JOINT VARIATIONAL SEGMENTATION OF CT-PET DATA FOR TUMORAL LESIONS

Julien Wojak, Elsa D. Angelini, Isabelle Bloch

Institut Telecom, Télécom ParisTech, CNRS LTCI, 75013 Paris, France

ABSTRACT

Medical imaging, used for both diagnosis and therapy planning, is evolving towards multi-modality acquisition protocols. Manual segmentation of 3D images is a tedious task and prone to inter- and inter-experts variability. Moreover, the automatic segmentation exploiting the characteristics of multi-modal images is still a difficult problem. In this paper, we propose the use of a variational segmentation method, based on the minimization of the TV norm and a convex formulation, for segmenting thoracic pairs of PET and CT images, in the context of radiotherapy planning. We first highlight the limitations of a pure vectorial formulation of the variational segmentation method for PET and CT images. We then propose to better exploit the bi-modality by introducing a parameter which varies spatially depending on the PET intensity to adjust precisely the segmentation of CT images. Segmentation results on lung tumors and lymphatic nodes are shown, and comparisons performed with manual segmentations illustrate the quality of the results.

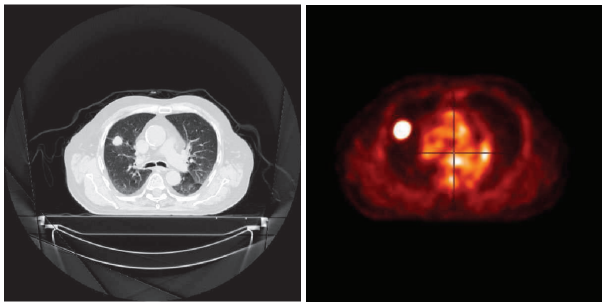
***Index Terms*—** Multi-modal segmentation, variational method, thoracic imaging, PET, CT

1. INTRODUCTION

In clinical practice, clinicians are used to work with multiple imaging modalities to perform a diagnosis on a tumoral pathology. Multimodal imaging is also very common for MRI brain imaging combining T1 and T2-weighted protocols. In this paper we focus on PET/CT image volumes for thoracic tumor or lymph node diseases. The underlying clinical goal is to provide an accurate and reproducible segmentation tool enabling the definition of tumoral target volumes in radiotherapy planning. The CT and PET imaging modalities provide complementary anatomical (CT) and functional (PET) information. In the past, the major shortcoming to joint quantitative analysis of these two modalities was the need for an accurate registration. Nowadays, combined PET-CT scans provide co-registered images. Fusing the information from the two modalities remains difficult, due to the coarser spatial resolution of the PET data, and breathing motion artifacts due to long PET acquisition times. Therefore, while there is an important amount of previous work dealing with the segmentation of either CT or PET images, very few works have addressed the joint segmentation problem. Among the few papers dealing with this problem, we can cite [1] where multimodal image segmentation was based on the active contours of Chan and Vese [2] formulated for multi-channel images. Promising results were obtained but the method was very sensitive to the initialization setup. In [3] after detecting tumors as local maxima in the PET image, the segmentation was performed with a joint likelihood ratio test. In [4] a classification framework based on MAP-MRF models was applied on PET-CT vectorial images.

Acknowledgments: This work was partially funded by the Medicen Pôle de Compétitivité within the Miniara project.

In all these previous works, perfect registration of the two modalities was assumed, which is not always a realistic hypothesis, especially in thoracic images where breathing movements are difficult to control. In this paper we propose to perform a CT image segmentation, guided by the PET information using a recently proposed variational formulation with a global minimum. The proposed segmentation framework controls the regularity of the segmented structures while easily incorporating PET information but remaining robust to potential mis-registration of the image data.



(a) (b)

Fig. 1. Clinical case with a lung tumor: (a) CT modality and (b) PET modality.

2. SEGMENTATION VIA TV SEMI-NORM MINIMIZATION AND FUZZY REGION COMPETITION

Variational image segmentation methods consist in finding contours or regions in a given image I by minimizing an appropriate energy functional. For example in the piecewise-constant Mumford-Shah image segmentation model from Chan-Vese [2], the functional is expressed as:

$$E_{CV}(C, c_1, ..., c_l) = \text{length}(C) + \lambda \sum_{l=1}^{L} \int_{\Omega_l} (I(x) - c_l)^2 \mathrm{d}\Omega \qquad (1)$$

where Ω_l, $l = 1..L$, is a partition of the image I domain Ω ($\Omega \subset \mathbb{R}^n$, $n = 2$ or 3) in L regions Ω_l, C being the set of curves that delimit the regions, and c_l are constant values characterizing the average value within the regions.

The solution is a partition of the image in L phases, each phase having a constant intensity (the optimal c_l value) and the length of the contours being constrained to avoid oscillations.

A common approach to minimize Equation (1) with two phases (foreground and background) is detailed in [2], embedding the contour formulation into a level-set framework and computing the Euler-Lagrange equations for the optimization of the energy parameters via gradient descent. The major drawback of this method is its sensitivity to local minima, making the quality of image segmentation sensitive to the initialization.

2.1. Fuzzy segmentation model

In [5, 6, 7] an alternative segmentation framework is proposed, directly looking for optimal Ω_i as an optimal function measuring the membership to each region (and no longer using the characteristic function). Let $u \in BV_{[0,1]}(\Omega)$ (the space of functions of bounded variations taking their values in $[0, 1]$) be a fuzzy membership function suited for the segmentation problem, e.g. $u \approx 0$ in the background and $u \approx 1$ in the foreground. The following functional is proposed (using similar notations as in Equation (1)) :

$$\min_{u\in BV_{[0,1]}(\Omega),c1,c2} E_{TV_g}(u, c1, c2) = \min_{u\in BV_{[0,1]}(\Omega),c_1,c_2} \int_\Omega g|\nabla u|\mathrm{d}\Omega + \lambda \int_\Omega u(I - c_1)^2\mathrm{d}\Omega + \lambda \int_\Omega (1 - u)(I - c_2)^2\mathrm{d}\Omega \quad (2)$$

The regularity of the segmented regions is controlled by constraining the TV semi-norm of the membership function. As in [6], the total variation (TV) semi-norm is weighted by a contour map g (for example $g = \frac{1}{1+|\nabla I|}$) to allow less regularization where contours are clearly visible.

The parameter λ controls the trade-off between data fidelity and regularity terms. This image segmentation formulation has two very important properties: (i) For fixed c_i, the problem (2) is convex in u, and therefore has a global minimum. There exist very efficient algorithms to minimize (2) with respect to u. (ii) The global minimizer of u is guaranteed to be close to a characteristic function. A contour can therefore be simply obtained by thresholding u at any level α with $0 < \alpha < 1$. Proofs of these properties are detailed in [5, 6, 8]. These properties have guided our choice of this type of method.

2.2. Efficient minimization

To solve Equation (2), an alternate minimization is iterated, first fixing u and optimizing c_i and then fixing c_i and optimizing u, until convergence. Minimization with respect to c_i is performed with these explicit formulations:

$$c_1 = \frac{\int_\Omega uI\mathrm{d}\Omega}{\int_\Omega u\mathrm{d}\Omega} \text{ and } c_2 = \frac{\int_\Omega (1-u)I\mathrm{d}\Omega}{\int_\Omega (1-u)\mathrm{d}\Omega} \quad (3)$$

When the c_i are fixed, the minimizer u^* of Equation (2) is the same as the minimizer of:

$$\min_{u\in BV_{[0,1]}(\Omega)} E_{TV_g}(u) = \int_\Omega g|\nabla u|\mathrm{d}\Omega + \lambda \int_\Omega ur\mathrm{d}\Omega \quad (4)$$

with $r = (I - c_1)^2 - (I - c_2)^2$.
As in [6], other region competition functions r could be used, such as statistical comparison of estimated region densities.
To minimize Equation (4) with respect to u under the constraint $u \in BV_{[0,1]}(\Omega)$, the following equivalent unconstrained problem can be written (the proof is given in [8]):

$$\min_u E_{TV_g}(u) = \int_\Omega g|\nabla u|\mathrm{d}\Omega + \int_\Omega \lambda ur + \beta\nu(u)\mathrm{d}\Omega \quad (5)$$

where the penalty term ν is given by $\nu(t) = \max(0, |2t - 1| - 1)$ and $\beta > \frac{1}{2}|r|_\infty$. The problem (5) could be solved with Euler-Lagrange equations and gradient descent. However this method is slow and since the terms $|\nabla u|$ and $\nu(t)$ are not differentiable, they must be regularized, which implies a slower gradient descent. To avoid these two problems, as in [6], it is preferable to use the fast duality projection algorithm proposed in [9]. This algorithm is based on a weak approximation of Equation (5) by introducing an auxiliary variable v:

$$\min E_{TV_{g2}}(u, v) = \min \int_\Omega g(x)|\nabla u|\mathrm{d}\Omega + \frac{1}{2\theta}\int_\Omega |u - v|^2\mathrm{d}\Omega + \int_\Omega \lambda vr + \beta\nu(v)\mathrm{d}\Omega \quad (6)$$

This approximation is suitable if the minimizers u^* and v^* are almost identical w.r.t. the L^2 norm. To ensure this condition, θ must be set to a small value. Equation (6) is minimized with an alternate scheme described in Algorithm 1. Thanks to this formulation, since the functional 2 is convex, we have a robust method to segment an image in two constant piecewise phases. Moreover, thanks to the algorithm 1 the minimization is very fast.

Algorithm 1 Minimization

Require: $I, u, \theta, \tau, g, \lambda$
  repeat
    compute c_1 and c_2 by (3)
    $r := \lambda[(I - c_1)^2 - (I - c_2)^2]$
    $v := \max(\min(u - \theta r, 1), 0)$
    $p^0 := 0$
    repeat
      $p^{n+1} := \frac{p^n + \frac{1}{8}\nabla(div(p^n) - \frac{v}{\theta})}{1 + \frac{\frac{1}{8}\nabla(div(p^n) - \frac{v}{\theta})|}{g}}$
    until $p^{n+1} \approx p^n$
    $u := v - div(p)$
  until convergence
  return u

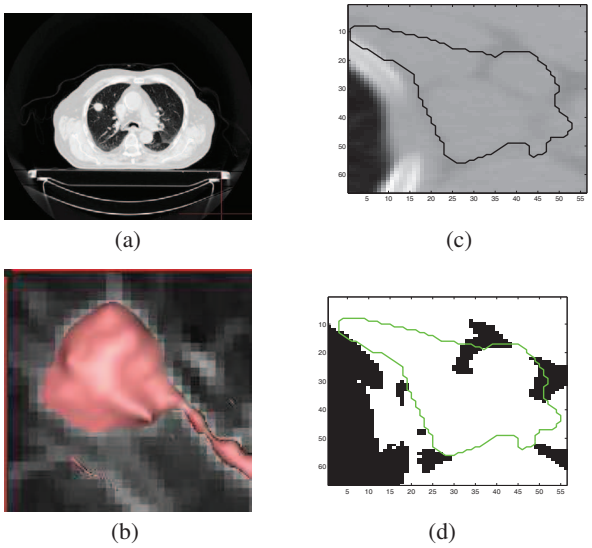


Fig. 2. Left: (a) Lung CT data and (b) automated tumor segmentation with prior masking of non-lung tissues. Right: (c) Thoracic lymph node CT data and (d) comparison of manual segmentation (green contour) and automated segmentation (binary mask).

As illustrated in Figure 2, this segmentation framework provides very accurate delineation of highly-contrasted stuctures on CT images but is not able to extract low-contrast structures such as lymph nodes or to well differentiate the tumor from adjacent blood vessels. PET information mainly consists of increased intensity values in active parts of tumors or pathological lymph nodes. However, due to breathing movements and acquisition artifacts, these intense areas are often larger than the corresponding anatomical structures. For this reason, we would like to incorporate PET information in our segmentation process, while remaining robust to uncertainty and mis-registration of the PET image data content.

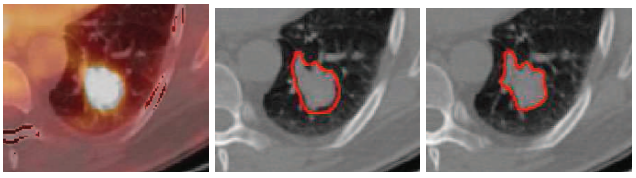
3. PET-BASED LOCAL CONSTRAINTS ON THE FUZZY REGION COMPETITION

3.1. Direct vectorial approach

For lung tumors it is often difficult to visually separate the tumor from the surrounding vessels or mediastinal tissues in CT image data. In this case, PET images can provide useful information to guide the segmentation on the CT images, highlighting only tumoral tissues.
Since the CT spatial resolution is much higher than the PET resolution, and in order to better preserve the 3D continuity of anatomical structures, we resampled the PET images via bilinear interpolation to match the CT spatial resolution in terms of voxel sizes.
After resampling, a straightforward extension of the proposed segmentation framework can be tested, reformulating the energy functional to handle vectorial images, and leading to the following energy minimization problem:

$$\begin{aligned}&\min_{u\in BV_{[0,1]}(\Omega)} E_{TV_g}(u,\mathbf{c_1},\mathbf{c_2}) = \\ &\min_{u\in BV_{[0,1]}(\Omega)} \int_\Omega g|\nabla u|\mathrm{d}\Omega \\ &\qquad + \int_\Omega u||\Lambda(\mathbf{I}-\mathbf{c_1})||^2\mathrm{d}\Omega + \int_\Omega (1-u)||\Lambda(\mathbf{I}-\mathbf{c_2})||^2\mathrm{d}\Omega\end{aligned} \tag{7}$$

where $\mathbf{c_1} = (c_{1_1}, c_{1_2})^T, \mathbf{c_2} = (c_{2_1}, c_{2_2})^T, \mathbf{I} = (I_{CT}, I_{PET})^T$ are vectors in $\mathbb{R}^2$ and $\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. While this method is faster than [1], it suffers from the same sensitivity to the intrinsic misregistration between the CT and the PET data. This is illustrated in Figure 3 (b) where the vectorial segmentation result for a lung tumor corresponds to the mean shape between the PET and the CT image information.



(a) (b) (c)

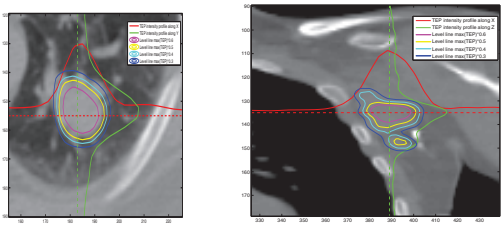
Fig. 3. Averaging effect: (a) superimposed PET and CT images, (b) vectorial segmentation with fixed parameters, (c) segmentation with a spatially varying $\lambda(x)$.

3.2. Locally adaptative use of PET information

The PET information can be introduced in the segmentation method by favoring the CT-based segmentation of tumors or lymph nodes in bright PET areas. We propose to implement this idea by using a spatially varying parameter λ.
We therefore reformulate the energy functional in Equation 2 as :

$$\begin{aligned}&\min_{u\in BV_{[0,1]}(\Omega)} E_{TV_g}(u,c1,c2) = \int_\Omega g|\nabla u|\mathrm{d}\Omega \\ &\qquad + \int_\Omega \lambda(x)u(I-c_1)^2\mathrm{d}\Omega + \int_\Omega \lambda(x)(1-u)(I-c_2)^2\mathrm{d}\Omega\end{aligned} \tag{8}$$

The function $\lambda(x)$ must be designed so that it takes high values where we want the data fidelity term to be more important than the regularization term and the segmentation result to closely follow the CT information. Conversely when $\lambda(x)$ takes low values, the data fidelity term is less important than the regularization term and the oscillations of the object (e.g. tumor) contours are minimized. The minimization process remains the same as the one described in Section 2.2 since $\lambda(x)$ is not optimized but pre-defined.
For radiotherapy planning, clinicians use the PET data to segment the tumors or the lymph nodes. In [10] several lung tumors segmentation methods for PET images were discussed and evaluated. Most of the methods are based on simple image intensity thresholding. The most simple threshold values are set between 60% and 50% of the maximal uptake PET intensity. In [11] a threshold was computed from the mean PET intensity inside the whole body, based on a specific regression formula. In [12] a threshold at 60% of the maximal uptake intensity was criticized and a thresholding taking into account background PET intensity was proposed. The study in [10] revealed that thresholding at 60% leads to too small delineations of the tumors. We therefore considered that all voxels with intensity larger than 60% were inside the pathological structure (tumor or lymph node) to design our function $\lambda(x)$. Moreover, the PET intensity signal is intense inside the pathological structure and decreases around it, as illustrated in Figure 4 on intensity profiles accross a tumor and a lymph node. From these observations, we assessed that $\lambda(x)$ should be an increasing function of the PET intensity with the following properties:
$-$ $\lambda(x)$ should take a constant high value M for x greater than 60% of the PET maximum intensity value;
$-$ $\lambda(x)$ should take a low value m close to 0 for low PET values;
$-$ $\lambda(x)$ should be an increasing function between m and M. The growth rate we propose depends on the confidence we have in the intermediate PET values.
A sigmoid function is proposed to define $\lambda(x) = M\frac{1}{1+ae^{-bI(x)}} + m$ (where $I(x)$ represents the PET intensity centered in 0). The parameter a is set to 1, M which is the upper bound of the sigmoid is chosen such that the data fidelity term and the regularization term are comparable, m is the lower bound of the sigmoid and is set to 10^{-6} to have a vanishing data fidelity term when the PET signal is not intense, and b which controls the slope of the sigmoid transition is chosen in the empirical interval $[0.0055, 0.012]$ (with $b \in [0.0055, 0.012]$ the similarity index between a manual segmentation and an automated segmentation was always superior to 0.5).



(a) (b)

Fig. 4. PET intensity profiles overlaid on CT data for: (a) a lung tumor and (b) a lymph node. Level lines of PET intensity are superimposed for values between 0.6×max(PET) in red and 0.3×max(PET) in blue.

4. CLINICAL EXPERIMENTS & RESULTS

We evaluated the proposed segmentation framework on one case with lymphoma (presenting infected enlarged lymph nodes) and five cases with lung cancer (presenting tumors). Some visual results are

illustrated in Figures 3(c) and 6. To quantify the quality of the segmentation results, we computed the similarity, the sensitivity and the specificity indices comparing to manual segmentation from a clinical expert. They are reported in Table 1. All results were satisfactory for all cases, with higher accuracy on lung tumors.

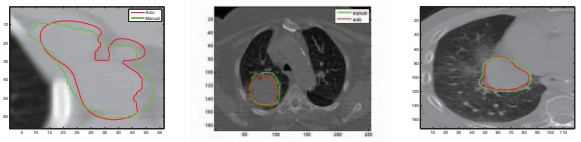
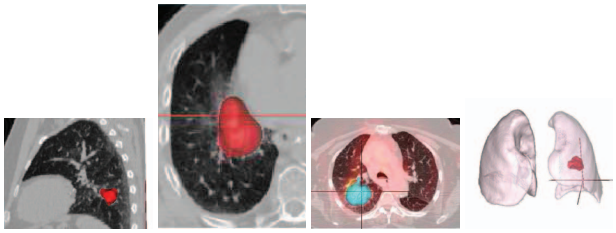


Fig. 5. Segmentation results showing the manual segmentation (green) and the automatic segmentation (red) on three pathological cases: (left) thoracic lymph node, (middle-right) two large lung tumors in contact with the lung parenchyma.



(a) (b) (c) (d)

Fig. 6. 3D views of segmentation results. (a-b) A tumor superimposed on orthogonal slices of a CT scan, (c) the tumor superimposed on the fused PET-CT scan images, (d) 3D view of the tumor along with the lungs masks.

	Similarity index	Sensitivity	Specificity
tumor 1	0.85	0.76	0.96
tumor 2	0.79	0.68	0.96
tumor 3	0.81	0.78	0.89
tumor 4	0.90	0.86	0.95
tumor 5	0.82	0.79	0.84
lymph node	0.72	0.60	0.92

Table 1. Quantitative results: comparaison of manual and automatic segmentations.

We also performed tests on images for which manual segmentations were not available. Overall qualitative results were very satisfactory in terms of robustness of the segmentation framework to low-contrast contours and mis-registration between the two imaging modalities. However in difficult cases where the tumor is at the bottom of the lungs, close to the liver, some difficulties remain. Indeed PET intensity is also very high in the liver and does no allow us to discriminate the tumor in the lung. This case could be solved by pre-segmenting the liver using for example the method described in [13].

5. CONCLUSION

We have proposed in this work an original extension of a two-phase fuzzy variational segmentation framework to handle multi-modal 3D medical images. Our method takes into account the specificity of the imaging modalities as it allows to manage uncertainties on PET image information to guide a precise segmentation of a co-registered CT data set, via the manipulation of spatially variable parameters in the energy functional being minimized. Preliminary clinical results have shown the robust behavior of the proposed formulation to challenging pathological cases with low-contrast CT image information.

6. REFERENCES

[1] I. El Naqa, D. Yang, and J. O. Deasy, "Automated estimation of the biophysical target for radiotherapy treatment planning using multimodality image analysis," in *International Conference on Image Processing*, 2007, pp. V: 533–536.

[2] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[3] V. Potesil, X. Huang, and X. Zhou, "Automated tumor delineation using joint PET/CT information," in *Proc. of SPIE, Medical Imaging: Computer-aided Diagnosis*, 2007, vol. 6514.

[4] H. Grubben, P. Miller, G. G. Hanna, K. J. Carson, and A. R. Hounsell, "MAP-MRF segmentation of lung tumours in PET/CT image," in *International Symposium on Biomedical Imaging*, 2009, pp. 290–293.

[5] X. Bresson, S. Esedoglu, P. Vandergheynst, J. P. Thiran, and S. J. Osher, "Fast global minimization of the active contour/snake model," *Journal of Mathematical Imaging and Vision*, vol. 28, no. 2, pp. 151–167, June 2007.

[6] B. Mory and R. Ardon, "Fuzzy region competition: A convex two-phase segmentation framework," in *Scale Space and Variational Methods in Computer Vision*, 2007, pp. 214–226.

[7] B. Mory, R. Ardon, and J. P. Thiran, "Variational segmentation using fuzzy region competition and local non-parametric probability density functions," in *International Conference on Computer Vision*, 2007, pp. 1–8.

[8] T. Chan, S. Esedoglu, and M. Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," Tech. Rep., UCLA CAM Report 04-54, 2004.

[9] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 89–97, 2004.

[10] Y. Venel, H. Garhi, A. de Muret, J.-L. Baulieu, I. Barillot, and C. Prunier-Aesch, "Comparison of six methods of segmentation of tumor volume on the 18F-FDG PETscan with reference histological volume in nonsmall cell bronchopulmonary cancers," *Médecine Nuclaire*, vol. 32, no. 6, pp. 339 – 353, 2008.

[11] Q. C. Black, I. S. Grills, L. L. Kestin, C. O. Wong, J. W. Wong, A. A. Martinez, and D. Yan, "Defining a radiotherapy target with positron emission tomography," *International Journal of Radiation Oncology Biology Physics*, vol. 60, no. 4, pp. 1272 – 1282, 2004.

[12] U. Nestle, S. Kremp, A. Schaefer-Schuler, C. Sebastian-Welsch, D. Hellwig, C. Rubes, and C-M Kirsch, "Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer," *The Journal of nuclear medicine*, vol. 46, no. 8, pp. 1342 – 1348, 2005.

[13] T. Heimann and B. van Ginneken etal, "Comparison and evaluation of methods for liver segmentation from CT datasets," *IEEE Trans. Medical Imaging*, vol. 28, no. 8, pp. 1251–1265, Aug. 2009.