Random walkers on morphological trees: A segmentation paradigm

Francisco Javier Alvarez Padilla, Barbara Romaniuk, Benoît Naegel, Stéphanie Servagi-Vernat, David Morland, Dimitri Papathanassiou, Nicolas Passat

Random walkers on morphological trees: A segmentation paradigm

Francisco Javier Alvarez Padilla[a,b], Barbara Romaniuk[a], Benoît Naegel[c], Stéphanie Servagi-Vernat[d], David Morland[a,e], Dimitri Papathanassiou[a,e], Nicolas Passat[a]

[a] *Université de Reims Champagne Ardenne, CReSTIC EA 3804, 51097 Reims, France*
[b] *Universidad de Guadalajara, Guadalajara, Mexico*
[c] *Université de Strasbourg, CNRS, ICube, Strasbourg, France*
[d] *Département de Radiothérapie, Institut Godinot, Reims, France*
[e] *Département de Médecine Nucléaire, Institut Godinot, Reims, France*

Abstract

The problem of image segmentation is often considered in the framework of graphs. In this context, two main paradigms exist: in the first, the vertices of a non-directed graph represent the pixels (leading e.g. to the watershed, the random walker or the graph cut approaches); in the second, the vertices of a directed graph represent the connected regions, leading to the so-called morphological trees (e.g. the component-trees or the trees of shapes). Various approaches have been proposed for carrying out segmentation from images modeled by such morphological trees, by computing cuts of these trees or by selecting relevant nodes from descriptive attributes. In this article, we propose a new way of carrying out segmentation from morphological trees. Our approach is dedicated to take advantage of the morphological tree of an image, enriched by multiple attributes in each node, by using maximally stable extremal regions and random walker paradigms for defining an optimal cut leading to a final segmentation. Experiments, carried out on multimodal medical images emphasize the potential relevance of this approach.

Keywords: Segmentation, tree of shapes, component-tree, mathematical morphology, region-based attributes, PET/CT, multimodality

1. Introduction

Segmentation is one of the most crucial issues in image analysis. As a mid-level image processing task, consisting of partitioning the image support into several parts presenting —spectrally and/or semantically— homogeneous properties, segmentation constitutes a mandatory prerequisite to most high-level image analysis or computer vision procedures. The difficulty of segmentation lies in many factors, e.g., the ill-posed definition of "homogeneousness" in images, the variability of images for a same observed scene, or the rapid evolution of image acquisition devices, that continuously induces new methodological challenges.

In this context, a wide range of approaches derived from various domains (statistics, signal processing, machine learning, optimization...) have been involved in the development of image segmentation methods and tools. Within this large population, an important family is composed by the approaches developed in the framework of graph theory.

Basically, an image composed of pixels can be modeled as a graph, where each pixel is a vertex of the graph,

Email addresses: francisco.alvarez.udg@gmail.com (Francisco Javier Alvarez Padilla), barbara.romaniuk@univ-reims.fr (Barbara Romaniuk), b.naegel@unistra.fr (Benoît Naegel), stephanie.servagiver@reims.unicancer.fr (Stéphanie Servagi-Vernat), david.morland@reims.unicancer.fr (David Morland), d.papathanassiou@reims.unicancer.fr (Dimitri Papathanassiou), nicolas.passat@univ-reims.fr (Nicolas Passat)

whereas the spatial neighbouring between two pixels corresponds to an edge in the graph. In other words, the spatial support of the image is modeled by a non-directed graph $G = (V, E)$ with pixels/vertices in V and adjacency/edges in E. The spectral information carried by the image is then modeled as a valuation function $v : V \to \mathbb{V}$ such that $v(x) \in \mathbb{V}$ represents the value of the pixel/vertex x, sometimes enriched by a second valuation function $d = E \to \mathbb{V}$ such that $d((x, y))$ represents the derivative between the values of two adjacent pixels/vertices x and y.

Relying on this framework, image segmentation could then benefit from the usual methodologies on graphs, for instance, region-growing (Adams and Bischof, 1994; Sethian, 1999), optimal path finding (Udupa and Samarasekera, 1996; Heijmans et al., 2005), minimal spanning tree computation (Cousty et al., 2009; Najman et al., 2013). Beyond these approaches more or less directly inherited from the usual algorithmics on graphs, more sophisticated methodologies were progressively proposed and enriched. The most popular is certainly the watershed (Vincent and Soille, 1991), that led to many variants (Vachier and Meyer, 2005; Machairas et al., 2015). One can also cite more recent contributions proposed, e.g., by Boykov et al. (2001) (graph-cut), Felzenszwalb and Huttenlocher (2004), Grady (2006) (random walker), or Couprie et al. (2011) (power watershed). Many of these graph-based approaches were defined or further reformulated in the framework of mathematical morphology, see e.g. (Najman and Cousty, 2014) for a comprehensive survey.

In the meantime, a second family of graph-based methods was developed, also mainly in the framework of mathematical morphology where it contributed to the development of the so called connected operators (Salembier and Serra, 1995). Whereas the above family relies on a pointwise paradigm (1 pixel = 1 vertex), with non-directed graphs modeling the image, this second family relies on a region-based paradigm (1 connected component = 1 vertex), with directed graphs modeling hierarchies of partitions of the images.

In practice, these methods aim at defining hierarchies of regions of the images, generally organized as trees (i.e., rooted, connected, acyclic graphs) where each vertex (also called "node") is a connected subset of the graph $G = (V, E)$ whereas the edges model the inclusion between connected subsets/nodes within the tree. Various kinds of such structures, often called *morphological trees*, were developed, including non-exhaustively: the (gray-level) component-tree (Salembier et al., 1998) and its multivalued (Kurtz et al., 2014) and incremental variants (Xu et al., 2016); the (gray-level) tree of shapes (Monasse and Guichard, 2000) and its multivalued variant (Carlinet and Géraud, 2015); the binary partition tree (Salembier and Garrido, 2000) and variants such as the α-tree (Soille, 2008) or the multifeature tree (Randrianasoa et al., 2018); the hyperconnection tree (Perret et al., 2012); or the hierarchical watershed (Perret et al., 2018). It is worth mentioning that some non-tree morphological hierarchies were also proposed, generally defined as directed acyclic graphs (e.g., component-graphs (Passat and Naegel, 2014; Passat et al., 2019), braids (Kiran and Serra, 2015; Tochon et al., 2019)) or as multigraph-like structures (e.g., component-hypertrees (Passat and Naegel, 2011; Morimitsu et al., 2020), asymmetric hierarchies (Perret et al., 2015)), leading to a more complex but richer modeling of spectral and/or spatial information.

Important efforts were geared towards handling such hierarchical structures, both for their construction (Carlinet and Géraud, 2014), and their actual manipulation with segmentation purposes, often leading to (quasi-)linear time costs, and opening the way to interactive applications (Westenberg et al., 2007; Passat et al., 2011). In particular, two main paradigms were developed, the first for defining relevant nodes within such hierarchies, via attribute-based filtering (Breen and Jones, 1996; Jones, 1999), the second for defining optimal cuts of the trees, corresponding to partitions of the underlying image (Guigues et al., 2006; Serra, 2011). These strategies led to many methods and tools dedicated to the segmentation of different kinds of images, in many contexts (e.g., medical imaging, remote sensing, document analysis, computer vision).

In this article, we describe a simple, yet efficient way of carrying out segmentation from morphological trees. In particular, the novelties of the proposed paradigm are the following. On the one hand, we aim at taking advantage of many images of a same scene, by computing vectorial attributes. On the other hand, we propose to carry out node selection within these trees with the random walker approach, coupled with maximally stable extremal regions (MSER) (Matas et al., 2004), relying on

the graph-structure of the tree, whereas random walker usually acts on the graph modeling the image.

In Section 2, we recall basic notions on morphological hierarchies and random walkers, in order to make the article mostly self-included. In Section 3, we describe the proposed method for image segmentation from a morphological hierarchy, emphasizing the tree modeling of the image, the MSER analysis of the nodes and the final tree cut computation by random walker. In Section 4, we exemplify the use of this segmentation approach in the case of medical imaging, namely for multimodal Positron Emission Tomography / X-ray Computed Tomography (PET / CT). Section 5 provides concluding remarks on the main perspectives offered by this approach.

2. Basic Notions

From now on, an image is considered as a non-directed graph $G = (V, E)$ where the vertices $x \in V$ correspond to the pixels, whereas the edges $(x, y) \in E$ correspond to the adjacency links between some pixels/vertices x and y. We also consider that the images are grey-level ones, i.e. that the valuation on the pixels/vertices is defined by $v : V \to \mathbb{Z}$.

2.1. Morphological hierarchies: Component-tree and tree of shapes

In the sequel we consider two of the most usual morphological hierarchies, namely the component-tree and the tree of shapes, which are illustrated in Figure 1.

Component-tree. Popularized by Salembier et al. (1998), the component-tree models the valued graph (V, E, v) by storing all the connected components induced by all the subgraphs (V_λ, E_λ) defined by either (1) $V_\lambda = \{x \in V \mid v(x) \geq \lambda\}$ or (2) $V_\lambda = \{x \in V \mid v(x) \leq \lambda\}$, and $E_\lambda = \{(x, y) \in E \mid x, y \in V_\lambda\}$. In case (1) (resp. (2)) the component-tree is also called the max-tree (resp. min-tree). The set of all these connected components (also called nodes) is denoted $\Theta^\star$ (where $\star$ is either $\geq$ or $\leq$). It is organized with respect to the inclusion relation $\subseteq$ on sets. More precisely, the component-tree $\tau^\star$ of (V, E, v) is defined as the Hasse diagram of the partially ordered set $(\Theta^\star, \subseteq)$. It is indeed a tree, i.e., a rooted, connected, acyclic, directed graph, with a root equal to V.

Tree of shapes. Popularized by Monasse and Guichard (2000), the tree of shapes can be seen as a self-dual version of the component-tree. More precisely, the set of nodes Θ of the tree of shapes of (V, E, v) is defined as the union $\Theta = H(\Theta^\leq) \cup H(\Theta^\geq)$ of the nodes of the min- and max-trees, pre-processed by a hole-filling procedure H. (For instance, the node $\{D, H\} \in \Theta^\geq$ of the max-tree of Figure 1(1.e), obtained from V_2 in Figure 1(1.c), leads to the node $H(\{D, H\}) = \{D, G, H, I\} \in \Theta$). Such as for the component-tree, these nodes are organized with respect to the inclusion relation. In other words, the tree of shapes ϕ of (V, E, v) is defined as the Hasse diagram of the partially ordered set $(\Theta, \subseteq)$. It is, of course, also a tree, with a root equal to V.

2.2. A graph-based segmentation paradigm: The random walker

The random walker, introduced by Grady (2006) is a marker-based segmentation approach, dedicated to process an image modeled as a valued graph (V, E, v). Given subsets of non-overlapping vertices V_F and V_B of V, the purpose is to determine for each remaining vertex $x \in V \setminus (V_F \cup V_B)$, if x is closer from V_F or V_B in (V, E).

This notion of closeness is related to the trajectory of a walker starting from x and trying to reach V_F or V_B by randomly progressing over the edges of E with a probability determined by a valuation $d : E \to \mathbb{R}_+$. In practice, the probability of the walker to pass through an edge $(x, y) \in E$ (i.e., to go from x to y) is inversely proportional to d, which is itself defined as a gradient-like measure between $v(x)$ and $v(y)$. In other words, the more (resp. less) similar the values of x and y, the higher (resp. lower) the probability that the walker goes from x to y.

This edge-valued graph can be expressed as a positive semidefinite combinatorial Laplacian matrix. By analogy with the computation of potentials in electrical circuits, it is possible to determine the probability to reach V_F before V_B for each vertex $x \in V$ by solving a combinatorial Dirichlet problem. In particular this can be done by solving a linear equation system of size proportional to the size of the graph, but highly sparse due to the low combinatorics of the adjacency relation on vertices. The result of this process provides a fuzzy segmentation of the vertices, that can be turned into a crisp segmentation by a 0.5 thresholding.

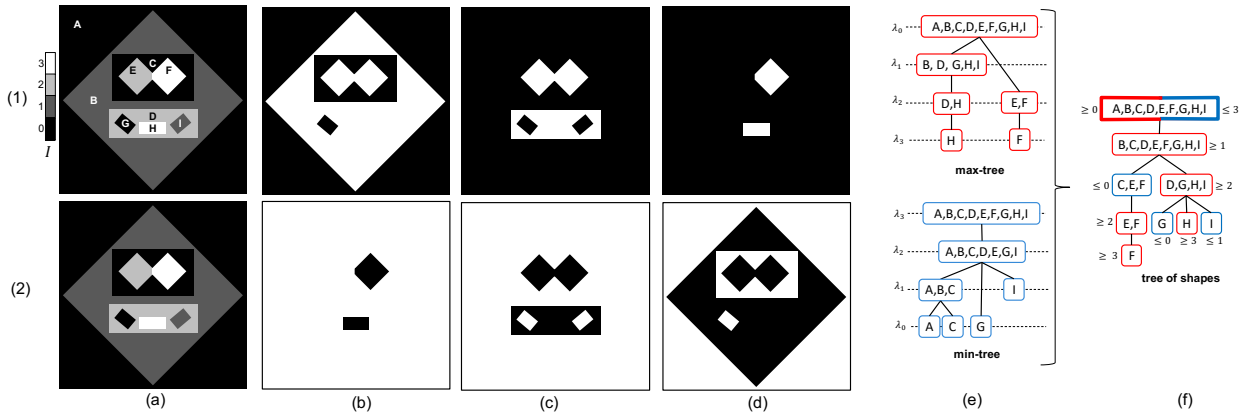

Figure 1: (a) A grey-level image defined on a set of values $[0, 3] \subset \mathbb{Z}$. Line (1): In white, the sets V_1 (b), V_2 (c) and V_3 (d) in the case of the max-tree (the set V_0, equal to the whole image support is not depicted); the connected components of these binary images form the set of nodes $\Theta^{\geq}$. Line (2): In white, the sets V_2 (b), V_1 (c) and V_0 (d) in the case of the min-tree (the set V_3, equal to the whole image support is not depicted); the connected components of these binary images form the set of nodes $\Theta^{\leq}$. (e.1) The Hasse diagram of $(\Theta^{\geq}, \subseteq)$, namely the max-tree of the image (a). (e.2) The Hasse diagram of $(\Theta^{\leq}, \subseteq)$, namely the min-tree of the image (a). (f) The Hasse diagram of $(\Theta, \subseteq)$, namely the tree of shapes of the image (a).

3. Method

We now describe our segmentation method. It is composed of three successive steps: (1) modeling of the image/graph as a tree, and computation of multiple attributes for each node; (2) filtering of the tree, for discarding non-relevant node based on MSER analysis; and (3) random walker segmentation of the remaining tree for determining the probability of each node to be preserved or removed in the segmented image.

3.1. Tree modeling and vectorial node valuation

The initial image, defined as a valued graph (V, E, v) is first modeled as a tree structure, namely a component-tree τ or a tree of shapes ϕ. The choice of the tree depends on the kind of image, and more precisely on the correlation between the regions of interest and their spectral properties. Indeed, if these regions are either of maximal or minimal grey-level value, we may model the image by a component-tree (a max-tree if they have maximal values, or a min-tree if they have minimal values). However, if the regions are of extremal values (i.e. possibly both minimal or maximal values), we may consider a tree of shapes, that builds upon the isocoutours within the image, in a more symmetric way than component-trees.

The built morphological hierarchy, noted ψ (that can be either τ or ϕ), is composed of nodes that represent connected regions within the image/graph. The goal of the segmentation process consists of determining which of these regions (and by side effect, which of the vertices/pixels) have to be preserved or removed in the resulting image.

To guide this decision, each node is enriched with attributes that characterize the corresponding region from various points of view (e.g., geometrically, morphologically, spatially, spectrally). This strategy was initially introduced by Breen and Jones (1996) for the so-called attribute opening. By considering a single attribute, it is possible to define a scalar mapping from the nodes of ψ to $\mathbb{R}$, opening the way to simple, threshold-based node selection. This strategy was considered, in particular by Jones (1999) for attribute-based filtering of trees.

It is however possible to define no longer one, but many attributes, for each node, leading to —more informative— vectorial attributes (Urbach et al., 2005), at the cost of a more complex handling of the induced attribute space. This is the strategy we consider here, where we compute, at each node N of the tree, a vector $A(N) = [a_1, a_2, ..., a_n] \in \mathbb{R}^n$ of n scalar attributes.

3.2. MSER tree filtering

Before actually carrying out node segmentation within the morphological hierarchy, we first aim at extracting from this tree a reduced data structure, composed of a subset of nodes that present relevant properties in terms of stability with respect to the chosen attributes. In other words, we aim at removing "noisy" nodes, in order to make the further segmentation more robust, from a more compact tree, thus improving both computational cost and accuracy.

To this end, our approach builds upon the notion of Maximally Stable Extremal Regions (MSER) democratized by Matas et al. (2004). Basically, MSER analysis determines the nodes that exhibit a sufficient stability with respect to a given attribute, based on a finite difference analysis along the branches of the tree, from the more distal nodes up to the root.

In their seminal work, Matas et al. (2004) considered in particular the volume-based stability, by assuming that a node N_i is stable if the estimator $s_{\text{vol}}(N_i) = |N_{i+\Delta} \setminus N_{i-\Delta}|/|N_i|$ reaches a local minimum at i, where $\Delta \in \mathbb{N}$ represents the gap between the node and those on the branch of the tree, considered for comparison purpose. In our implementation, we consider the formulation of MSER such as stated by Xu et al. (2014). Additionally, we extend the general metric as follows:

$$s_h(N_i) = \sum_{j=-\Delta}^{\Delta-1} |A(N_{i+1-j})[h] - A(N_{i-j})[h]| \qquad (1)$$

(If the h-th attribute is incremental, such as for the volume attribute, then we can normalize the result by $A(N_i)[h]$.)

Using this formula, it is possible to carry out MSER analysis for all the n attributes of the $A(N)$ vector for each node N of the tree.

For determining if a node is stable for a given attribute, we compute its MSER score for all the paths induced by the branches of its subtree. The node is then considered as stable for the attribute if it is considered as stable for at least one path, following a variant of "max" policy.

Finally, in order to fuse the results regarding the different attributes in a regularized way, a majority vote is performed, for discriminating the nodes that are mostly stable with respect to the different attributes from those that are marginally stable. The least stable nodes, characterized by a minority of local minima over the n MSER analyses, are pruned from the tree. This process leads to a tree with a lower number of nodes, presenting a greater relevance with respect to the attributes.

3.3. Random walker tree cut computation

The last step then consists of (i) labeling some nodes of the tree, and (ii) propagating these labels onto the initially non-labeled nodes, in order to finally obtain a completely labeled tree, inducing a labeled image corresponding to the final segmentation.

Step (i) can be carried out in various ways. The chosen policy is application-dependent, and not discussed here. An example of voxel-based node labeling is given in Section 4.1.1.

The crucial part of the process is indeed Step (ii), illustrated in Figure 2. At this stage, we consider that we have $k \geq 2$ distinct labels: one background label B and $k-1$ foreground label(s) F_i ($1 \leq i \leq k-1$).

In order to carry out label propagation, we first take advantage of the structure of the tree, that guarantees the uniqueness of the path between two nodes. In particular, we apply the following rules:

1. If a node N has a foreground label F_i, then all the nodes within the subtree rooted at N have the same label F_i.

2. If a node N has the background label B and none of the nodes between N and the root of the tree has a foreground label F_i, then all the nodes on this path are labeled as B.

Once these two rules have been applied, the remaining unlabeled regions of the tree correspond to sets of linear paths bounded on their proximal side by nodes B and on their distal side by nodes F_i.

We consider a random walker paradigm for each of these linear paths, composed of p nodes, connected pairwise by $p-1$ edges. The purpose is then to label the $p-2$ internal nodes, i.e., to determine the unique frontier between the region labeled B and the one labeled F_i. To this end, the random walker is applied on the path, with the two extremal nodes as seeds, and with a valuation d defined on each edge (N_i, N_j) as follows:

$$d((N_i, N_j)) = 1 - d_M(\beta.A(N_i), \beta.A(N_j)) \qquad (2)$$

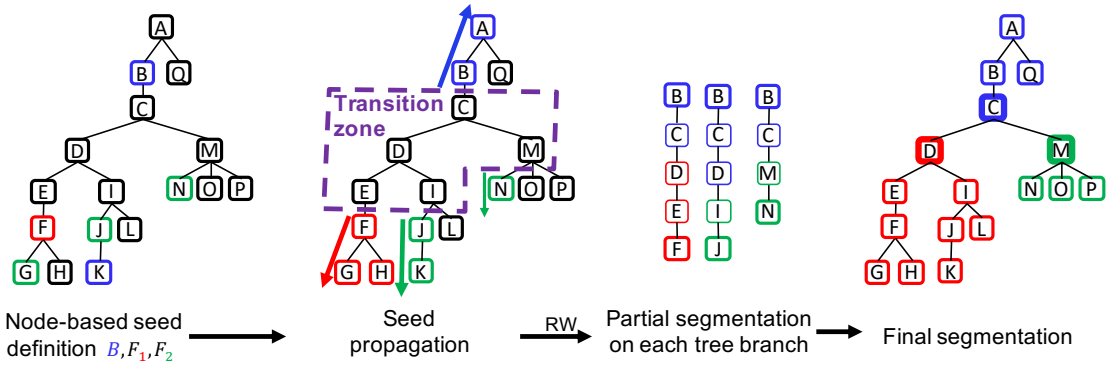

Figure 2: Successive steps of the random walker segmentation of the tree. From left to right: seed definition; label propagation within the proximal and distal parts; random walker segmentation in the non-labeled branches; final labeling by merging of the random walker results. See Section 2.2.

where $\beta \in [0,1]^n$ with $\sum_{h=1}^{n} \beta[h] = 1$ is a vector weighting the different attributes of the nodes; $\beta.A(N_\star) = [\beta[h].N_\star[h]]_{h=1}^{n}$; and d_M is the Mahalanobis distance, used in order to normalize the results despite potentially heterogeneous attributes.

Since a node can belong to many paths, it may happen that it is labeled in different ways by the random walkers. The disambiguation is carried out by considering that a label F_i has a higher priority than the label B, and that rules (1), (2) stated above still apply. When $k > 2$, a last ambiguous case may occur if one node with its parent node unanimously labeled B is assigned different F_i labels in different paths. In that case, the disambiguation is made by considering the label F_i that received the highest probability.

4. Application Example: PET / CT Image Segmentation

We now illustrate the behaviour of our method in the emerging field of tumor segmentation from PET / CT imaging. This is a complex task, due to tumor properties and imaging limitations. A tumor is often irregular in terms of shape and size, and variable in terms of position and cellular composition. Positron Emission Tomography (PET), coupled with X-ray Computed Tomography (CT), is widely used for tumor detection and analysis. Indeed, PET provides high spectral resolution, whereas CT provides high spatial resolution, allowing to obtain reliable information from both morphological and functional points of view.

Tumor segmentation from PET / CT images is a recent research area. A survey of the first methods considered to tackle that issue may be found in (Foster et al., 2014). Many of those actually used in clinical routine rely on thresholding of the PET images, either with fixed or adaptive values (Nestle et al., 2007; Yu et al., 2009), tuned with respect to settings on physical phantoms.

Our purpose is here to illustrate the potential relevance of our approach (in particular with respect to the choice of tree and attributes), but not to quantitatively compare it to other related methods. Such experiments deserve a dedicated study, and constitute a perspective work (see Section 5).

4.1. *Instanciation of the method*

PET / CT is considered here for tumor analysis in the context of radiotherapy. In such case, both tumor active viable areas (visible in PET) and the whole tumor, including necrotic parts (visible in CT) have to be considered. Multimodality is exploited by taking advantage of the spatial correspondence between PET and CT. The tree computed from the PET image is used as tree of reference for the segmentation process. The tree computed from the CT image is used to bring additional information via attributes computed from its nodes. Beforehand, PET is

rescaled to the higher CT resolution in order to actually reach a voxel-to-voxel correspondence.

4.1.1. Seed selections

The segmentation process requires the definition of seeds. It is convenient for the user to define these seeds directly in the image (voxel selection), whereas they will have to be embedded in the tree (node selection) for carrying out the random walker segmentation.

The voxel seed selection is carried out by defining two kinds of markers in the image: foreground (F), for tumors, and background (B) for non-tumor areas. In the PET image, tumors exhibit high intensity values, due to high biomarker uptake. Thus, a simple double-thresholding process can be considered by setting:

$$F = \{x \mid PET(x) > T_F \cdot \max_y\{PET(y)\}\} \quad (3)$$

$$B = \{x \mid PET(x) < T_B \cdot \max_y\{PET(y)\}\} \quad (4)$$

where $PET(x)$ is the value of the voxel x in the PET image, and $T_B, T_F \in [0, 1]$, with $T_B < T_F$, ensuring that $F \cap B = \emptyset$.

This voxel-based seed definition then has to be translated into a node-based seed definition. Since a voxel generally belongs to many nodes (forming a same branch of the tree), the correspondence between a voxel seed and a node seed is not one-to-one. In order to tackle this issue, the following policy is applied. Let $x \in F$ be a foreground seed voxel. Let $\{N_i(x)\}_{i=0}^{k}$ ($k \geq 0$) be the set of all the nodes containing x, such that for all $0 \leq i < j \leq k$, we have $N_j(x) \subset N_i(x)$. (Note that $N_k(x)$ is then the smallest node containing x, whereas $N_0(x)$ is the largest one, namely the root of the tree.) We define

$$N_F(x) = \arg\max_{\subseteq}\{N_i(x) \mid N_i(x) \cap B = \emptyset\} \quad (5)$$

In other words, we choose as foreground seed node $N_F(x)$ associated to the voxel x the largest node that contains x without containing any background seed voxels. It may happen that the set $\{N_i(x) \mid N_i(x) \cap B = \emptyset\}$ be empty. In such case, the seed voxel does not induce the definition of a seed node. Following the same notations, we define for each background seed voxel x the background seed node $N_B(x)$ by substituting min to max and B to F in Equation (5).

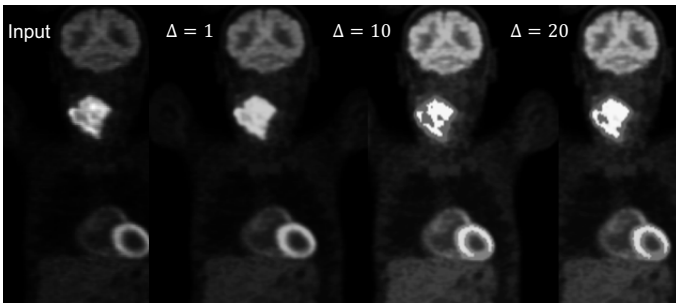

Figure 3: Filtering of the tree of shapes ϕ(PET,–) for various values of Δ, visualized via the reconstructed image. For $\Delta = 1$, 10 and 20, the ratio of discarded nodes is 98.32%, 99.17%, and 99.53%, respectively.

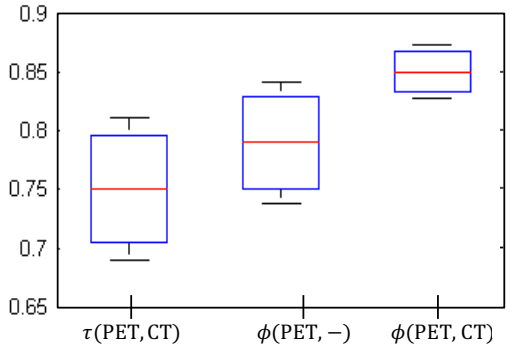

Figure 4: Dice scores (box plot) of the segmentation results for the various kinds of trees (see Section 4.2).

4.1.2. Attributes

For the proposed experiments, four attributes are considered, with respect to the properties of the sought structures of interest in the PET / CT images: difference of values between a node and its parent (C); number of voxels in the node (V); distance between the barycenter of the node and that of its parent node (D); compactness (K), defined as $\pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}/S$ (with S the number of voxels on the boundary of the node).

4.2. Results and discussion

The data involved in the proposed experiments are 33 ^{18}F-FDG PET / CT images (with 8 CT injected with iodinated contrast agent, namely CE-CT). They were acquired from several patients (12 lung cancers, 8 head and

neck cancers, 3 breast cancers, 3 lymphomas, 2 hepatic cancers, 3 cervix cancers, 1 colon cancer and 1 sarcoma).

Patients received an intravenous dose of FDG (3 MBq/kg). Portal phase CT was performed first in auto mA mode with adaptive statistical iterative reconstruction, native collimation of 16×1.25 mm, resolution matrix of $512 \times 512 \times 343$ with a voxel size of $0.97 \times 0.97 \times 2.5$ mm^3. PET acquisition was then performed 60 minutes after FDG injection, covering the area from the skull to proximal thighs. PET was reconstructed using an iterative algorithm (OSEM 24 subsets, 2 iterations) and post-processed with a Butterworth filter (cut-off frequency: 6.4 mm), obtaining a $256 \times 256 \times 263$ matrix with a voxel size of $2.73 \times 2.73 \times 3.27$ mm^3. For the case of ^{18}F-FDG PET / CE-CT, an intravenous dose of iodinated contrast agent (Ioversol – Optiject 350 mg I/ml) was injected 70 seconds before CT phase.

Three kinds of morphological hierarchies were considered (one monomodal, two multimodal):

- τ(PET,CT): a component-tree built from the PET image, and enriched with attributes from the CT image;
- ϕ(PET,–): a tree of shapes built from the PET image;
- ϕ(PET,CT): a tree of shapes built from the PET image, and enriched with attributes from the CT image.

For a tree $\psi(X, Y)$, built from image X and enriched with attribute values from image Y, we considered the four attributes V, C, D, K for image X and the attribute C for image Y.

A majority vote MSER filtering was applied on ψ with $\Delta = 1$ (see Equation (1)). This value was experimentally set by observing its impact on the filtered tree and its underlying image. Figure 3 exemplifies the influence of Δ in the case of ϕ(PET,–).

Two labels are defined, leading to a binary segmentation paradigm: F for foreground (tumor), and B for background. For seed selection, the parameters T_B and T_F were experimentally set to 0.15 and 0.60, respectively (see Equations (3–4)). The weight vector $\beta \in [0, 1]^5$ was set to $[1/h]_{i=1}^5$ (see Equation (2)).

With these settings, one can observe that the best results were obtained for ϕ(PET,CT), by considering the Dice score as quality metric for comparing the segmentation results and expert-defined ground-truths; see Figure 4. On the one hand, these results emphasize the relevance of using the tree of shapes instead of the component-tree, even for modeling PET images (that should present regions of interest only at locally maximal values). This derives from the fact that lower contrast necrosis areas, forming "holes" in the PET high contrast areas, are captured by the tree of shapes whereas they are not by the component-tree. On the other hand, these results also shed light on the relevance of considering multimodal attributes, since the segmentation results are better with ϕ(PET,CT) than with ϕ(PET,–), i.e. by considering attributes from both PET and CT instead of PET only. This is illustrated in Figure 5.

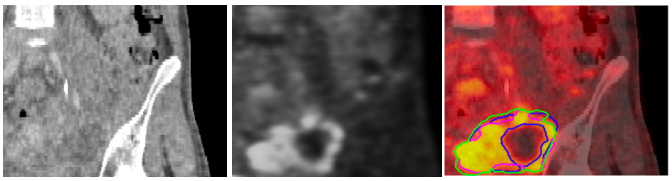

Figure 5: From left to right: contrast-enhanced CT, PET and fusion of the two modalities; coronal view of the left side of the pelvic area. In the rightmost image, coloured boundaries correspond to the segmentation obtained from τ(PET,CT) (blue), ϕ(PET,–) (pink) and ϕ(PET,CT) (green).

5. Conclusion

Our proposed approach constitutes an attempt of embedding the random walker —and more generally graph-cut approaches initially designed for non-directed graphs— in hierarchical structures. This paradigm was proposed here in the case where these hierarchies are indeed trees. However, it could be extended to the case where such hierarchies are more complex as trees, e.g., directed acyclic graphs such as the component-graphs (Grossiord et al., 2019). This will constitute the main methodological perspective of this work.

From an applicative point of view, the next step will be to assess the relevance of the proposed approach in the context of PET / CT image segmentation. Indeed, in the framework of graph-based approaches, Bagci et al. (2013) proposed to perform random walker segmentation separately in various modalities and then to fuse the results. Alternatively, Ju et al. (2015) proposed to consider the

graph cut approach for bimodal segmentation from a co-graph modeling the PET / CT images. In the framework of morphological hierarchies, Grossiord et al. (2020) and Urien et al. (2017) used a single-attribute component-tree analysis for PET images.

Vectorial attributes obtained from the component-trees of PET images were also involved by Grossiord et al. (2017) for feeding a random forest classification. Comparing our approach with these different methods will be a way of assessing the advantages of coupling graph-based techniques with morphological hierarchies in the context of multimodal medical imaging.

From a theoretical point of view, a longer time perspective will consist of investigating the potential links that may exist between the proposed approach and the general framework of power watersheds (Couprie et al., 2011; Najman, 2017) that already allowed to successfully unify various kinds of graph-based segmentation paradigms.

Acknowledgements

This work was partly funded by the "*Programme Investissement d'Avenir*", run by the French *Agence Nationale de la Recherche* (ANR-11-INBS-0006).

References

Adams, R., Bischof, L., 1994. Seeded region growing. IEEE Trans. on Pattern Analysis and Machine Intelligence 16, 641–647.

Bagci, U., Udupa, J., Mendhiratta, N., Foster, B., Xu, Z., Yao, J., Chen, X., Mollura, D., 2013. Joint segmentation of anatomical and functional images: Applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images. Medical Image Analysis 17, 929–945.

Boykov, Y., Veksler, O., Zabih, R., 2001. Fast approximate energy minimization via graph cuts. IEEE Trans. on Pattern Analysis and Machine Intelligence 23, 1222–1239.

Breen, E.J., Jones, R., 1996. Attribute openings, thinnings, and granulometries. Computer Vision and Image Understanding 64, 377–389.

Carlinet, E., Géraud, T., 2014. A comparative review of component tree computation algorithms. IEEE Trans. on Image Processing 23, 3885–3895.

Carlinet, E., Géraud, T., 2015. MToS: A tree of shapes for multivariate images. IEEE Trans. on Image Processing 24, 5330–5342.

Couprie, C., Grady, L.J., Najman, L., Talbot, H., 2011. Power watershed: A unifying graph-based optimization framework. IEEE Trans. on Pattern Analysis and Machine Intelligence 33, 1384–1399.

Cousty, J., Bertrand, G., Najman, L., Couprie, M., 2009. Watershed cuts: Minimum spanning forests and the drop of water principle. IEEE Trans. on Pattern Analysis and Machine Intelligence 31, 1362–1374.

Felzenszwalb, P.F., Huttenlocher, D.P., 2004. Efficient graph-based image segmentation. International Journal of Computer Vision 59, 167–181.

Foster, B., Bagci, U., Mansoor, A., Xu, Z., Mollura, D., 2014. A review on segmentation of positron emission tomography images. Computers in Biology and Medicine 50, 76–96.

Grady, L., 2006. Random walks for image segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence 28, 1768–1783.

Grossiord, É.., Naegel, B., Talbot, H., Najman, L., Passat, N., 2019. Shape-based analysis on component-graphs for multivalued image processing. Mathematical Morphology – Theory and Applications 3, 45–70.

Grossiord, É., Passat, N., Talbot, H., Naegel, B., Kanoun, S., Tal, I., Tervé, P., Ken, S., Casasnovas, O., Meignan, M., Najman, L., 2020. Shaping for PET image analysis. Pattern Recognition Letters 131, 307–313.

Grossiord, É., Talbot, H., Passat, N., Meignan, M., Najman, L., 2017. Automated 3D lymphoma lesion segmentation from PET/CT characteristics, in: ISBI, pp. 174–178.

Guigues, L., Cocquerez, J.P., Le Men, H., 2006. Scale-sets image analysis. International Journal of Computer Vision 68, 289–317.

Heijmans, H.J.A.M., Buckley, M., Talbot, H., 2005. Path openings and closings. Journal of Mathematical Imaging and Vision 22, 107–119.

Jones, R., 1999. Connected filtering and segmentation using component trees. Computer Vision and Image Understanding 75, 215–228.

Ju, W., Xiang, D., Zhang, B., Wang, L., Kopriva, I., Chen, X., 2015. Random walk and graph cut for co-segmentation of lung tumor on PET-CT images. IEEE Trans. on Image Processing 24, 5854–5867.

Kiran, B.R., Serra, J., 2015. Braids of partitions, in: ISMM, pp. 217–228.

Kurtz, C., Naegel, B., Passat, N., 2014. Connected filtering based on multivalued component-trees. IEEE Trans. on Image Processing 23, 5152–5164.

Machairas, V., Faessel, M., Cárdenas-Peña, D., Chabardès, T., Walter, T., Decencière, E., 2015. Waterpixels. IEEE Trans. on Image Processing 24, 3707–3716.

Matas, J., Chum, O., Urban, M., Pajdla, T., 2004. Robust wide-baseline stereo from maximally stable extremal regions. Image and Vision Computing 22, 761–767.

Monasse, P., Guichard, F., 2000. Scale-space from a level lines tree. Journal of Visual Communication and Image Representation 11, 224–236.

Morimitsu, A., Passat, N., Alves, W.A.L., Hashimoto, R.F., 2020. Efficient component-hypertree construction based on hierarchy of partitions. Pattern Recognition Letters 135, 30–37.

Najman, L., 2017. Extending the power watershed framework thanks to γ-convergence. SIAM Journal on Imaging Sciences 10, 2275–2292.

Najman, L., Cousty, J., 2014. A graph-based mathematical morphology reader. Pattern Recognition Letters 47, 3–17.

Najman, L., Cousty, J., Perret, B., 2013. Playing with Kruskal: Algorithms for morphological trees in edge-weighted graphs, in: ISMM, pp. 135–146.

Nestle, U., Schaefer-Schuler, A., Kremp, S., Groeschel, A., Hellwig, D., Rübe, C., Kirsch, C.M., 2007. Target volume definition for 18F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer. European Journal of Nuclear Medicine and Molecular Imaging 34, 453–462.

Passat, N., Naegel, B., 2011. Component-hypertrees for image segmentation, in: ISMM, pp. 284–295.

Passat, N., Naegel, B., 2014. Component-trees and multivalued images: Structural properties. Journal of Mathematical Imaging and Vision 49, 37–50.

Passat, N., Naegel, B., Kurtz, C., 2019. Component-graph construction. Journal of Mathematical Imaging and Vision 61, 798–823.

Passat, N., Naegel, B., Rousseau, F., Koob, M., Dietemann, J.L., 2011. Interactive segmentation based on component-trees. Pattern Recognition 44, 2539–2554.

Perret, B., Cousty, J., Guimarães, S.J.F., Santana Maia, D., 2018. Evaluation of hierarchical watersheds. IEEE Trans. on Image Processing 27, 1676–1688.

Perret, B., Cousty, J., Tankyevych, O., Talbot, H., Passat, N., 2015. Directed connected operators: Asymmetric hierarchies for image filtering and segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence 37, 1162–1176.

Perret, B., Lefèvre, S., Collet, C., Slezak, E., 2012. Hyperconnections and hierarchical representations for grayscale and multiband image processing. IEEE Trans. on Image Processing 21, 14–27.

Randrianasoa, J.F., Kurtz, C., Desjardin, E., Passat, N., 2018. Binary Partition Tree construction from multiple features for image segmentation. Pattern Recognition 84, 237–250.

Salembier, P., Garrido, L., 2000. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. IEEE Trans. on Image Processing 9, 561–576.

Salembier, P., Oliveras, A., Garrido, L., 1998. Anti-extensive connected operators for image and sequence

processing. IEEE Trans. on Image Processing 7, 555–570.

Salembier, P., Serra, J., 1995. Flat zones filtering, connected operators, and filters by reconstruction. IEEE Trans. on Image Processing 4, 1153–1160.

Serra, J., 2011. Hierarchies and optima, in: DGCI, pp. 35–46.

Sethian, J.A., 1999. Fast marching methods. SIAM Review 41, 199–235.

Soille, P., 2008. Constrained connectivity for hierarchical image partitioning and simplification. IEEE Trans. on Pattern Analysis and Machine Intelligence 30, 1132–1145.

Tochon, G., Dalla Mura, M., Veganzones, M.A., Géraud, T., Chanussot, J., 2019. Braids of partitions for the hierarchical representation and segmentation of multimodal images. Pattern Recognition 95, 162–172.

Udupa, J.K., Samarasekera, S., 1996. Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation. CVGIP: Graphical Model and Image Processing 58, 246–261.

Urbach, E.R., Boersma, N.J., Wilkinson, M.H.F., 2005. Vector-attribute filters, in: ISMM, pp. 95–104.

Urien, H., Buvat, I., Rougon, N., Soussan, M., Bloch, I., 2017. Brain lesion detection in 3D PET images using max-trees and a new spatial context criterion, in: ISMM, pp. 455–466.

Vachier, C., Meyer, F., 2005. The viscous watershed transform. Journal of Mathematical Imaging and Vision 22, 251–267.

Vincent, L., Soille, P., 1991. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Trans. on Pattern Analysis and Machine Intelligence 13, 583–598.

Westenberg, M.A., Roerdink, J.B.T.M., Wilkinson, M.H.F., 2007. Volumetric attribute filtering and interactive visualization using the max-tree representation. IEEE Trans. on Image Processing 16, 2943–2952.

Xu, Y., Géraud, T., Najman, L., 2016. Connected filtering on tree-based shape-spaces. IEEE Trans. on Pattern Analysis and Machine Intelligence 38, 1126–1140.

Xu, Y., Monasse, P., Géraud, T., Najman, L., 2014. Tree-based morse regions: A topological approach to local feature detection. IEEE Transactions on Image Processing 23, 5612–5625.

Yu, W., Fu, X., Zhang, Y., Xiang, J., Shen, L., Jiang, G., Chang, J., 2009. GTV spatial conformity between different delineation methods by 18 FDG PET/CT and pathology in esophageal cancer. Radiotherapy and Oncology 93, 441–446.