pixels

1Method introduced in [24], to simulate user interaction of interactive segmentation.

$C_i(L) = \{(s, t) \in \xi | L(s) = 0, L(t) = 1\}$, and an outer-cut boundary with arcs from object to background pixels $C_o(L) = \{(s, t) \in \xi | L(s) = 1, L(t) = 0\}$. Consequently, we consider two different types of energy, $E_i$ (Eq. 2) and $E_o$ (Eq. 3).

$$E_i(L, G) = \min_{(s, t) \in C_1(L)} \omega(s, t)$$

$$E_o(L, G) = \min_{(s, t) \in C_0(L)} \omega(s, t)$$

We use a digraph, where $\omega(s, t)$ is a combination of a regular undirected dissimilarity measure $\psi(s, t)$, multiplied by an orientation factor (i.e., $\omega(s, t) = \psi(s, t) \times (1 + \alpha)$ if $I(s) > I(t)$ and $\omega(s, t) = \psi(s, t) \times (1 - \alpha)$ otherwise). Several different procedures can be adopted for $\psi(s, t)$, such as the absolute value of the difference of image intensities (i.e., $\psi(s, t) = |I(t) - I(s)|$). Note that we have $\omega(s, t) \neq \omega(t, s)$ when $\alpha > 0$.

The OIFT is built upon the IFT framework by considering one of the following path functions in a symmetric digraph:

$$f_{\text{max}}^{E_{\text{direct}}}(\pi_{r \rightarrow s}, \langle s, t \rangle) = \begin{cases} -1 & \text{if } t \in S_1 \cup S_2 \\ +\infty & \text{otherwise} \end{cases}$$

$$f_{\text{max}}^{E_{\text{direct}}}(\pi_{r \rightarrow s}, \langle s, t \rangle) = \begin{cases} \text{Expr1} & \text{if } r \in S_1 \\ \text{Expr2} & \text{if } r \in S_2 \end{cases}$$

$$\text{Expr1} = \max\{f_{\text{max}}^{E_{\text{direct}}} (\pi_{r \rightarrow s}, \langle s, t \rangle), 2 \times \omega(t, s) + 1\}$$

$$\text{Expr2} = \max\{f_{\text{max}}^{E_{\text{direct}}} (\pi_{r \rightarrow s}, \langle s, t \rangle), 2 \times \omega(s, t)\}$$

The segmentation using $f_{\text{max}}^{E_{\text{direct}}} \text{ or } f_{\text{max}}^{E_{\text{direct}}} \text{ favors transitions}$ from dark to bright pixels, and $f_{\text{max}}^{E_{\text{direct}}} \text{ or } f_{\text{max}}^{E_{\text{direct}}} \text{ favors}$ the opposite orientation, according to Theorem 1. In the case of multiple candidate segmentations with the same energy, $f_{\text{max}}^{E_{\text{direct}}} \text{ produces a better handling of the tie zones}$ than $f_{\text{max}}^{E_{\text{direct}}} \text{.}$
Theorem 1 (Inner/outer-cut boundary optimality). For two given sets of seeds \( S_o \) and \( S_b \), any spanning forest computed by the IFT algorithm for function \( f_{\max}^{S, S_o} \) or \( f_{\max}^{S, S_b} \) defines an optimum cut that maximizes \( E_i \) among all possible segmentation results satisfying the hard constraints. Any spanning forest computed by the IFT algorithm for function \( f_{\max}^{S, S_o} \) or \( f_{\max}^{S, S_b} \) defines an optimum cut that maximizes \( E_o \) among all possible segmentation results satisfying the hard constraints (seeds).

In our experiments, we used 20 real volumetric MR images of the foot in 3D. We computed the mean performance curve (Dice coefficient) for the methods: Iterative Relative Fuzzy Connectedness (IRFC), IFT with \( f_{\max} \) \( (IFT^{\max}_{F}) \), Power Watershed \( (PW_{q=2}) \), and OIFT using \( f_{\max}^{S, S_b} \) \( (OIFT^{\max}_{inner}) \), \( f_{\max}^{S, S_o} \) \( (OIFT^{\max}_{outer}) \), \( f_{\omega}^{S, S_b} \) \( (OIFT_{\omega}^{inner}) \) and \( f_{\omega}^{S, S_o} \) \( (OIFT_{\omega}^{outer}) \) \( [17] \). We used different seed sets obtained by eroding and dilating the ground truth (Figure 4). The experimental accuracy curves with the Sobel gradient (Figure 5) show that whenever the object presents transitions from dark to bright pixels, as it is the case with the bones talus and calcaneus, \( f_{\max}^{S, S_b} \) and \( f_{\omega}^{S, S_b} \) give the best accuracy results. Note also that \( f_{\max}^{S, S_o} \) and \( f_{\omega}^{S, S_o} \) present the worst accuracy values, by specifying the wrong orientation.

Fig. 4. (a) Ground truth of the talus in an MR image of a foot. (b) Seed sets obtained by eroding and dilating the ground truth. (c) Segmentation by IRFC. (d) An improved result by exploiting the boundary polarity using \( f_{\max}^{S, S_b} \).

VI. SHAPE CONSTRAINTS VIA NSCF

Shape constraints, such as the star-convexity prior introduced by Veksler \( [25] \), can limit the search space of possible delineations to a smaller subset, thus eliminating false candidate boundaries. In this context, a point \( p \) is said to be visible to \( c \) via a set \( \mathcal{O} \) if the line segment joining \( p \) to \( c \) lies in the set \( \mathcal{O} \). An object \( \mathcal{O} \) is star-convex with respect to center \( c \), if every point \( p \in \mathcal{O} \) is visible to \( c \) via \( \mathcal{O} \) (Figure 6). In the case of multiple stars, a computationally tractable definition, was proposed in \( [24] \), using a Geodesic Star Convexity (GSC) constraint in the segmentation by \( \min -cut/\max -flow \).

In \( [18] \), we proposed an IFT extension that incorporates the GSC constraint, favoring the segmentation of objects with more regular shape, resulting in a novel method called \( OIFT \) with Geodesic Star Convexity Constraints \( (GSC-IFT) \). In this method, the set of star centers is taken as the set of internal seeds \( (S_o) \), and the line segments are the paths that form a spanning forest rooted at the internal seeds. The forest topology is controlled by a parameter \( \beta \). For lower values of \( \beta \) \( (\beta \approx 0.0) \), the method imposes more star regularization to the boundary of the object (Figures 7 and 8), and for higher values of \( \beta \), it allows a better fit to the curved protrusions and indentations of the boundary.

Thereafter, in \( [19] \) we proposed the novel method called \( OIFT \) with Geodesic Star Convexity \( (GSC-OIFT) \), which incorporate Gulshan’s geodesic star convexity prior in the OIFT approach for interactive image segmentation, in order to simultaneously handle boundary polarity and shape constraints (Theorem 2). This method permits the customization of the segmentation by IFT to better match the features of a particular target object (Figure 9). We constrain the search for optimum result, that maximize the graph-cut measures \( E_i \) (Eq. 2) or \( E_o \) (Eq. 3), only to segmentations that satisfy the geodesic star convexity constraint. We compute a geodesic forest \( P_{\sum} \) for \( f_{\sum} \) \( [10] \) by the regular IFT algorithm, using only \( S_o \) as seeds, for the given digraph \( G \), obtaining two sets of arcs \( \xi_{P_{\sum}}^{t} = \{(s, t) \in \xi \mid s = P_{\sum}(t)\} \) and \( \xi_{P_{\sum}}^{o} = \{(s, t) \in \xi \mid t = P_{\sum}(s)\} \). The GSC constraint is violated when \( C_i(L) \cap \xi_{P_{\sum}}^{t} \neq \emptyset \) or \( C_o(L) \cap \xi_{P_{\sum}}^{o} \neq \emptyset \).
for higher values of \( \alpha \) obtained by eroding and dilating the ground truth. Note that CT studies to segment the liver. Figure 11a shows the mean constraints by the geodesic star convexity.

Among all possible segmentation results satisfying the shape constraints, so that the accuracy tends to decrease (Figures 11b-d). Figure 12 shows some results in the case of user-selected constraints. Similarly, the segmentation computed over \( G' \) by the IFT algorithm for function \( f_{\text{max}} \) defines an optimum cut in the original graph \( G \), that maximizes \( E_o(L,G) \) among all possible segmentation results satisfying the shape constraints by the geodesic star convexity, and the seed constraints. Similarly, the segmentation computed by the IFT algorithm for function \( f_{\text{max}}^{S_o,S_b} \), over a modified graph \( G' = (\mathcal{V}, \xi, \omega') \); with weights \( \omega'(s,t) = -\infty \) for all \((s,t) \in \xi_{\text{sum}}^P\) and \( \omega'(s,t) = \omega(s,t) \) otherwise; defines an optimum cut in the original graph \( G \), that maximizes \( E_i(L,G) \) among all possible segmentation results satisfying the shape constraints by the geodesic star convexity.

In our experiments, we used 40 image slices of 10 thoracic CT studies to segment the liver. Figure 11b shows the mean accuracy curves for the same images assuming different seed sets obtained by eroding and dilating the ground truth. Note that for higher values of \( \alpha \), GSC–OIFT imposes less shape constraints, so that the accuracy tends to decrease (Figures 11b-d). Figure 12 shows some results in the case of user-selected markers.

VII. CONCLUSION AND FUTURE WORKS

The proposed extension GSC–OIFT includes the IFT with \( f_{\text{max}} \), OIFT and GSC–IFT as particular cases, depending on the configuration of its parameters \( \alpha \) and \( \beta \). Note that the adaptive functions presented in Section IV can’t be reduced to a GSC–OIFT computation. As a result of the theoretical foundation proposed in this work, four conference papers were published in international events of high regard [17], [18], [19], [20], and one journal paper was published in the IEEE Transactions on Image Processing (impact factor: 3.111) [16]. This work has also allowed new achievements that were recently published, such as [26] and [27].

As future work, we intend to combine the proposed methods with statistical models to automatically define seeds for automatic segmentation.

ACKNOWLEDGMENT

The authors thank FAPESP grant #2011/50761-2, #2012/06911-2, CNPq (305381/2012-1, 486083/2013-6), CAPES, NAP eScience - PRP - USP, FINEP project (1266/13) in biomedical engineering and Dr. J.K. Uduapa (MIPG-UPENN) for the images.
References


