Hierarchies and shape-space for PET image segmentation

Eloïse Grossiord, Hugues Talbot, Nicolas Passat, Michel Meignan, Pierre Tervé, Laurent Najman

HIERARCHIES AND SHAPE-SPACE FOR PET IMAGE SEGMENTATION

É. Grossiord[1,4]*, H. Talbot*[1]*, N. Passat*[2]*, M. Meignan*[3]*, P. Tervé*[4]*, L. Najman*[1]

[1] Université Paris-Est, ESIEE-Paris, LIGM, CNRS, France
[2] Université de Reims Champagne-Ardenne, CReSTIC, France
[3] The Lymphoma Academic Research Organisation (LYSARC), Lyon, France
[4] KeoSys, Nantes, France

ABSTRACT

Positron Emission Tomography (PET) image segmentation is essential for detecting lesions and quantifying their metabolic activity. Due to the spatial and spectral properties of PET images, most methods rely on intensity-based strategies. Recent methods also propose to integrate anatomical priors to improve the segmentation process. In this article, we show how the hierarchical approaches proposed in mathematical morphology can efficiently handle these different strategies. Our contribution is twofold. First, we present the component-tree as a relevant data-structure for developing interactive, real-time, intensity-based segmentation of PET images. Second, we prove that thanks to the recent concept of shaping, we can efficiently involve a priori knowledge for lesion segmentation, while preserving the good properties of component-tree segmentation. Preliminary experiments on synthetic and real PET images of lymphoma demonstrate the relevance of our approach.

Index Terms— Positron Emission Tomography, segmentation, component-tree, shaping, mathematical morphology, lymphoma.

1. INTRODUCTION

Positron Emission Tomography (PET) using 18-fluorine fluorodeoxyglucose (^{18}F-FDG) is recognized as the modality of choice for lymphoma imaging, due to its high sensitivity and specificity. PET imaging is now routinely used, not only to detect tumor lesions, but also to assess their metabolic activity, allowing diagnosis, staging and treatment response evaluation. Thus, efficient PET image analysis is essential in clinical practice. However, tumor segmentation is challenging, considering the limitations the modality suffers from: poor spatial resolution compared to anatomical imaging; partial volume effects and intrinsic noise that lowers the quality of reconstructed images; and physiological artifacts. Besides, the segmentation of lymphoma is a difficult task because of multiple tumors location, intensity distribution and contrast to its surrounding tissues. As a consequence, the development of efficient and easy-to-use image processing and analysis tools for guiding the expert's diagnosis is highly relevant.

In contrast to other standard 3D imaging modalities, PET provides metabolic activity visualization, characterized by the intensity of an injected radiotracer. Based on this assertion, most approaches proposed to analyse PET images only rely on this intensity information. On the one hand for metabolic assessment, intensity-based criteria such as the Standardized Uptake Value (SUV) [1] have been considered. On the other hand, for the detection and segmentation of lesions, intensity-based approaches have been developed. Indeed, the segmentation methods designed during the last ten years were mostly based on thresholding [2], region-growing [3], classification (FCM [4], FLAB [5]), watershed, or basic mathematical morphology pipelines (see [6] for a recent survey). Practically, such methods generally lead to interactive tools in clinical routines, where the expert user provides regions of interest (e.g., bounding-boxes) and/or seeds, and tune threshold values in order to delineate lesions in the chosen area(s). Despite the low quality of PET images, and based on the assertion that the intensity provides indirect spatial information on organs or lesions gathering the sought metabolic activity, few strategies were designed to integrate anatomical priors in order to improve the segmentation process [7, 8]. In contrast to the above intensity-based methods, they generally consider multimodal images (e.g., PET/CT or MRI/PET) to collect supplementary anatomical information, and rely on non-interactive and time-consuming strategies (e.g., deformable models, optimization schemes).

Our purpose is to propose a segmentation strategy combining the advantages of both approaches: the ability to rely on simple and real-time interaction paradigms, based on intensity and thresholding; and the ability to embed anatomical priors that improve the accuracy of segmentation, in particular by discriminating actual lesions from organs fixating the radiotracer. To reach that goal, we propose for the first time to consider the morphological hierarchy paradigm [9, Ch. 7] for designing PET-oriented connected operators. In Sec. 2, we propose to use the notion of component-tree [10] as a hierarchical data-structure allowing us to perform PET segmentation. Indeed, the component-tree is well-fitted, both for handling 3D images where the structures of interest correspond to extremal intensity values (as already proved in angiographic imaging [11]) and for developing interactive and real-time segmentation methods [12]. In Sec. 3, we describe a strategy based on the recently introduced notion of the shapings [13], that consists of using a two-layer component-tree, namely a "tree of tree", to embed additional information such as geometrical priors. Since the induced data-structure remains a tree, the desired properties described in Sec. 2, namely user interaction, thresholding paradigm, and real-time computation, remain valid, while naturally handling higher-level knowledge. In Sec. 4, we validate this methology on phantom images and on 3D PET images of lymphoma, where results are compared with expert's manual delineation.

2. COMPONENT-TREES AND GENERALIZED THRESHOLDING

2.1. Component-trees

From a structural point of view, the component-tree [10] of an image I is a data-structure $\mathcal{T}$ that models all the connected components (i.e., the maximally connected regions) of I, obtained from its successive binary level (thresholded) sets. More formally, the

component-tree $\mathcal{T}$ is composed of a set of nodes Θ, namely the connected components of the binary level sets, and of edges between these nodes, that correspond to the transitive opening of the partial order relation of inclusion $\subseteq$ on Θ. Each node $K \in \Theta$ is associated to the (maximal) value v of the level set where the associated connected component appears.

Various algorithms exist for computing the component-tree of an image, and the most efficient run in quasi-linear time [14]. In addition, retrieving an image from its component-tree is a trivial process that can be carried out in linear time. Processing an image via its component-tree is then a low-cost operation.

Moreover, the component-tree is a lossless image model. Indeed, if we consider the image I in its functional form, i.e., as a mapping $I : \Omega \to V$, where Ω is the image support, i.e., the set of its voxels, and V is its set of grey-level values, then the image I can be expressed as the supremum of the nodes of $\mathcal{T}$:

$$I = \bigvee_{K \in \Theta} C_K \qquad (1)$$

where $C_K : \Omega \to V$ is the cylinder function such that $C_K(x) = v$ for any $x \in K$ with v the maximal value that defines K as a connected component of the level set of I at threshold v, while $C_K(x) = 0$ anywhere else.

The composition formula (1) leads to a well-defined image for Θ, but also for any subset of nodes of Θ. In other words, it is possible to filter the image I by discarding some of the nodes, and then reconstructing a resulting image from the preserved nodes. Each point $x \in \Omega$ in the filtered image then presents a value that is lower or equal to the initial image; the induced operators are then anti-extensive.

This anti-extensive filtering paradigm was formalized in [10, 15], and is basically composed of three successive steps:

(i) construction of the component-tree $\mathcal{T}$ associated to I;

(ii) reduction of Θ, leading to a reduced component-tree $\widehat{\mathcal{T}}$; and

(iii) reconstruction of a filtered image $\widehat{I}$ induced from $\widehat{\mathcal{T}}$.

$$
\begin{array}{ccc}
I & \xrightarrow{\;(i)\;} & \mathcal{T} \\[2pt]
\downarrow & & \downarrow{\scriptstyle(ii)} \\[2pt]
\widehat{I} \leq I & \xleftarrow{\;(iii)\;} & \widehat{\mathcal{T}}
\end{array}
\qquad (2)
$$

2.2. PET image segmentation with component-trees

In the case of PET, the component-tree of an image is organized from the lowest to the highest values. More precisely, for a given PET image I, the root of its component-tree $\mathcal{T}$, i.e., the node K that is equal to the whole support Ω, is obtained from the lowest level set, at value 0. Conversely, the leaves of $\mathcal{T}$, i.e., the nodes K that are minimal elements for the inclusion relation $\subseteq$, are the flat zones of the image of (locally) maximal values, and then correspond to high metabolic activity areas of lesions and hyperfixating organs.

The main strength of the component-tree, is that it intrinsically models the space of *all* the potential thresholding operations that can be carried out on an image. In addition, the choice of a solution within this data-structure can be performed in linear time, by scanning the (decreasing) grey-level values along the branches of the tree, from its leaves to its root. In particular, this process is compliant with all major strategies, namely fixed thresholding (horizontal cut in the tree), adaptive thresholding (branch-specific cut in the tree), or bounding-box thresholding (explicit choice of leaves or branches to cut in the tree).

Component-trees then constitute a highly relevant model for intensity-based analysis of PET images. Its low space and time cost properties are crucial to perform PET image segmentation. In the next section, we propose a framework for performing component-tree-based segmentation by embedding a priori geometric information, without losing these complexity properties.

3. FROM INTENSITY TO ATTRIBUTE SEGMENTATION: THE SHAPING PARADIGM

3.1. Shaping: Trees of trees

Another strength of component-trees, that derives from the decomposition of the image into elementary patterns both in terms of space and intensity, is the ability to associate an attribute $a \in A$ to each node of the tree, that models additional information.

This attribute-based approach has been considered in the anti-extensive filtering framework described in Diag. (2). The choice of the nodes to preserve or discard depends on a Boolean predicate $\rho : A \to \{\text{true}, \text{false}\}$ acting on the attributes. When the attribute is increasing (i.e., $\rho(K)$ is true for a given node K implies the same for all the nodes between K and the root), the attribute can be handled similarly to the grey-level of the nodes, in the sense where the node selection proceeds as a simple thresholding. For instance, this is the case for attributes such as the number of voxels of the node, used for area/volume denoising filters. However, most attributes are not increasing, in general. This means that the validity of the predicate for a given node does not imply its validity for the rest of the branch. This is not an issue in the context of automated processing, where the tree is fully scanned, and ad hoc pruning policies can be applied a posteriori to homogenize the results in the different branches of the tree (see [10, 15] for a discussion on this topic).

In our case, as we wish to carry out threshold-based interactive segmentation, working with increasing attributes is crucial. To tackle this issue, we propose to rely on the recently introduced notion of *shapings* [13]. Basically, shapings consists of considering the component-tree $\mathcal{T}$ as an image. Indeed, from a functional point of view, it can be defined as a mapping $\mathcal{T} : \Theta \to A$, where voxels are replaced by nodes, while grey-levels are replaced by values of the attribute. This approach is tractable only if the space of the attribute values A is equipped with a total order relation, i.e., can be modeled as (a subset of) $\mathbb{R}$ or $\mathbb{Z}$; this is however the case of most attributes, in particular numerical ones.

In this case, the "image" $\mathcal{T}$ can be thresholded, and for any level set, the connected components, i.e., the sets of nodes that are linked by edges in $\mathcal{T}$, can be computed and ordered with respect to the inclusion relation. It is then possible to build a component-tree of the component-tree $\mathcal{T}$. This "tree of tree", noted $\mathcal{TT}$, can be processed as any other component-tree, and we can perform anti-extensive filtering with it. The virtue of this new tree is that the attribute computed from the nodes of $\mathcal{T}$ is now increasing in $\mathcal{TT}$, which can then be processed by a thresholding approach. The overall procedure remains quasi-linear in time and space, since we only duplicate the standard component-tree anti-extensive filtering process, as illustrated in Diag. (3):

$$
\begin{array}{ccccc}
I & \xrightarrow{\;(i)\;} & \mathcal{T} & \xrightarrow{\;(i)\;} & \mathcal{TT} \\[2pt]
\downarrow & & & & \downarrow{\scriptstyle(ii)} \\[2pt]
\widehat{I} \leq I & \xleftarrow{\;(iii)\;} & \widehat{\mathcal{T}} & \xleftarrow{\;(iii)\;} & \widehat{\mathcal{TT}}
\end{array}
\qquad (3)
$$

3.2. Attribute-based segmentation of PET images

We consider the paradigm of shapings to develop an interactive attribute-based procedure for discriminating active lesions from hyperfixating organs in PET images, in the context of lymphoma. As stated above, an intensity-based strategy, relying either on a single component-tree or any other – classification, region-growing, watershed – intensity-based approach, may over-emphasise the high intensity areas, including true positives, but also many false positives.

Our working hypothesis, that is relevant in the case of lymphoma – where the number of lesions is high and their shape is compact, compared to other fixating organs (e.g., heart, kidneys, etc.) – is that a compacity/elongation criterion can allow us to better discriminate the active lesions.

Then, we propose to define an attribute $\mathcal{E} : 2^\Omega \to \mathbb{R}$ that associates to any connected subset of Ω, a numerical value describing how the shape is compact, i.e., similar to a sphere, or elongated. This attribute is defined by computing the eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ associated to the eigenvectors of the 3×3 matrix of inertia of the (binary) shapes $K \subseteq \Omega$. We define

$$\mathcal{E}(K) = \lambda_3 / \lambda_1 \qquad (4)$$

This is the ratio between the lowest and highest eigenvalues, lying in $[0, 1]$. If $\mathcal{E}(K)$ is close to 1, then λ_3 and λ_1 (and thus λ_2) present similar values, and K then has a compact shape. When $\mathcal{E}(K)$ decreases toward 0, K becomes more elongated.

The PET image is first processed to build its component-tree $\mathcal{T}$, based on the total order of the metabolic signal. As stated above, in this first tree, the nodes that are close to the leaves correspond to the areas of highest intensity, but can be either lesions or hyperfixating organs. For each node $K \in \Theta$ of this tree, the attribute value $\mathcal{E}(K) \in [0, 1]$ is then computed and stored at the node.

The second component-tree $\mathcal{TT}$ is then defined from the "image" $\mathcal{T} : \Theta \to [0, 1]$, that takes its values in the shape space induced by $\mathcal{E}$. In this new tree, the root is the node of maximal value, while the leaves are those of locally minimal values. This second tree $\mathcal{TT}$ is a min-tree (i.e., a component-tree based on $(\mathbb{R}, \geq)$), while $\mathcal{T}$ was a max-tree (i.e., a component-tree based on $(\mathbb{R}, \leq)$). The nodes that are close to the leaves then correspond to the more elongated areas, i.e., the sets of connected components of Ω that are of least relevance, and should then be removed preferentially.

In order to segment the active lesions versus the hyperfixating organs, we propose two distinct strategies. The first is a thresholding along the branches of the tree $\mathcal{TT}$, thus pruning the distal parts of the tree, composed of the nodes that have a compactness value $\mathcal{E}$ lower than a chosen value $C \in [0, 1]$. The second is a differential thresholding that considers the "gap" between the compactness value $\mathcal{E}_K$ of the node K and the value $\mathcal{E}_L$ of the leave of its branch. This gap $\mathcal{E}_K - \mathcal{E}_L$ still varies in $[0, 1]$, with 0 values at the leaves. This criterion remains increasing, thus authorising a (relative) thresholding approach similar to the first strategy, but with a finer behaviour. This latter strategy is considered for the experiments described in the next section.

4. EXPERIMENTS AND RESULTS

4.1. Phantom images

To validate our approach, we first used the NEMA 2007 IEC image phantom. This phantom contains six spherical "lesions" with diameters 10, 13, 17, 22, 28 and 37 mm. The spheres were filled with varying ^{18}F concentrations to three different signal to background ratios

(SBRs), namely 2.59:1, 5.06:1 and 8.40:1. Data were acquired on a Philips GEMINI TF Time-of-Flight PET/CT scanner. The PET images were obtained with the 3D LOR-RAMLA reconstruction and CT-based attenuation correction. They present a spatial resolution of 2.0 mm (Fig. 1, first row).

For comparison, all images were registered to the highest contrast image. Reference segmentation was obtained by closest interactive thresholding to the known geometry on this image. The ROC curves of Fig. 2 measure the set differences between the reference segmentation and our shapings result. It illustrates the impact of the signal loss on the segmentation results. The obtained results are near perfect for the image with a very high SBR of 8.40:1. However, even for more realistic SBRs, including that of value 2.59:1, which is low with respect to real cases, the obtained results remain satisfactory, thus underlining the relevance of the shapings approach.

4.2. Real images

Finally, we qualitatively evaluated our method on a series of 13 PET images of lymphoma patients presenting significant FDG-PET tumor radioactivity uptake before radiotherapy. The acquisitions were carried out on several PET/CT scanners (Philips ALLEGRO and GEMINI DUAL GS, Siemens Biograph Sensation 16, GE DISCOVERY ST) with their associated reconstruction algorithms (RAMLA, TF MLEM, OSEM). They were acquired following standard protocol for PET cancer imaging: PET acquisition using a scan time of 3 minutes per bed position, between the base of the skull and mid-tigh, one hour after the peripheral intravenous injection of an ^{18}F-FDG dose of 4 to 6 MBq/kg in patients, fasted for at least 6 hours before FDG injection. PET image were obtained with CT-based attenuation correction and their spatial resolution (FWHM) varied from $4 \times 4 \times 2$ mm^3 to $4 \times 4 \times 4$ mm^3.

Our approach was applied to these images and compared to global thresholding. An example of result is illustrated in Fig. 3. We observe that the proposed method improves the robustness to signal heterogeneity in the image; this fact derives from the component-tree based approach, that intrinsically provides adaptive thresholding. The shapings paradigm also allows us to take into account geometric priors, improving in particular the robustness to texture effects, for instance here in the right lung. Furthermore, shapings share with the global thresholding the advantages of low parametricity, easy interaction and real-time computing.

5. CONCLUSION AND PERSPECTIVES

We have introduced the first application of morphological hierarchies for the segmentation of PET images. We have shown that the component-tree handles the principal approaches of intensity-based segmentation, subsuming the thresholding strategies considered until now. We have also proposed a solution based on the shapings paradigm to embed anatomical priors, while preserving a threshold-based strategy, that allows us to satisfy important properties such as real-time computing, and user-friendly interaction. For future work, we plan to use several attributes, to enrich the anatomical priors guiding the segmentation. Second, we also plan to use morphological (CT, MRI) data, in addition to PET images, as already proposed in recent works [7, 8]. These are indeed prerequisites to retrieve accurate anatomical information, both from geometrical and relational points of view. As a side effect, we will need to handle vectorial attributes (resp. multimodal imaging), and thus to define the attribute space of $\mathcal{T}$ (resp. I) in a partial order framework, which are not handled by component-trees. To tackle this issue, we will cross the con-
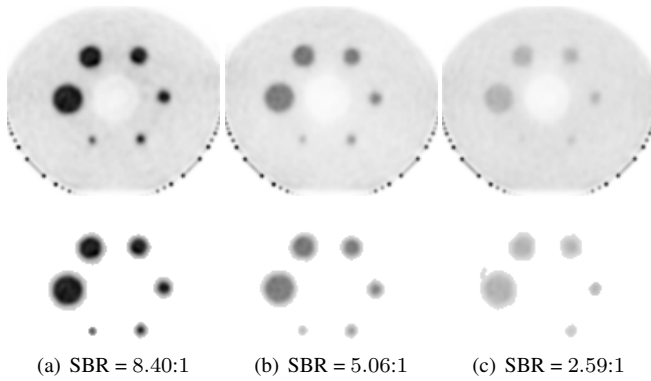
(a) SBR = 8.40:1 (b) SBR = 5.06:1 (c) SBR = 2.59:1

Fig. 1. First row: phantom PET images, with three different SBRs. Second row: segmentation results. MIP visualization. See Sec. 4.1.
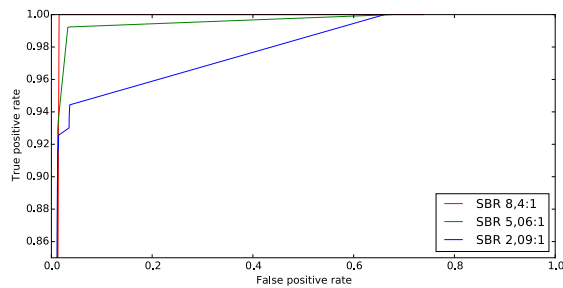

Fig. 2. ROC curves for the relative thresholding based on attribute-based approach, for various SBRs (see Fig. 1).
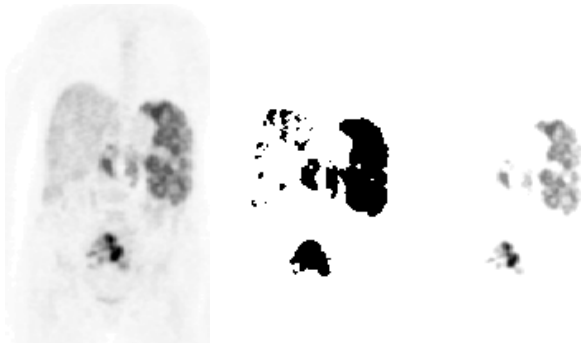

Fig. 3. Left: PET image. Middle: optimal global thresholding. Right: relative shaping segmentation. Maximum intensity projection. See Sec. 4.2.

cept of shapings with the recently introduced notion of component-graph [16], in order to develop the next generation of hierarchical approaches for PET image segmentation.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] S.-C. Huang, "Anatomy of SUV," *Nucl Med Biol*, vol. 27, pp. 643–646, 2000.

[2] U. Nestle, S. Kremp, and A. Schaefer-Schuler et al., "Comparison of different methods for delineation of ^{18}F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer," *J Nucl Med*, vol. 46, pp. 1342–1348, 2005.

[3] H. Li, W. L. Thorstad, and K. J. Biehl et al., "A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours," *Med Phys*, vol. 35, pp. 3711–3721, 2008.

[4] H. Zaidi, M. Diaz-Gomez, and A. Boudraa et al., "Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging," *Phys Med Biol*, vol. 47, pp. 1143–1160, 2002.

[5] M. Hatt, C. Cheze le Rest, and A. Turzo et al., "A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET," *IEEE T Med Imaging*, vol. 28, pp. 881–893, 2009.

[6] T. Shepherd, M. Teräs, and R. Beichel et al., "Comparative study with new accuracy metrics for target volume contouring in PET image guided radiation therapy," *IEEE T Med Imaging*, vol. 31, pp. 2006–2024, 2012.

[7] J. Wojak, E. D. Angelini, and I. Bloch, "Joint variational segmentation of CT-PET data for tumoral lesions," in *ISBI, Proc.*, 2009, pp. 217–220.

[8] U. Bagci, J. K. Udupa, and N. Mendhiratta et al., "Joint segmentation of anatomical and functional images: Applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images," *Med Image Anal*, vol. 17, pp. 929–945, 2013.

[9] L. Najman and H. Talbot, Eds., *Mathematical Morphology: From Theory to Applications*, ISTE/J. Wiley & Sons, 2010.

[10] P. Salembier, A. Oliveras, and L. Garrido, "Anti-extensive connected operators for image and sequence processing," *IEEE T Image Process*, vol. 7, pp. 555–570, 1998.

[11] A. Dufour, O. Tankyevych, and B. Naegel et al., "Filtering and segmentation of 3D angiographic data: Advances based on mathematical morphology," *Med Image Anal*, vol. 17, pp. 147–164, 2013.

[12] N. Passat, B. Naegel, and F. Rousseau et al., "Interactive segmentation based on component-trees," *Pattern Recogn*, vol. 44, pp. 2539–2554, 2011.

[13] Y. Xu, T. Géraud, and L. Najman, "Morphological filtering in shape spaces: Applications using tree-based image representations," in *ICPR, Proc.*, 2012, pp. 485–488.

[14] E. Carlinet and T. Géraud, "A comparative review of component tree computation algorithms," *IEEE T Image Process*, vol. 23, pp. 3885–3895, 2014.

[15] R. Jones, "Connected filtering and segmentation using component trees," *Comput Vis Image Und*, vol. 75, pp. 215–228, 1999.

[16] N. Passat and B. Naegel, "Component-trees and multivalued images: Structural properties," *J Math Imaging Vis*, vol. 49, pp. 37–50, 2014.