Hierarchical forest attributes for multimodal tumor segmentation on FDG-PET/contrast-enhanced CT

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

HIERARCHICAL FOREST ATTRIBUTES FOR MULTIMODAL TUMOR SEGMENTATION ON FDG-PET/CONTRAST-ENHANCED CT

Francisco Javier Alvarez Padilla[1], Barbara Romaniuk[1], Benoît Naegel[2], Stéphanie Servagi-Vernat[1,3], David Morland[1,4], Dimitri Papathanassiou[1,4] and Nicolas Passat[1]

[1] Université de Reims Champagne-Ardenne, CReSTIC, France
[2]Université de Strasbourg, CNRS, ICube, France
[3]Radiotherapy Department, Institut Jean-Godinot, Reims, France
[4]Nuclear Medicine Department, Institut Jean-Godinot, Reims, France

ABSTRACT

Accurate tumor volume delineation is a crucial step for disease assessment, treatment planning and monitoring of several kinds of cancers. However, this process is complex due to variations in tumors properties. Recently, some methods have been proposed for taking advantage of the spatial and spectral information carried by coupled modalities (e.g., PET-CT, MRI-PET). Simultaneously, the development of attribute-based approaches has contributed to improve PET image analysis. In this work, we aim at developing a coupled multimodal / attribute-based approach for image segmentation. Our proposal is to take advantage of hierarchical image models for determining relevant and specific attribute from each modality. These attributes then allow us to define a unique, semantic vectorial image. Sequentially, this image can be processed by a standard segmentation method, in our case a random-walker approach, for segmenting tumors based on their intrinsic attribute-based properties. Experimental results emphasize the relevance of computing region-based attributes from both modalities.

Index Terms— Image segmentation, multimodality, nuclear imaging, computed tomography, hierarchical models, region-based attributes

1. INTRODUCTION

Medical imaging has revolutionized the management of many diseases and allows nowadays the analysis in vivo of human body from anatomical structures to physiological processes. It plays a central part in cancer for diagnosis, treatment planning, following-up and / or the evaluation of treatment response. Despite its complexity, tumor evaluation can be characterized by four main factors: cellular activity, anatomical position, volume and extension (metastasis). In particular,

these parameters can be computed from PET/CT data. However, these imaging modalities present intrinsic restrictions. A low tumor-to-background contrast may be encountered in CT, meanwhile PET has a low spatial resolution that leads to partial volume effects. Nevertheless, the use of both anatomical and functional information allows to partially overcoming these drawbacks. In clinical routine, the delineation of tumors is performed in a mostly manual way to guide the outcome. This modus operandi is slow and prone to intra / inter operator errors / variability. To improve robustness and ergonomy, some methods have been introduced [1]. They can be classified with respect to their methodology: thresholding [2, 3], learning-based [4], boundary-based (level set [5], gradient-based [6]), statistical scopes (FLAB [7], fuzzy-c-means [8], Random Walker [9]), mathematical morphology (watershed [10], component-tree [11]) and hybridization [12, 13]. Most of these methods have been designed for processing PET images because of its high sensitivity and specificity to tumor biomarker metabolism. However, two inclusive notions are gaining share to enrich image information concept: multimodality and region-based features. Multimodality [14] refers to the use of semantically different images as data sources. On the other hand, region-based features can be computed [15] for image characterization. Different strategies have shown their performance to define relevant image regions for feature extraction such as using sliding windows [16], supervoxels [17, 18] or hierarchical image models (HS). Among these models, the component-tree [19] allows one to spectrally and spatially represent an image by a tree structure derived from the inclusion relation on the binary level-sets of the image. Alternatively, the tree of shapes [20] provides another HS whose regions represent the iso-contours. Both hierarchical data-structures are convenient to represent objects of interest that are connected and spectrally homogeneous. The characterization of these regions can be derived also from specific criteria. Indeed, region-based attributes can be computed for comparison / characterization purposes (area, shape, contrast, for instance). The fact of

merging regional-based attributes and multimodality scopes can improve target definition. In this work, we propose a way to exploit hierarchical multimodal structures for this purpose. Relevant attributes are computed from a HS of each image modality. They are analyzed by the Maximally Stable Extremal Regions algorithm [21] in order to build a vectorial image profile. This framework is used for random walk segmentation arguing a better performance than with scalar images.

2. FRAMEWORK

Our methodology is divided in four main steps: hierarchical structure computing, attribute computing, vectorial image restitution and vectorial segmentation (see Fig. 1).
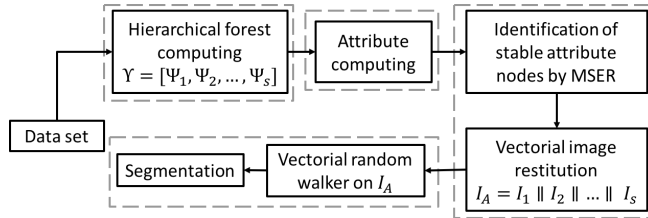

Fig. 1: Algorithm pipeline.

Hierarchical structure computing From a data set of n images, a s number of HS Ψ are built obtaining a set of trees, forming a hierarchical forest $\Upsilon = [\Psi_1, \Psi_2, .., \Psi_s]$ (see Sec. 2.1).

Attribute computing The nodes Θ of Ψ can be characterized by features (conventionally called attributes h) to increase regional comparison points [22] (see Sec. 2.2).

Vectorial image restitution The attributes of the extremal nodes of the trees Ψ are evaluated by the MSER algorithm (see Sec. 2.3) to find their stability area within the branches. A pixel / voxel $x \in X$ can belong to various nodes within Ψ. MSER is a crucial step to select in a branch, the node which better defines x for each attribute h. Each stable attribute value is then embedded into the initial image support for building a vectorial image. In particular, a new attribute image $\overrightarrow{i}$ is built on X. At each point, it is composed of a vectorial value $\overrightarrow{A} = [h_1, h_2, ..., h_n]$ obtaining a vectorial image $\overrightarrow{i}$. All vectorial images obtained from the different trees are then merged for creating a global vectorial image $\overrightarrow{I}$.

Vectorial segmentation The proposed vectorial image $\overrightarrow{I}$ can be processed by any segmentation method that can handle vectorial values. We opted for the Random Walker (RW) for its robustness to noise, as well as its ability to detect weak contours and simultaneously identify multiple objects.
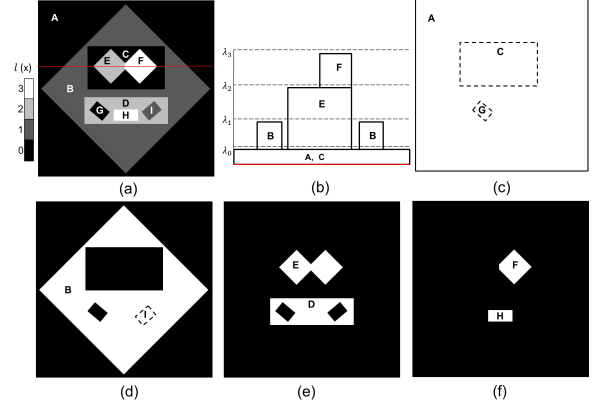

Fig. 2: (a) A gray-level image I. (b) 1D representation of I along the line $l(x)$ (in red, in (a)). (c–f) The four $\lambda_*(I)$. Each connected component is labeled by a letter (A, B, ...). See Fig. 3 for various tree computations of this image.

2.1. Hierarchical image structures

A hierarchical structure, or tree, is a data-structure that models –generally in a lossless fashion– the spatial and spectral information of an image, in a multiscale way. This partition leads to the definition of image regions or nodes. A node $k \in \Theta$ of such a tree is defined as a maximally connected set of points which present a homogeneous distribution of values in a threshold set of the image. These nodes are hierarchically organized via an order relation, e.g., the inclusion relation, or the nested relation of their frontiers.

Two classical hierarchical models are considered in this work: the component-tree (C-T) τ and the tree of shapes Φ.
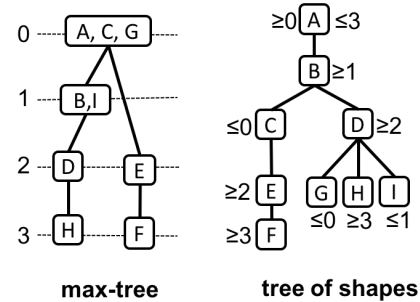
Fig. 3: An example of two HS built from Fig. 2 (a).

Component-tree For a C-T τ, an image I is partitioned by a thresholding protocol (considering the $\geq$ order for a max-tree, $\leq$ for a min-tree) at each intensity level, defining the binary levels $\lambda_*(I)$ (see Fig. 2). In other words, a component-tree can be seen as the collection of all thresholding segmented regions which are related by the spatial inclusion. This tree can be computed in quasi-linear time [23].

Tree of shapes The ToS of an image I is defined as a self-dual version of component-trees; basically, it consists of

merging the max- and min-trees. More precisely, the ToS Φ (see Fig. 3) aims at structuring the connected regions associated to the nodes by considering the inclusion relation on these regions where the "holes" have been filled.

2.2. Attributes

An attribute h is considered as a feature (generally expressed by a scalar value) which is representative of an image region. It can rely on different notions, e.g., spectral attributes: contrast (spectral distance between the minimum and the maximum intensity), average intensity; spatial attributes: area (number of pixels/voxels), barycenter coordinates; geometrical attributes: compactness and elongation; hybrid attributes: volume (area×intensity value), volumetric contrast (volume×contrast). A low difference between attributes of adjacent nodes can generally signify that the associated regions are strongly correlated. For instance, if the object of interest (OOI) has a high contrast, spectral attributes are appropriate; if the OOI is considered circular 2D (spherical 3D), barycenter coordinates, compactness and elongation parameters are relevant.

2.3. Maximally Stable Extremal Regions

The MSER is a stability metric algorithm that seeks the stable attribute nodes of a hierarchical structure. This analysis process begins from the extremal nodes k_{i*} (leaves of Ψ), up to the root, trough the branches. The stability state is achieved if and only if the estimator $s(k_i) = |k_{i+\Delta}(h) - k_{i-\Delta}(h)|$ has a local minimum at k_i where $\Delta \in \mathbb{N}$ denotes the evaluation range. If h is an incremental attribute (e.g., area, volume), $s(k_i)$ must be divided between $|k_i(h)|$ for normalization.

2.4. Vectorial Random Walker

The RW [24] is a graph-based algorithm inspired from electrical circuit analogy, and adapted for the segmentation of images. An image $\vec{I}$ can be modeled as a graph, where each voxel represents a node $v \in V$ with a vector of attributes $\vec{A}$. Nodes are spatially binded by edges e. This relationship between two neighbor nodes v_i and v_j is thus expressed by e_{ij} and weighted by a Gaussian function:

$$w_{ij} = exp(-\beta \|\vec{A}_i - \vec{A}_j\|^2), \qquad (1)$$

where $\vec{A}_i$ is the vector of attributes h in the node v_i and β is a weight parameter. This framework allows a Random Walker to evaluate the least costly way to reach a target, defined by seeds / markers defined on the graph. The output of the RW algorithm is expressed by a map of probabilities of belonging to each class; voxels are then assigned as members of the class with the highest probability value.

3. METHODOLOGICAL ASPECTS ON MEDICAL IMAGES

We now present the instanciation of the above framework for tumor segmentation from PET-CT data.

3.1. HS type selection

In functional images (FI), such as PET, tumor areas (according to metabolism of biomarkers) have a high intensity. In order to take advantage of this feature, we opt for the max-tree as HS. The root of τ is then the node k at the level line λ_0 (the support Ω of I) and the leaves of τ represent the locally maximal values.

In anatomical imaging (AI), such as CT or MRI, tumoral tissues are not necessarily local maxima (or local minima) because of their contrast absorption by vascularization is considered irregular. Therefore, the ToS is a potential HS to handle this constraint. Tumors are thus evaluated by isocontours, ignoring their irregular content, placing their regions in the terminal branches of the tree.

3.2. Attributes selection

This point is crucial given that input attributes have a great impact on segmentation result. Their selection can be different between imaging modalities according to their specifications. Spectral attributes are more relevant for FI than for AI, while geometrical and spatial attributes are of greater interest on AI than on FI. However, the fact of using similar attributes does not mean redundancy because each image and HS have its own properties. Therefore, we considered the following attributes, for AI: contrast, area, barycenter coordinates, compactness, elongation, and for FI: contrast, area, intensity.

3.3. MSER evaluation range

The scalar free parameter Δ defines the inferior and superior limits, through the branches, to determine stability. The $\pm\Delta$ range must be inside the tumor interval (lower band b_l and upper band b_u) at HS to avoid ambiguity. Thus, Δ limits can be measured according to image spectral resolution: $\Delta < (b_u - b_l)/2$. For PET, tumors have been defined between 40–50% of maximal intensity (I_{max}) depending on the kind of cancer [1], thus, $b_l = [0.4, 0.5] \times I_{max}$ and $b_u = I_{max}$. For CT, we set b_l and b_u in Hounsfield units depending of tumor type and according to contrast absorption rate.

3.4. RW parameters definition

There are two parameters to set before running the RW algorithm: class markers and weighting norm.
Class markers. Two classes are defined: tumor zone f and background b. Their voxel markers m are define on FI (projected to AI) by a threshold protocol: $m_b = I(x) < t_b \times$

$max(I)$ and $m_f > t_f \times max(I)$, where $t_b < t_f$, t represents a percentage and $max(I)$ the maximum intensity value of I.

Weighting norm. In the norm $\beta \|\vec{A}_i - \vec{A}_j\|^2$, we propose to use the Mahalanobis distance to normalize the data scale, since $\vec{A}$ is composed of heterogeneous scalar values:

$$d_m(v_i, v_j) = \sqrt{\left(\vec{A}_i - \vec{A}_j\right)^T \Sigma^{-1} \left(\vec{A}_i - \vec{A}_j\right)}, \qquad (2)$$

where Σ^{-1} corresponds to the inverse covariance matrix. The free parameter β is set at 1.

4. EXPERIMENTATION

Data collection Our data set consists of 7 ^{18}F-FDG PET-CT with iodinated contrast enhancement (CE) of 7 different patients (2 colon cancer, 1 head and neck cancer, 1 endometrial cancer, 1 ovarian cancer, 1 lymphoma and 1 cervix cancer). These acquisitions were carried out by using a Discovery 710 PET/CT scanner (General Electrics, Milwaukee, USA) at the Cancerology Institute Jean-Godinot in Reims, France.

Patients were asked to fast for at least 6 hours and received an intravenous dose of FDG (3 MBq/kg). Portal phase CT was performed first, 70 seconds after an intravenous dose of iodinated contrast agent (Optiject 350). CT was acquired in auto mA mode with adaptative statistical iterative reconstruction (ASIR), 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. 3D time of flight mode, 6 to 7 bed-steps of 2 minutes, covering the area from the base of the skull through the proximal femur. PET was reconstructed using an iterative algorithm (OSEM 24 subsets, 2 iterations) and post-filtered 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.

Protocol PET and CE–CT images were processed according to their acquisition parameters such as image position and voxel size. Images were considered registered by co-acquisition. Thereby, we obtained 7 couples of images that were spatially corresponding. A C-T τ was built from PET and a ToS Φ from CE-CT in order to compose the forest $\Upsilon = [\tau_{PET}, \Phi_{CE-CT}]$. The attributes mentioned in Sec. 3.2 were computed. For Random Walker markers (see Sec. 3.4), t_f was set at 0.25 and t_b at 0.75 for all the experiments.

Results and discussion Results were compared with the Random Walker method applied both on mono-modal images: PET and CE-CT; and multimodal images: PET/CE-CT. The Dice similarity coefficient (DSC) was selected to measure the spatial overlap (false positives and false negatives) between a ground-truth (provided by a radiotherapist) and the computed results (see Fig. 4 for an illustrative example). The DSC scores are summarized as mean and standard deviation values, in the next Table.

Table 1: Different results of Random Walker applied on mono/multimodal intensity values; and our vectorial image from hierarchical forest attributes (HFA).

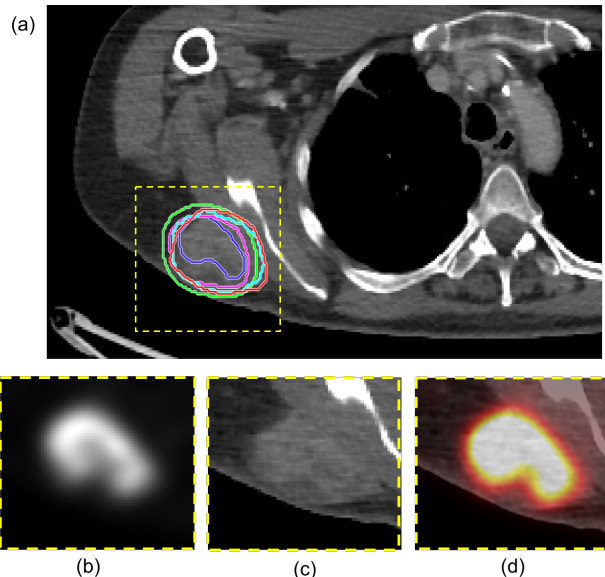| Monomodal	DSC ($\mu \pm \sigma$)	Multimodal	DSC ($\mu \pm \sigma$)
PET	0.716 ± 0.065	PET/CT	0.802 ± 0.038
CT	0.789 ± 0.053	HFA PET/CT	0.856 ± 0.036

Fig. 4: (a) CE–CT slice with boundary results: ground-truth in red, PET in blue, CE–CT in green, PET/CT in magenta, HFA in cyan. Multimodal image crops: (b) PET, (c) CE–CT and (d) PET/CE–CT fusion.

Results tend to show that multimodality has a better performance than imaging modalities individually processed, which corroborates results previously obtained in the literature. Indeed, their functional / morphological complementarity allows for a better delineation of tumors. We also observe that the use of region-based attributes representing stable image regions, seems to allow us to obtain slightly improved results. Indeed, the notion of attribute value stability within hierarchical structures allows us to better discriminate relevant properties from semantic noise. In addition, attributes characterize regions which, only by intensity, are considered similar. Finally, the versatility of relevant attributes and imaging modalities selection, allows to adapt the algorithm according to inherent target properties.

This will constitute our next work for consolidating the proposed framework. Some experiments will be proposed for a larger data-set of multimodal data. In addition, a ground-truth will be provided by several medical experts to consider inter-operator variability. Finally, a more detailed evaluation of the relevance of attributes (number, type) would be required in order to assess their respective implication in target discrimination process.

5. REFERENCES

[1] Foster B., Bagci U., Mansoor A., Xu Z. and Mollura D. *A review on segmentation of positron emission tomography images.* Comput Biol Med, 50:76–96, 2014.

[2] Daisne J. F., Sibomana M., Bol A., Doumont T., Lonneux M., and Grégoire V. *Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms.* Radiother Oncol, 69:247–250, 2003.

[3] Nestle U., Schaefer-Schuler A., Kremp S., Groeschel A., Hellwig D., Rübe C. and Kirsch C.-M. *Target volume definition for 18F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer.* Eur J Nucl Med Mol I, 34:453–462, 2007.

[4] Bi L., Kim J., Feng D. and Fulham M. *Multi-stage thresholded region classification for whole-body PET-CT lymphoma studies.* MICCAI, 569–576, 2014.

[5] Cheng-Liao J. and Qi J. *Segmentation of mouse dynamic PET images using a multiphase level set method.* Phys Med Biol, 55:6549–6569, 2010.

[6] Wanet M., Lee A., Weynand B., De Bast M., Pncelet A., Lacroix V., Coche E., Gégorie V. and Geets X. *Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer: A comparison with threshold-based approaches, CT and surgical specimens.* Radiother Oncol, 98:117–125, 2011.

[7] Hatt M., Cheze le Rest C., Descourt P., Dekker A., De Ruysscher D., Oellers M., Lambin P., Pradier O. and Visvikis D. *Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications.* Int J Radiat Oncol, 77:301–308, 2010.

[8] Belhassen S. and Zaidi H. *A novel fuzzy C-means algorithm for unsupervised heterogeneous tumor quantification in PET.* Med Phys, 37:1309–1324, 2010.

[9] Bagci U., Udupa J., Mendhiratta N., Foster B., Xu Z., Yao J., Chen X. and Mollura D. *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, 17:929–945, 2013.

[10] Tylski P., Bonniaud G., Decencière E., Stawiaski J., Coulot J., Lefkopoulos D. and Ricard M. ^{18}F-FDG PET images segmentation using morphological watershed: A phantom study. NSS/MIC, 4:2063–2067, 2006.

[11] Grossiord É., Talbot H., Passat N., Meignan M., Tervé P. and Najman L. *Hierarchies and shape-space for PET image segmentation.* ISBI, 1118–1121, 2015.

[12] Dewalle-Vignion A.-S., Betrouni N., Lopes R., Huglo D., Stute S. and Vermandel M. *A new method for volume segmentation of PET images, based on possibility theory.* IEEE T Med Imaging, 30:409–423, 2011.

[13] Grossiord E., Talbot H., Passat N., Meignan M. and Najman L. *Automated 3D lymphoma lesion segmentation from PET/CT characteristics.* ISBI, 174–178, 2107.

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

[15] Lartizien C., Rogez M., Susser A., Giammarile F., Niaf E. and Ricard F. *Computer aided staging of lymphoma patients with FDG PET/CT imaging based on textural information.* ISBI, 118–121, 2012.

[16] Ahmadvand P., Duggan N., Bénard F. and Hamarneh G. *Tumor lesion segmentation from 3D PET using a machine learning driven active surface.* MLMI Workshop, 271–278, 2016.

[17] Machairas V., Baldeweck T., Walter T. and Decencière E. *Hierarchical multi-scale supervoxel matching using random forest for automatic semi-dense abdominal image registration.* ISBI, 1409–1413, 2017.

[18] Conze P.-H., Noblet V., Rousseau F., Heitz F., Blasi V., Memeo R. and Pessaux P. *Scale-adaptive supervoxel-based random forests for liver tumor segmentation in dynamic contrast-enhanced CT scans.* Int J Comput Ass Rad, 12:223–233, 2017.

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

[20] Monasse P. and Guichard F. *Fast computation of a contrast invariant image representation.* IEEE T Image Process, 9:860–872, 2000.

[21] Matas J., Chum O., Urban M. and Pajdla T. *Robust wide baseline stereo from maximally stable extremal regions.* BMVC, 384–396, 2002.

[22] Alvarez F., Grossiord É., Romaniuk B., Naegel B., Kurtz C., Talbot H., Najman L., Guillemot R., Papathanassiou D. and Passat N. *Multicriteria 3D PET image segmentation.* IPTA, 346–351, 2015.

[23] Carlinet E. and Géraud T. *Comparative review of component-tree computation algorithms.* IEEE T Image Process, 23:3885–3895, 2014.

[24] Grady L. *Random walks for image segmentation.* IEEE T Pattern Anal, 28:1768–1783, 2006.