An Efficient Hierarchical Layered Graph Approach for Multi-Region Segmentation

Leissi M. Castañeda Leon*
Institute of Mathematics and Statistics
São Paulo, Brazil 05508–090
Email: leissicl@ime.usp.br

Krzysztof Chris Ciesielski
Department of Mathematics
West Virginia University, USA
Email: kcies@math.wvu.edu

Paulo A. Vechiatto de Miranda
Institute of Mathematics and Statistics
São Paulo, Brazil 05508–090
Email: pmiranda@vision.ime.usp.br

***Abstract*—We proposed a novel efficient seed-based method for the multiple region segmentation of images based on graphs, named *Hierarchical Layered Oriented Image Foresting Transform* (HLOIFT). It uses a tree of the relations between the image objects, represented by a node. Each tree node may contain different individual high-level priors and defines a weighted digraph, named as *layer*. The layer graphs are then integrated into a hierarchical graph, considering the hierarchical relations of inclusion and exclusion. A *single* energy optimization is performed in the hierarchical layered weighted digraph leading to globally optimal results satisfying all the high-level priors. The experimental evaluations of HLOIFT and its extensions, on medical, natural and synthetic images, indicate promising results comparable to the state-of-the-art methods, but with lower computational complexity. Compared to hierarchical segmentation by the min-cut/max-flow algorithm, our approach is less restrictive, leading to globally optimal results in more general scenarios, and has a better running time.**

I. Introduction

The image segmentation task is not only one of the most fundamental and challenging problems in image processing and computer vision, but also has impact in different research areas such as Medicine, Neurology and Artificial Intelligence [1]–[3]. For example, in medical imaging, image segmentation can be used to isolate the regions corresponding to different organs in Magnetic Resonance Images (MRI), helping to analyze their forms, volumes and textures for the diagnostic of pathologies [4]. In this case, we usually have critical obstacles to face such as noise, low contrast and regions complexity. The incorporation of prior knowledge is useful for more accurate results.

Currently, graph-based methods are commonly used, where the image is modeled as a connected graph, because graphs can naturally represent image parts and their relationships [5]. Then, the image segmentation task can be interpreted as a graph partition problem subject to hard constraints, such as seed pixels, given by an algorithm or a user, in the image domain for the foreground regions and background (anywhere outside the foreground). Examples of graph-based methods are watershed from markers [6], random walks [7], fuzzy connectedness [8], graph cuts (GC) [9], grow cut [10], minimum barrier distance [11] and image foresting transform (IFT) [12], [13].

* This work is related to her Ph.D. thesis

For the multiple region segmentation, each region may present its own distinctive features, requiring different priors to guide the segmentation process, e.g. shape constraints [14]–[16], convexity prior [17], and boundary polarity [18], [19], allowing the customization of the segmentation to a given target region. Also, it is advantageous to explore the structural interaction between the different regions in the image, whenever it is possible. However, many existing classical approaches do not include any form of structural information together into a single energy optimization.

Most of the methods for multi-region segmentation including structural information, such as inclusion or exclusion interactions between regions, are based on graph-cut optimization and are performed by a min-cut/max-flow algorithm [20]–[22]. However, their globally optimal results are restricted only to some particular cases. For example, they cannot represent the inclusion of a pair of adjacent regions in a third region because it cannot be converted to a submodular energy. Also they usually have a high computational cost. The methods based in LOGISMOS [23], [24] require an approximated pre-segmentation whenever the regions present complex shapes. Then, the fast segmentation obtained by our proposed method could also be used as a starting point for LOGISMOS.

In the context of segmentation by the *Image Foresting Transform* (IFT) framework, in order to incorporate structural information among regions, the methods usually employ Fuzzy Object Models (FOMs) [25]–[31]. However, these approaches are based on separate IFT executions per region, that do not incorporate structural information and high-level priors into a single energy optimization, limiting their potential.

A. Proposal

In this work, we circumvent the aforementioned problems, by proposing a hierarchical layered graph-based approach for the multiple region segmentation problem, named as *Hierarchical Layered OIFT* (HLOIFT). We formulate the integration of individual region constraints such as boundary polarity and shape constraint together with geometric priors such as inclusion and exclusion constraints between regions, within a single energy which is optimized by a new algorithm with proof of correctness, overcoming the mentioned limitations from previous works and conserving a low computational cost of *Oriented Image Foresting Transform* (OIFT) [19].

B. Contributions

Therefore, our general main contributions are as follows.

- **Theoretical:** We propose a new method for multi-region segmentation which may include high-level priors for image regions and the hierarchical constraints between them.
- **Generality:** Our approach is less restrictive than most methods in use and leads to globally optimal results in more general scenarios.
- **Complexity:** Our method has lower computational complexity as compared to methods based on the min-cut/max-flow algorithm.

These contributions are resumed in one conference paper published in a international event [32], one journal paper submitted to the Journal of Visual Communication and Image Representation and which is still unpublished [33], one book chapter [34] as part of the workshops proceedings of the European Conference in Computer Vision (ECCV) and three poster presentation [35]–[37] in international conferences.

C. Outline

For the sake of completeness in presentation, Section II includes some required definitions and the relevant previous work of image segmentation by OIFT with some used priors. Section III presents our new algorithm HLOIFT and its extension to work with 3D volumetric images and considering superpixels instead of pixels. In Section IV are some experimental evaluations of HLOIFT and its extensions, comparing it against IFT and hierarchical segmentation by the min-cut/max-flow algorithm, showing the HLOIFT potential. Finally, our conclusions are stated in Section V.

II. Background

An *(2-dimensional) image* can be interpreted as a weighted digraph (i.e., directed graph) $G = (\mathcal{N}, \mathcal{A}, \omega)$ whose nodes (vertices) $\mathcal{N}$ are the pixels in the image domain $\mathcal{I} \subset Z^2$, the arcs/edges listed in $\mathcal{A}$ are the ordered pairs of pixels $(s,t) \in \mathcal{I}^2$ (usually, in 2D images, identified with either 4- or 8-neighborhoods), and the weight map ω associates to each arc $(s,t) \in \mathcal{A}$ a value $\omega(s,t) \in [-\infty, \infty]$ (usually defined as $\omega(s,t) = \|I(t) - I(s)\|$, i.e. a dissimilarity measure). We use the notations $t \in \mathcal{A}(s)$ and $(s,t) \in \mathcal{A}$ to indicate that t is adjacent to s.

A digraph G is *symmetric* if for all $(s,t) \in \mathcal{A}$, the pair (t,s) is also an arc of G, that is, $(t,s) \in \mathcal{A}$. All digraphs we consider below are symmetric. Notice, that we may have $\omega(s,t) \neq \omega(t,s)$.

In binary segmentation, whose goal is to separate an object/region O_1 from its background $O_0 = \mathcal{I} \setminus O_1$, we consider two non-empty seed sets $\mathcal{S}_0 \cup \mathcal{S}_1 \subset \mathcal{I}$, aiming for $O_1 \supset \mathcal{S}_1$ and $O_0 \supset \mathcal{S}_0$, assuming that $\mathcal{S}_1 \cap \mathcal{S}_0 = \emptyset$. In particular, the region O_1 is identified with its *labeling* $L\colon \mathcal{I} \to \{0,1\}$, which is the indicator function of O_1, i.e. $O_1 = \{v \in \mathcal{I}\colon L(v) = 1\}$ and $O_0 = \{v \in \mathcal{I}\colon L(v) = 0\}$. All this notation may be easily extended for multiple region segmentation.

A. Oriented Image Foresting Transform (OIFT)

The Oriented Image Foresting Transform (OIFT) method is build upon the Image Foresting Transform(IFT) framework [12] which is a modification of Dijkstras shortest paths algorithm [38] allowing multiple sources and general connectivity functions (Monotonically Increasing (MI) functions). The OIFT explores the region-contour orientation, in connected and symmetric digraphs using non-smooth connectivity functions (NSCF) [19], and some optimal criteria based on a cut measure on the graph. It improves the segmentation results because it helps to distinguish between two similar and nearby boundary segments with opposite orientations from distinct regions. However it was only defined for binary segmentation.

The resulting image segmentation by OIFT is a global optimum solution by **maximizing** the graph-cut measure given by:

$$\varepsilon_{min}(L) = \min\{\omega(s,t)\colon (s,t) \in \mathcal{A} \;\&\; L(s) > L(t) = 0\}, \tag{1}$$

subject to the seed constraints.

B. High-level priors

We briefly describe some individual and structural priors used, such as boundary polarity, shape priors and geometric interactions.

1) Boundary polarity: To explore the boundary polarity the arcs weights $\omega(s,t)$ are defined as a combination of an undirected dissimilarity measure $\psi(s,t)$ between neighboring pixel s and t, multiplied by an orientation factor, as follows:

$$\omega(s,t) \;=\; \begin{cases} \psi(s,t) \times (1+\alpha) & \text{if } I(s) > I(t), \\ \psi(s,t) \times (1-\alpha) & \text{if } I(s) < I(t), \\ \psi(s,t) & \text{otherwise,} \end{cases} \tag{2}$$

where $\alpha \in [-1,1]$ and we usually have $\psi(s,t) = |I(t) - I(s)|$. Other options for $\psi(s,t)$ are discussed in [39], [40]. Note that, in general, we have $\omega(s,t) \neq \omega(t,s)$ for $\alpha \neq 0$. For $\alpha > 0$, the segmentation by OIFT favors transitions from bright to dark pixels, and for $\alpha < 0$ favors the opposite orientation. Finally, for multi-region segmentation we consider multiples α_i values, each associated to a different region.

2) Geodesic star convexity: The *geodesic star convexity prior* (GSC) corresponds to a discrete version of the *star shape prior* (SSP) [16], by considering shortest paths in the image graph, returned by IFT with the additive path-cost functions, as line segments. It considers all the given seeds as center points of a region, prioritizing the segmentation of the region with more regular shape [41]. The geodesic star convexity prior is obtained by setting the weights of some arcs in the created digraph to $-\infty$, according to the scheme proposed in [41]. Moreover, it is still possible to simultaneously handle boundary polarity and shape priors [41].

3) Geometric interactions: The contribution of [20], for multi-region segmentation, is a binary multi-layered formulation, defining a layer for each region, encoding only geometric interactions between the different regions in the image. These interactions refer to the *inclusion* of a region within another

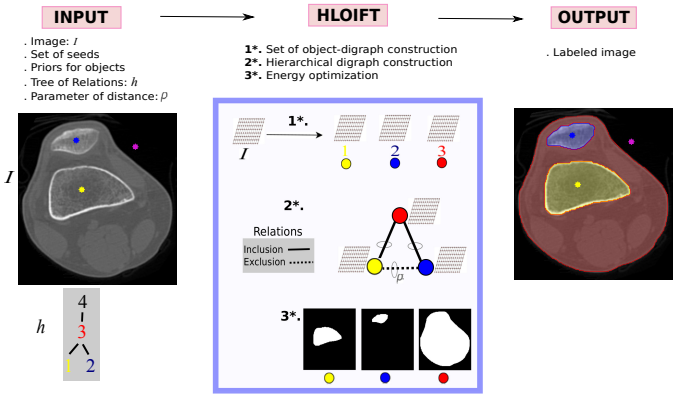


Fig. 1. Overview of our framework. Given the input parameters, a hierarchical weighted digraph of layers (digraphs) is constructed using the *inclusion* (solid line) and *exclusion* (dashed line) hierarchical constraints between objects, and a graph-cut measure is optimized by our algorithm. Finally, we have a labeled image as output.

and the *exclusion* of them, while enforcing a minimal distance between the region contours in different layers for the inclusion case. However, in [20], their globally optimal results are restricted only to some particular cases, because it often results in bad local minima due to complexities of interaction constraints.

III. Our Method

We propose the ***Hierarchical Layered Oriented Image Foresting Transform*** (HLOIFT) as a new seed-based method for multi-object segmentation. Figure 1 shows an overview of our framework. For a given input image, seeds sets for some objects, and the tree of relations between objects, the HLOIFT method has the following steps: (1) Each layer is constructed as a weighted digraph representing one object with its own priors (described in Subsection III-A). (2) HLOIFT defines a setup for the inter-layer connections representing the hierarchical constraints, such as inclusion and exclusion relations (described in Subsection III-B). (3) HLOIFT uses an extension of the OIFT algorithm to compute an optimal cut over the hierarchical layered digraph, giving as output a labeled image (described in Subsection III-C).

A. Set of layer digraph construction

We first create a set of m *layers*, where each layer $\mathcal{H}_i$ represents a single region O_i, $i \in \mathcal{L} = \{1, \ldots, m\}$, of an (*n-dimensional*) image. A *layer* $\mathcal{H}_i = (\mathcal{N}_i, \mathcal{A}_i, \omega_i)$ is a weighted digraph, where $\mathcal{N}_i = \{i\} \times \mathcal{I}$ and each node $t = (i, v) \in \mathcal{N}_i$ correspond to the pixel $p(t) = v$ and $\lambda(t) = i$ means that t belongs to the ith layer of the graph. We define the intralayer adjacency $\mathcal{A}_i$ on $\mathcal{N}_i = \mathcal{I}$ as $\mathcal{A}_\mathcal{I}$, usually being the 4- or 8-neighborhood adjacency in the case of 2D images. Similarly, an intralayer weight function ω_i for every $(s, t) \in \mathcal{A}_i$ (given, for example, by (2)). Of course, ω_i should highlight the priors for O_i whenever it is appropriate. We used the *boundary polarity* priors, and the *geodesic star convexity prior*, prioritizing the O_i with more regular shape, as mentioned before.

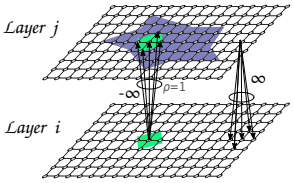


Fig. 2. Illustration of inclusion inter-layer arc construction, involving two regions O_i and O_j (O_j is the parent of O_i, i.e., $h(i) = j$), where $\omega(s,t) = -\infty$ and $\omega(t,s) = \infty$ for $\lambda(s) = i$ and $\lambda(t) = j$.

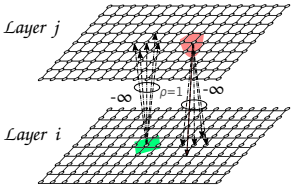


Fig. 3. Illustration of exclusion inter-layer arc construction, involving two regions O_i and O_j (O_i and O_j are siblings), where $\omega(s,t) = -\infty$ for all inter-layer arcs (s,t).

B. Hierarchical digraph construction

In this step, HLOIFT generates a *hierarchical layered* weighted digraph $\mathcal{H}$ as the union of all layered graphs $\mathcal{H}_i$, $i = 1, \ldots, m$, with additional *interlayer* arcs connecting only some of the distinct layers, based on the requirements imposed by the priors h (a *tree*) and ρ (distance parameter).

The hierarchy prior (h) between any pair $\langle O_i, O_j \rangle$ of regions is understood as follows. If $O_{m+1} = \mathcal{I}$ (the image domain and the root of the tree). Then $h(i) = j$ (**inclusion**) if, and only if, O_j is the smallest of the regions properly containing O_i, and we will refer to O_j as the *parent* of O_i. We say that the regions O_i and O_j, with $O_i \cap O_j = \emptyset$, are *siblings* (**exclusion**), provided $i, j \in \mathcal{L}$ are distinct and $h(i) = h(j)$.

Specifically, we declare that for $s \in \mathcal{N}_i$ and $t \in \mathcal{N}_j$, with $i, j \in \mathcal{L}$ being distinct, the pair (s, t) is an arc in $\mathcal{H}$ (i.e., (s, t) belongs to $\mathcal{A}$ of $\mathcal{H}$) if, and only if, $p(s) - p(t) \leq \rho$ and either O_i and O_j are siblings, or one of them is the parent of the other. We will use the distance parameter $\rho \geq 0$, where

(C) for siblings O_i and O_j we will assume that $\|s - t\| > \rho$ for every $s \in O_i$ and $t \in O_j$, while for parent-offspring pair $\langle O_j, O_i \rangle$ that $t \in O_j$ whenever there exists an $s \in O_i$ with $\|s - t\| \leq \rho$.

The weights of the inter-layer arcs for *inclusion*, is $\omega(t, s) = \infty$ and $\omega(s, t) = -\infty$, and for *exclusion* ($\mathcal{A}_s$) as $\omega(s, t) = \omega(t, s) = -\infty$. Figures 2 and 3 show the arcs and their corresponding weights, with $\rho = 1$ of inter-layer arcs using a 4-neighborhood adjacency for inclusion and exclusion cases respectively.

C. Energy optimization

Finally, we execute the HLOIFT defined by Algorithm 1 and a labeled image is obtained. The algorithm is applied to

the hierarchical layered graph $\mathcal{H}$ constructed above and its output *maximizes* a single energy ε^h_{min} defined to ensure that the output satisfies also the hierarchical constraints imposed by h and ρ. Specifically, for a binary map $X\colon \mathcal{N} \to \{0,1\}$ (segmentation) the energy ε^h_{min} of X is defined as

$$\varepsilon^h_{min}(X) \quad = \quad \min\{\varepsilon^{incl}_{min}(X), \varepsilon^{excl}_{min}(X)\}, \tag{3}$$

where $\varepsilon^{incl}_{min}(X) = \min\{\omega(s,t)\colon (s,t) \in \mathcal{A} \setminus \mathcal{A}_s \ \&\ X(s) > X(t)\}$, and $\varepsilon^{excl}_{min}(X) = \min\{\omega(s,t)\colon (s,t) \in \mathcal{A}_s \ \&\ X(s) = X(t) = 1\}$.

As result we have the Theorem 1 which proof of correctness is described in [33].

Theorem 1 (Cut optimality by HLOIFT)**.** *For every image $(\mathcal{I}, I)$, a hierarchy tree h, a distance parameter $\rho \geq 0$, and a sequence $\langle \mathcal{S}_0, \ldots, \mathcal{S}_m \rangle$ of seed sets consistent with (C) (with respect to h and ρ), the binary map $X\colon \mathcal{N} \to \{0,1\}$ computed by Algorithm 1 maximizes the energy $\varepsilon^h_{min}(X)$ given by (3) among all solutions satisfying the seed constraints and the consistency requirement (C).*

The Algorithm 1 starts initializing the binary and cost map together with the priority queue, given the input set of seeds. In lines 7-10 indicates that for the inclusion case, if the seeds for the parent region are given, then they are propagated as background seeds to the layers of their corresponding children regions. The condition on line 16 indicates, if we have only one region (i.e., $\mathcal{L} = \{1\}$), then $X = L$, or when tree h has no siblings, the condition from line 19 is never satisfied, so lines 19-20 can be removed. The weights of inter-layer arcs are chosen among $-\infty$ and ∞ so that when its value becomes $V(t)$, during the execution of line 17 in HLOIFT, the value of $V(t)$ (equal to $\omega(s,t)$ when $X(s) = 1$ and to $\omega(t,s)$ when $X(s) = 0$) is $-\infty$ if, and only if, the value of $X(s)$ together with the consistency requirement (C) uniquely determine the value of $X(t)$. In particular, $V(t) = -\infty$ will ensure that t will be chosen from the queue Q with the highest possible priority, so that the value of $X(t)$ can be affixed to that forced by (C) and $X(s)$ before the consistency of X could be jeopardized. Specifically, for inclusion case if s is removed from Q in line 11 with $X(s) = 1$, then the required final value $X(t) = 1$ will be ensured soon after by the highest priority of $V(t) = \omega(s,t) = -\infty$. Similarly, if t is removed from Q in line 8 with $X(t) = 0$, then the eventual value $X(s) = 0$ required by (C) will be ensured by the highest priority of $V(s) = \omega(s,t) = -\infty$. Thus for the exclusion case, if s is removed from Q in line 11 with $X(s) = 1$, then the eventual value $X(t) = 0$ required by (C) will be ensured by the highest priority of $V(t) = \omega(s,t) = -\infty$. Similarly, when t is removed from Q in line 11 with $X(t) = 1$. In case when, for the removed s, we have $X(s) = 0$, HLOIFT will simply not update any information on t. In this case, in line 16 the condition is satisfied when s and t are siblings and $p(s) \in O_{\lambda(s)}$ as $X(s) = 1$. Therefore, lines 19-20 ensure that for the output X we have no $(s,t) \in \mathcal{A}_s$ with $X(s) = X(t) = 1$, that is, $\varepsilon^{excl}_{min}(X) = -\infty$.

Algorithm 1. – HLOIFT ALGORITHM

INPUT: Hierarchical layered digraph $\mathcal{H} = (\mathcal{N}, \mathcal{A}, \omega)$, build using the hierarchical tree h and the distance parameter ρ; the seed sets $\langle \mathcal{S}_0, \ldots, \mathcal{S}_m \rangle$.
OUTPUT: The binary map $X\colon \mathcal{N} \to \{0,1\}$ identifying segmentation of regions.
AUXILIARY: Priority queue Q, variable tmp, the cost map $V\colon \mathcal{N} \to [-\infty, \infty]$, and an array of status $S\colon \mathcal{N} \to \{0,1\}$, where $S(t) = 1$ for processed nodes and $S(t) = 0$ for unprocessed nodes.

```
1.  For each t ∈ N and i ∈ L do
2.      Set S(t) ← 0 and V(t) ← ∞;
3.      If p(t) ∈ S_0 then
4.          V(t) ← −∞, X(t) ← 0, insert t in Q;
5.      If p(t) ∈ S_i and λ(t) = i then
6.          V(t) ← −∞, X(t) ← 1, insert t in Q.
7.          For each j ∈ L do
8.              If h(j) = i and λ(t) = j then
9.                  V(t) ← −∞, X(t) ← 0, insert t in Q.
10. While Q ≠ ∅ do
11.     Remove s from Q such that V(s) is minimum;
12.     Set S(s) ← 1;
13.     For each (s,t) ∈ A such that S(t) = 0 do
14.         If X(s) = 1 then tmp ← ω(s,t)
15.         Else tmp ← ω(t,s);
16.         If tmp < V(t) and ¬[(s,t) ∈ A_s and X(s) = 0], then
17.             Set V(t) ← tmp and X(t) ← X(s);
18.             If t ∉ Q then insert t in Q;
19.             If (s,t) ∈ A_s and X(s) = 1 then
20.                 X(t) ← 0.
```

Concerning the computational complexity, HLOIFT is $\mathcal{O}(M + N)$, where N is the number of vertices in the graph $\mathcal{H}$ and M is the number of arcs in the graph $\mathcal{H}$, when $\mathcal{Q}$ is implemented using bucket sorting [12] and $\mathcal{O}(M + N \log N)$ (linearithmic time) if $\mathcal{Q}$ is a heap. The Graph cut computational complexity is $\mathcal{O}(\sqrt{M} * N^2) = O(N^{2.5})$ when $\mathcal{H}$ is a sparse graph, which is more than quadratic-time using a push-relabel based on the highest label node selection rule [42].

D. Extensions of HLOIFT

1) 3D HLOIFT: We propose the usage of HLOIFT for the segmentation of 3-dimensional images (volumes), where $\mathcal{I}$ is a finite set of *voxels* (i.e., $\mathcal{I} \subset Z^3$). From the implementation point of view of HLOIFT, the only required changes are that now each layer has intra-layer arcs in a 3D neighborhood and that the parameter ρ defines a sphere in space.

2) HLOIFT with superpixels: We also propose a **superpixel**-based adaptation of the HLOIFT method [34], leading to a more efficient and adequate solution in large images, reducing the running time conserving the results by HLOIFT. Superpixels can group pixels into perceptually meaningful atomic regions of similar and connected pixels being computationally more efficient than their pixel counterparts. In this case, we used the IFT-SLIC [43], to compute superpixels from a given input image, creating a *Region Adjacency Graph* (RAG) greatly reducing the number of graph elements. Then, we create each pair $(\mathcal{N}_i,\ \mathcal{A}_i)$ as an isomorphic copy of the created RAG, then the framework is similar to the HLOIFT considering ever superpixels instead of pixels.

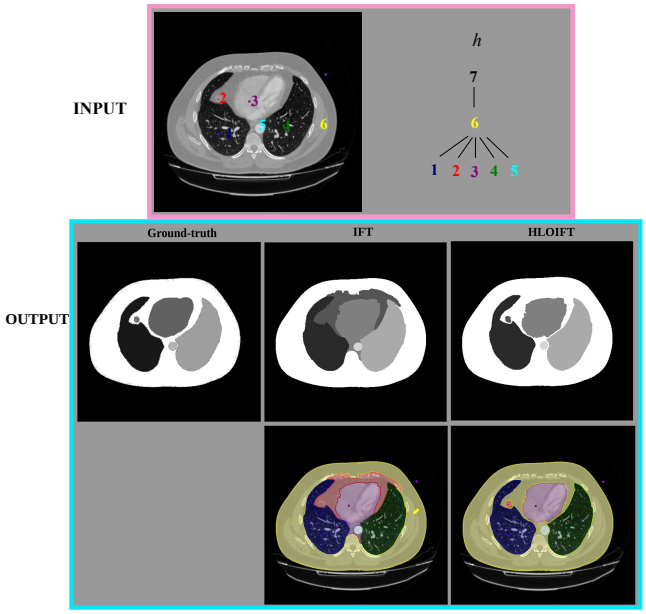


Fig. 4. An axial cross section of a thoracic-abdominal CT image segmentation. HLOIFT obtained a result similar to the given manual ground-truth.

IV. Experimental results

This section is divided in three parts, showing experimental results for the 2D HLOIFT, 3D HLOIFT and for the HLOIFT with superpixels respectively. We used medical and synthetic images, composed of multiple sophisticated regions. Also, the experiments presented here were conducted in a laptop Intel Core i3-5005U CPU 2.00GHz $\times 4$.

A. 2D HLOIFT

1) Qualitative comparison with the IFT method: In Figure 4, we use an axial cross section of a thoracic-abdominal CT image extracted from [44] to segment six regions: right lung (O_1), liver (O_2), heart (O_3), left lung (O_4), aorta (O_5) and the thoracic-abdominal region (O_6). We used ω_i with $\psi(s,t) = I(s) - I(t)$, $\rho = 3.5$, boundary polarity from dark to bright pixels for O_1, O_4 shape constraint by geodesic star convexity for O_2 and O_3, and boundary polarity from bright to dark pixels for O_5 and O_6. Clearly, the results obtained by HLOIFT are closer to the ground-truth compared to the IFT results.

2) Comparison with multi-object segmentation by min-cut/max-flow algorithm: We show a comparison between HLOIFT and the hierarchical segmentation of multiple regions min-cut/max-flow algorithm [20]. Figure 5 shows a flower segmentation using $\omega_i(s,t) = G(s) + G(t)$, where $G(.)$ denotes the maximum magnitude of the Sobel gradient for the three image channels. We used $\rho = 1.5$ and no region constraints (shape constraints/boundary polarity). In general the results are similar for both methods. The min-cut/max-flow algorithm generates results with a smoother contour, but sometimes this can lead to errors in some finer parts of the region, such as the petals (Figure 5b). Besides being able to compute globally optimal results with arbitrary hierarchy constraints, the proposed HLOIFT method also has a better running time compared to the min-cut/max-flow algorithm. Table I shows the running times for the flower segmentation using different image resolutions.

(a) (b) (c)

Fig. 5. Flower segmentation in two regions, the central part in cyan and the petals in yellow, using the inclusion relation. (a) The input image. (b) Result by the min-cut/max-flow algorithm in layered graphs. (c) Result by HLOIFT.

Image size (pixels)	HLOIFT (ms)	min-cut/max-flow (ms)
380×320	114.65	323.61
760×640	488.62	1,798.91
1520×1280	1,823.55	19,021.71

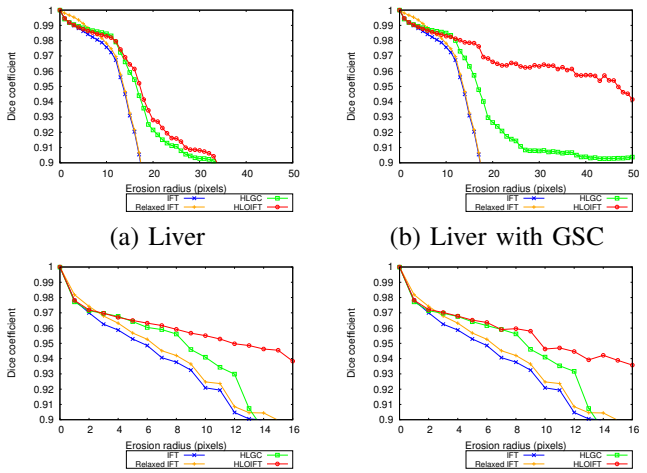
TABLE I
THE RUNNING TIMES FOR THE FLOWER SEGMENTATION BY HLOIFT AND THE MIN-CUT/MAX-FLOW ALGORITHM IN LAYERED GRAPHS USING DIFFERENT IMAGE SIZES.

3) Quantitative accuracy experiments: We compared HLOIFT, with the multi-region segmentation by IFT [45], its improved version by the relaxation procedure proposed by [46], and the hierarchical layered graph cut [20], denoted as HLGC, using two datasets of medical images into 2 datasets, and considering simple arc weights ($\psi(s,t) = |I(s) - I(t)|$).

The first dataset has 40 slices from thoracic CT studies of size 512×512 to segment the liver and the abdomen as its parent region. The second dataset has 40 real MR images to segment the talus and calcaneus bones as siblings, taking the foot region as their parent. The same seeds were used for all the methods, obtained by progressively eroding the ground truth regions and background for different radius sizes. For the bones we used $\alpha = -0.5$ and $\rho = 3$ pixels, while for the liver we used $\alpha = 0.9$ and $\rho = 5$ pixels. The mean accuracy curves according to the Dice coefficient are shown in Figure 6, being the results in the left column obtained without shape constraints and the ones in the right with Geodesic Star Convexity (GSC). HLOIFT had the best results in most cases. Note also that it could benefit more from the usage of the shape constraint by GSC for the liver compared to HLGC.

B. 3D HLOIFT

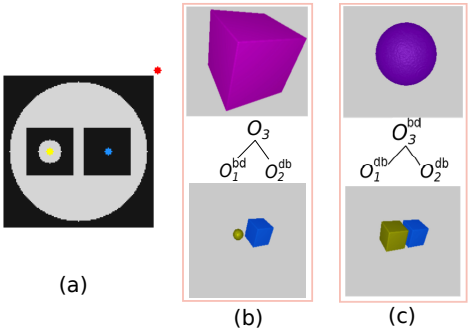
We presents an experimental evaluation of HLOIFT on 3-dimensional images (volumes) to assess its performance considering different high-level priors. In Figure 7, we use a synthetic image of size $150 \times 150 \times 150$ voxels composed of six distinct regions: Two dark cubes contained in a brighter ball, which in turn is contained inside a bigger cube, surrounded by a white background, and a small ball contained in one of the inner cubes. Then, we want to segment it into three regions, where O_1 and O_2 are *included* in O_3. We will highlight the



(a) Liver (b) Liver with GSC

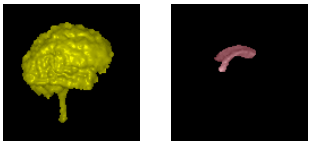
(c) Talus (d) Talus with GSC

Fig. 6. The mean curves of Dice accuracy for different methods.



(a) (b) (c)

Fig. 7. (a) Given the input image with six regions and seeds for O_1 (in yellow), O_2 (in blue) and background (in red), we obtain different results for O_1 and O_2 included in O_3, defining different boundary polarities. In (b) and (c) the 3D renderings of the regions is given.



(a) (b)

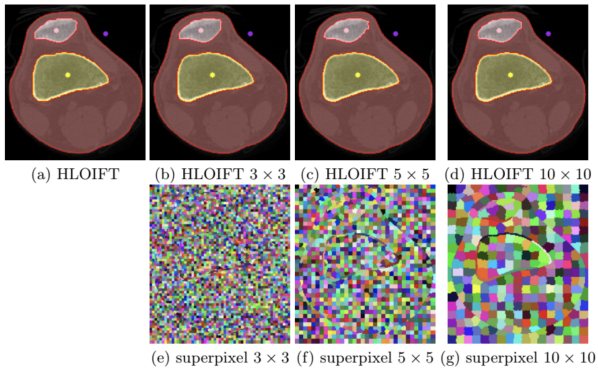
Fig. 8. 3D HLOIFT inclusion segmentation for (a) the brain and (b) ventricles.

boundary polarity prior, showing that using of prior "from bright to dark" (bd) or "from dark to bright" (db) for O_1 and O_3, gives different results for the same input image, which are shown in a slice view and the 3D renderings of the regions' surfaces as depicted in a red box.

In order to test our 3D HLOIFT in a real MR image, we used an image of $120 \times 120 \times 90$ voxels, to segment the brain (Figure 8a) and the lateral ventricles (Figure 8b) using the inclusion relation. The running time was 4.2 sec.

C. HLOIFT with superpixels

We present experimental results leading to a more efficient and adequate solution for multi-region segmentation in large images, showing the improvement in the running time.



(a) HLOIFT (b) HLOIFT 3×3 (c) HLOIFT 5×5 (d) HLOIFT 10×10

(e) superpixel 3×3 (f) superpixel 5×5 (g) superpixel 10×10

Fig. 9. The segmentation of a CT image of the knee for different superpixel sizes.

	171×193	342×386	684×772	1368×1544
IFT [45]	8.46	29.26	106.61	333.13
HLOIFT [33]	54.55	200.44	724.73	2,878.91
(10×10)	0.52	1.88	8.08	33.05
(5×5)	1.61	8.14	24.78	91.29
$3 \times 3()$	4.37	17.25	62.93	260.24

TABLE II
TIME IN MS FOR THE DIFFERENT METHODS AND IMAGE RESOLUTIONS.

In Table II, we show the execution time gains of the proposed approach in comparison to IFT [45] and the multiple region segmentation by the regular HLOIFT [33] without superpixels, for different image resolutions and superpixel sizes, for the segmentation of three regions in a CT image of the knee with inclusion and exclusion relations. The usage of superpixels in HLOIFT significantly reduced the size of the graph, resulting in a great saving of memory and computation time.Moreover, the segmentation results for different superpixel sizes were similar to those obtained by HLOIFT at the pixel level demonstrating the robustness of the proposed method (Figure 9).

V. CONCLUSIONS

We proposed a new graph-based algorithm, named as HLOIFT, for multi-region segmentation, allowing the integration of each region high-level priors with the hierarchical relations between them into a single energy optimization. Besides the theoretical contribution, our experiments show that good segmentation results can be obtained, even when considering a simple measure of intensity dissimilarity. Besides being faster than hierarchical min-cut/max-flow based approaches, it is also less restrictive, allowing globally optimal results for arbitrary hierarchies. Finally, we presented how the HLOIFT method can be easily extended to work with 3D images. Therefore, we also proposed a *superpixel*-based adaptation of the HLOIFT method, leading to a more efficient and adequate solution for multi-object segmentation in large images.

REFERENCES

[1] S. Gordon, I. Dolgopyat, I. Kahn, and T. R. Raviv, "Co-segmentation of multiple images into multiple regions: Application to mouse brain MRI," in *13th IEEE ISBI*, 2016, pp. 399–402.

[2] K. D. Toennies, *Guide to medical image analysis: methods and algorithms*. Springer Science & Business Media, 2012.

[3] E. Visser, M. C. Keuken, G. Douaud, V. Gaura, A.-C. Bachoud-Levi, P. Remy, B. U. Forstmann, and M. Jenkinson, "Automatic segmentation of the striatum and globus pallidus using mist: Multimodal image segmentation tool," *NeuroImage*, vol. 125, pp. 479 – 497, 2016.

[4] M. Jackowski and A. Goshtasby, *A Computer-Aided Design System for Segmentation of Volumetric Images*. Boston, MA: Springer US, 2005, pp. 251–272.

[5] S. Golodetz, I. Voiculescu, and S. Cameron, "Simpler editing of graph-based segmentation hierarchies using zipping algorithms," *Pattern Recognition*, vol. 70, pp. 44–59, 2017.

[6] J. Cousty, G. Bertrand, L. Najman, and M. Couprie, "Watershed cuts: Thinnings, shortest path forests, and topological watersheds," *PAMI, IEEE Transactions on*, vol. 32, no. 5, pp. 925–939, 2010.

[7] L. Grady, "Random walks for image segmentation," *PAMI, IEEE Transactions on*, vol. 28, no. 11, pp. 1768–1783, 2006.

[8] K. Ciesielski, J. Udupa, P. Saha, and Y. Zhuge, "Iterative relative fuzzy connectedness for multiple objects with multiple seeds," *Computer Vision and Image Understanding*, vol. 107, no. 3, pp. 160–182, 2007.

[9] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient ND image segmentation," *International journal of computer vision*, vol. 70, no. 2, pp. 109–131, 2006.

[10] X. Li, J. Chen, and H. Fan, "Interactive image segmentation based on grow cut of two scale graphs," in *Advances on Digital Television and Wireless Multimedia Communications*, W. Zhang, X. Yang, Z. Xu, P. An, Q. Liu, and Y. Lu, Eds. Springer Berlin Heidelberg, 2012, pp. 90–95.

[11] K. C. Ciesielski, R. Strand, F. Malmberg, and P. K. Saha, "Efficient algorithm for finding the exact minimum barrier distance," *Computer Vision and Image Understanding*, vol. 123, pp. 53 – 64, 2014.

[12] A. X. Falcão, J. Stolfi, and R. de Alencar Lotufo, "The image foresting transform: Theory, algorithms, and applications," *PAMI, IEEE Transactions on*, vol. 26, no. 1, pp. 19–29, 2004.

[13] K. C. Ciesielski, A. X. Falcão, and P. A. V. Miranda, "Path-value functions for which dijkstra's algorithm returns optimal mapping," *Journal of Mathematical Imaging and Vision*, Feb 2018.

[14] H. Isack, O. Veksler, M. Sonka, and Y. Boykov, "Hedgehog shape priors for multi-object segmentation," in *Proceedings of the IEEE Conference on CVPR*, 2016, pp. 2434–2442.

[15] V. Gulshan, C. Rother, A. Criminisi, A. Blake, and A. Zisserman, "Geodesic star convexity for interactive image segmentation," in *Proceedings of CVPR*, 2010, pp. 3129–3136.

[16] O. Veksler, "Star shape prior for graph-cut image segmentation," *Computer Vision–ECCV 2008*, pp. 454–467, 2008.

[17] L. Gorelick, O. Veksler, Y. Boykov, and C. Nieuwenhuis, "Convexity shape prior for binary segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 2, pp. 258–271, 2017.

[18] D. Singaraju, L. Grady, and R. Vidal, "Interactive image segmentation via minimization of quadratic energies on directed graphs," in *CVPR, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.

[19] P. A. Miranda and L. A. Mansilla, "Oriented image foresting transform segmentation by seed competition," *Image Processing, IEEE Transactions on*, vol. 23, no. 1, pp. 389–398, 2014.

[20] A. Delong and Y. Boykov, "Globally optimal segmentation of multi-region objects," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 285–292.

[21] A. Delong, L. Gorelick, O. Veksler, and Y. Boykov, "Minimizing energies with hierarchical costs," *International journal of computer vision*, vol. 100, no. 1, pp. 38–58, 2012.

[22] J. Ulén, P. Strandmark, and F. Kahl, "An efficient optimization framework for multi-region segmentation based on lagrangian duality," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 2, pp. 178–188, 2013.

[23] Y. Yin, X. Zhang, R. Williams, X. Wu, D. D. Anderson, and M. Sonka, "LOGISMOS–layered optimal graph image segmentation of multiple objects and surfaces: cartilage segmentation in the knee joint," *IEEE transactions on medical imaging*, vol. 29, no. 12, pp. 2023–2037, 2010.

[24] I. Oguz and M. Sonka, "LOGISMOS–B: layered optimal graph image segmentation of multiple objects and surfaces for the brain," *IEEE transactions on medical imaging*, vol. 33, no. 6, pp. 1220–1235, 2014.

[25] P. A. V. Miranda, A. X. Falcao, and J. K. Udupa, "Cloud bank: A multiple clouds model and its use in mr brain image segmentation," in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, June 2009, pp. 506–509.

[26] J. K. Udupa, D. Odhner, Y. Tong, M. M. S. Matsumoto, K. C. Ciesielski, P. Vaideeswaran, V. Ciesielski, B. Saboury, L. Zhao, S. Mohammadianrasanani, and D. Torigian, "Fuzzy model-based body-wide anatomy recognition in medical images," 2013.

[27] L. Rittner, J. K. Udupa, and D. A. Torigian, "Multiple fuzzy object modeling improves sensitivity in automatic anatomy recognition," in *In Proceedings of SPIE on Medical Imaging: Image Processing*, 2014.

[28] K. Sun, J. K. Udupa, D. Odhner, Y. Tong, and D. A. Torigian, "Automatic thoracic anatomy segmentation on ct images using hierarchical fuzzy models and registration," in *In Proceedings of SPIE on Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling*, 2014.

[29] J. K. Udupa, D. Odhner, L. Zhao, Y. Tong, M. M. Matsumoto, K. C. Ciesielski, A. X. Falcao, P. Vaideeswaran, V. Ciesielski, B. Saboury, S. Mohammadianrasanani, S. Sin, R. Arens, and D. A. Torigian, "Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images," *Medical Image Analysis*, vol. 18, no. 5, pp. 752 – 771, 2014.

[30] S. Mohammadianrasanani, "The use of a body-wide automatic anatomy recognition system in image analysis of kidneys," Master's thesis, School of Technology and Health, Royal Institute of Technology, 2013.

[31] Y. Tong, J. K. Udupa, D. Odhner, S. Sin, and R. Arens, "Abdominal adiposity quantification at MRI via fuzzy model-based anatomy recognition," in *In Proceedings of SPIE on Medical Imaging: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 2013.

[32] L. M. C. Leon and P. A. V. D. Miranda, "Multi-object segmentation by hierarchical layered oriented image foresting transform," in *2017 30th SIBGRAPI Conference*, Oct 2017, pp. 79–86.

[33] L. M. C. Leon, K. Ciesielski, and P. A. V. Miranda, "Efficient hierarchical multi-object segmentation in layered graph (submitted)," p. https://www.math.wvu.edu/ kcies/SubmittedPapers/SS29.HLOIFT.pdf.

[34] L. C. Leon and P. V. de Miranda, "Efficient interactive multi-object segmentation in medical images," in *Computer Vision – ECCV 2018 Workshops*. Springer International Publishing, 2019, pp. 705–710.

[35] L. M. C. Leon and P. A. V. Miranda, "A hierarchical layered graph approach for multi-label segmentation in 2d medical images," 2018.

[36] L. M. C. Leon, K. Ciesielski, and P. A. V. Miranda, "Extensions of the hierarchical graph approach in multi-region segmentation," LatinX in AI Workshop @ ICML 2019, Long Beach, United States.

[37] ——, "Extensions of the hierarchical graph approach in multi-region segmentation," Women in Computer Vision Workshop @ CVPR 2019, Long Beach, United States.

[38] E. W. Dijkstra, "A note on two problems in connexion with graphs," *NUMERISCHE MATHEMATIK*, vol. 1, no. 1, pp. 269–271, 1959.

[39] P. de Miranda, A. Falcão, and J. Udupa, "Synergistic arc-weight estimation for interactive image segmentation using graphs," *Computer Vision and Image Understanding*, vol. 114, no. 1, pp. 85 – 99, 2010.

[40] K. C. Ciesielski and J. K. Udupa, "Affinity functions in fuzzy connectedness based image segmentation i: Equivalence of affinities," *Comput. Vis. Image Underst.*, vol. 114, no. 1, pp. 146–154, Jan. 2010.

[41] L. Mansilla and P. Miranda, "Image segmentation by oriented image foresting transform with geodesic star convexity," in *CAIP*, vol. 8047, York, UK, Aug 2013, pp. 572–579.

[42] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE transactions on PAMI*, vol. 26, no. 9, pp. 1124–1137, 2004.

[43] E. B. Alexandre, A. S. Chowdhury, A. X. Falcao, and P. A. V. Miranda, "IFT-SLIC: A general framework for superpixel generation based on simple linear iterative clustering and image foresting transform," in *2015 28th SIBGRAPI Conference*, 2015, pp. 337–344.

[44] L. Soler, A. Hostettler, V. Agnus, A. Charnoz, J.-B. Fasquel, J. Moreau, A.-B. Osswald, M. Bouhadjar, and J. Marescaux, "3d image reconstruction for comparison of algorithm database : A patient-specific anatomical and medical image database," 2012. [Online]. Available: https://www.ircad.fr/research/3d-ircadb-02/

[45] P. A. Miranda and A. X. Falcão, "Links between image segmentation based on optimum-path forest and minimum cut in graph," *Journal of Mathematical Imaging and Vision*, vol. 35, no. 2, pp. 128–142, 2009.

[46] F. Malmberg, I. Nyström, A. Mehnert, C. Engstrom, and E. Bengtsson, "Relaxed image foresting transforms for interactive volume image segmentation," in *Proc.SPIE*, vol. 7623, 2010, pp. 7623 – 7623 – 11.