Generic and robust method for automatic segmentation of PET images using an active contour model

Mingzan Zhuang
Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands; Department of Radiation Oncology, Tumor Hospital of Shantou University Medical College, Shantou, Guangdong 515000, China; and The Key Laboratory of Digital Signal and Image Processing of Guangdong Province, Shantou University, Shantou, Guangdong 515000, China

Rudi A. J. O. Dierckx
Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands

Habib Zaidi[a)]
Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211 Geneva, Switzerland; Geneva Neuroscience Center, Geneva University, CH-1205 Geneva, Switzerland; and Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands

Purpose: Although positron emission tomography (PET) images have shown potential to improve the accuracy of targeting in radiation therapy planning and assessment of response to treatment, the boundaries of tumors are not easily distinguishable from surrounding normal tissue owing to the low spatial resolution and inherent noisy characteristics of PET images. The objective of this study is to develop a generic and robust method for automatic delineation of tumor volumes using an active contour model and to evaluate its performance using phantom and clinical studies.

Methods: MASAC, a method for automatic segmentation using an active contour model, incorporates the histogram fuzzy C-means clustering, and localized and textural information to constrain the active contour to detect boundaries in an accurate and robust manner. Moreover, the lattice Boltzmann method is used as an alternative approach for solving the level set equation to make it faster and suitable for parallel programming. Twenty simulated phantom studies and 16 clinical studies, including six cases of pharyngolaryngeal squamous cell carcinoma and ten cases of nonsmall cell lung cancer, were included to evaluate its performance. Besides, the proposed method was also compared with the contourlet-based active contour algorithm (CAC) and Schaefer's thresholding method (ST). The relative volume error (RE), Dice similarity coefficient (DSC), and classification error (CE) metrics were used to analyze the results quantitatively.

Results: For the simulated phantom studies (PSs), MASAC and CAC provide similar segmentations of the different lesions, while ST fails to achieve reliable results. For the clinical datasets (2 cases with connected high-uptake regions excluded) (CSs), CAC provides for the lowest mean RE ($-8.38\% \pm 27.49\%$), while MASAC achieves the best mean DSC (0.71 ± 0.09) and mean CE ($53.92\% \pm 12.65\%$), respectively. MASAC could reliably quantify different types of lesions assessed in this work with good accuracy, resulting in a mean RE of $-13.35\% \pm 11.87\%$ and $-11.15\% \pm 23.66\%$, a mean DSC of 0.89 ± 0.05 and 0.71 ± 0.09, and a mean CE of $19.19\% \pm 7.89\%$ and $53.92\% \pm 12.65\%$, for PSs and CSs, respectively.

Conclusions: The authors' results demonstrate that the developed novel PET segmentation algorithm is applicable to various types of lesions in the authors' study and is capable of producing accurate and consistent target volume delineations, potentially resulting in reduced intraobserver and interobserver variabilities observed when using manual delineation and improved accuracy in treatment planning and outcome evaluation. © *2016 American Association of Physicists in Medicine.* [http://dx.doi.org/10.1118/1.4954844]

1. INTRODUCTION

Hybrid molecular positron emission tomography/computerized tomography (PET/CT) imaging-guided radiation therapy makes use of advanced imaging technology for improved definition of tumor target volumes, thus relating the absorbed dose information to image-based patient representations. The information provided by PET images regarding metabolic activity and the spatial location of lesions is helpful in delineating the target tumor required

for assessment of response to treatment or radiation therapy treatment planning. However, one of the challenges of PET/CT imaging-guided target volume delineation is the lack of accurate standardized algorithms enabling the segmentation of target volumes, which is caused by the low spatial resolution and intrinsic noisy properties of PET images.

A number of PET segmentation algorithms have been developed during the last 15 years. These algorithms can be roughly divided into four categories: thresholding techniques, variational approaches, learning methods, and stochastic modeling-based methods. At the present time, there is no consensus on the best performing algorithm that can be adopted as standard for all or a number of indications.[1–3] Previous studies have shown a large variability in terms of computational complexity and amount of user interaction required by the various image segmentation techniques.[4–6] Recent reviews of state-of-the-art PET image segmentation techniques indicate that there is no optimal solution for all types of clinical oncology indications with respect to accuracy, precision, and efficiency.[7–9] As such, PET image segmentation remains an active and open research topic, leading to an ambient tendency toward yapetism (yet another PET image segmentation method).[9]

In this work, we focus on combining existing algorithms into a novel region-based method for automatic segmentation using an active contour model (MASAC) to enhance robustness and generalness rather than proposing yet another PET segmentation method. This algorithm incorporates both histogram fuzzy C-means (FCM) clustering and textural information to constrain the active contour to detect the boundaries in an accurate and robust way, without requiring the setting of user-defined parameters. Moreover, we make the algorithm more efficient and appropriate for parallel programming by taking the lattice Boltzmann method (LBM) as an alternative to solve the level set equation. The proposed model is quite general and can be applied in many clinical scenarios.

2. MATERIALS AND METHODS

2.A. PET image segmentation algorithm

The proposed algorithm can be divided into three main steps: firstly, we apply histogram FCM clustering, bilateral filter, and Gabor transformation to PET images, respectively. All three transformed images were used to replace the original PET images serving as input to the energy functional to overcome the problems caused by noisy characteristics and enhance reliability and robustness. Then, Otsu's thresholding method,[10] which selects from the gray level histogram the threshold minimizing the intraclass variance of the pixels, was used for the initial contour in the active contour model. Finally, the evolution equation of LBM is used to solve the level set equation. The flowchart of MASAC is shown in Fig. 1.

2.A.1. Histogram fuzzy C-means clustering

FCM is one of the most widely used algorithms for image segmentation[11,12] and its success is mainly attributable to the
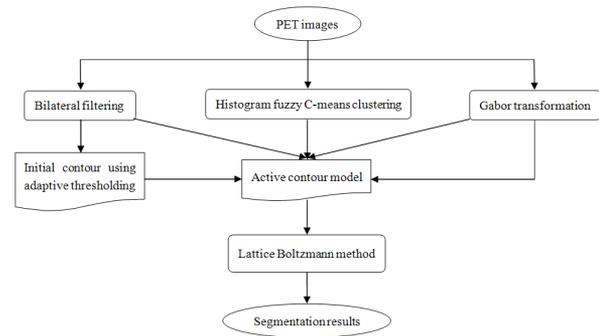

Fig. 1. Flowchart illustrating the various steps involved in the implementation of the MASAC PET image segmentation algorithm.

introduction of fuzziness for the membership of image voxels. To enhance efficiency, we perform clustering on the gray level histogram instead of image pixels following the methods proposed by Szilagyi *et al.*[13] and Zhao *et al.*[14] The gray level histogram of the images was constructed and smoothed using an adaptive kernel density estimator based on linear diffusion processes.[15,16] Thereafter, we perform clustering on the gray level histogram Hist to obtain the fuzzy membership images. The objective function used for histogram fuzzy C-means clustering is defined as

$$J = \sum_{k=1}^{c} \sum_{j=1}^{q} \text{Hist}(j) \mu_{kj}^{m} (j - \nu_k)^2, \tag{1}$$

where ν_k represents the center of the kth cluster, μ_{kj} denotes the fuzzy membership value of the gray level j with respect to the kth cluster, c is the number of clusters, q represents the number of the gray levels in the images, Hist(j) is the number of pixels with the gray level j, and m is the weighting exponent for the membership function. By using the Lagrange multiplier method to minimize Eq. (1), the update equations of membership function μ_{kj} and cluster center ν_k are obtained by the following expressions:

$$\mu_{kj} = \frac{1}{\sum_{l=1}^{c} \left(\dfrac{j - \nu_k}{j - \nu_l} \right)^{\frac{2}{m-1}}}, \tag{2}$$

$$\nu_k = \frac{\sum_{j=1}^{q} j \, \text{Hist}(j) \mu_{kj}^{m}}{\sum_{j=1}^{q} \text{Hist}(j) \mu_{kj}^{m}}. \tag{3}$$

The detailed procedures followed for the derivation of Eqs. (2) and (3) are given in Appendix A. Finally, the fuzzy membership images I_F could be easily achieved from the membership matrix μ.

2.A.2. Bilateral filter

The bilateral filter is a nonlinear, edge-preserving and noise-reducing smoothing filter, which employs both geometric closeness and photometric similarity of neighboring

pixels to construct its filter kernel.[17,18] The intensity value of each image pixel would be substituted by a weighted average intensity value from neighborhood pixels. This weight can be based on a Gaussian distribution. More specifically, the weights depend not only on the geometric distance between pixels, but also on the radiometric differences, such as the intensity differences in PET images. By systematically sweeping through image pixels, the bilateral filter could reduce noise while preserving sharp edges in images. Considering a gray-level image I, the output I_B of the bilateral filter is given by

$$
\begin{aligned}
I_B(p) = \frac{1}{\displaystyle\sum_{q \in S} G_{\sigma_s}(\|p-q\|)G_{\sigma_r}(|I(p)-I(q)|)} \\
\times \sum_{q \in S} G_{\sigma_s}(\|p-q\|)G_{\sigma_r}(|I(p)-I(q)|)I(q),
\end{aligned}
\tag{4}
$$

where S is a spatial neighborhood of p, σ_s is the parameter controlling the decay of the geometric factor, and σ_r to control the decay of the radiometric factor, while G_{σ_s} and G_{σ_r} are Gaussian smoothing functions for the spatial kernel and range kernel, respectively. Lee et al.[19] studied in detail the bilateral filter and gave some indications regarding the setting of the parameters σ_s and σ_r. In this work, we simply derive the total variance from the smoothed histogram of the images to determine the values of σ_s and σ_r by default, 3 for the images with low variance and 1 for images with high variance.

2.A.3. 3D Gabor function

The Gabor function has often been used in image processing. It has been demonstrated that it is particularly appropriate for texture representation and discrimination, because it is similar to the human visual system, as explained in the work of Kamarainen et al.[20] With the orientation angles φ set to 90°, θ to 0°, the 3D Gabor function in our work is defined as

$$
G(x,y,z) = S \times \exp\left(-\frac{1}{2\sigma^2}\left(x^2 + y^2 + z^2\right)\right) \times \cos(2\pi f x), \tag{5}
$$

where S is the normalization scale factor, σ is the size of the Gaussian window, and f is the frequency of the sinusoidal. The detailed procedure for implementation of the Gabor filter are given in Appendix B.

2.A.4. The energy function

Let $\phi : \Omega \to R$ be a level set function defined on a domain, the commonly used region-based energy, the Chan–Vese energy,[21] was used as the energy term E to minimize, which is given as follows:

$$
E(\phi) \triangleq \int_\Omega \left(H(\phi)(I-\mu)^2 + (1-H(\phi))(I-\upsilon)^2\right)dxdydz, \tag{6}
$$

where H is the Heaviside function, ϕ is the level set function, which is here a signed distance function, and μ and υ are the

mean values defined as

$$
\mu = \frac{\int_\Omega I(x,y,z)H(\phi)dxdydz}{\int_\Omega H(\phi)dxdydz}, \tag{7}
$$

$$
\upsilon = \frac{\int_\Omega I(x,y,z)(1-H(\phi))dxdydz}{\int_\Omega (1-H(\phi))dxdydz}. \tag{8}
$$

This energy E models the foreground and background as constant intensities represented by their means, μ and υ.

By using the histogram FCM membership images I_F, the bilateral smoothed images I_B, and the averaged Gabor transform images I_G as input in Eq. (6), we obtain the following energy function:

$$
\begin{aligned}
E(\phi) \triangleq &\int_\Omega \Big(H(\phi)(I_F(x,y,z)-\mu_F)^2 \\
&+ (1-H(\phi))(I_F(x,y,z)-\upsilon_F)^2\Big)dxdydz \\
&+ \int_\Omega \Big(H(\phi)(I_B(x,y,z)-\mu_B)^2 \\
&+ (1-H(\phi))(I_B(x,y,z)-\upsilon_B)^2\Big)dxdydz \\
&+ \int_\Omega \Big(H(\phi)(I_G(x,y,z)-\mu_G)^2 \\
&+ (1-H(\phi))(I_G(x,y,z)-\upsilon_G)^2\Big)dxdydz.
\end{aligned}
\tag{9}
$$

Thus the level set function ϕ can be obtained from the above defined energy functional by differentiation using the gradient descent method,

$$
\frac{\partial \phi}{\partial t} = -\frac{\partial E}{\partial \phi}. \tag{10}
$$

According to Eq. (9), Eq. (10) is equivalent to the following evolution equation:

$$
\begin{aligned}
\frac{\partial \phi}{\partial t} = &-\delta(\phi)\left((I_F-\mu_F)^2 - (I_F-\upsilon_F)^2\right) \\
&-\delta(\phi)\left((I_B-\mu_B)^2 - (I_B-\upsilon_B)^2\right) \\
&-\delta(\phi)\left((I_G-\mu_G)^2 - (I_G-\upsilon_G)^2\right),
\end{aligned}
\tag{11}
$$

where $\delta(\phi)$ is the Dirac delta function, defined as the derivative of the Heaviside function with respect to ϕ.

2.A.5. Lattice Boltzmann method

LBM is a relatively new simulation technique for complex fluid systems, which is solved to simulate the flow of a Newtonian fluid with collision models, such as Bhatnagar–Gross–Krook (BGK). By simulating streaming and propagation processes across a limited number of particles, the interaction of internal particles evinces a microcosm of this macroviscous fluid flow. This model has attracted the interest of image segmentation researchers because of its simplicity and highly parallelizable nature.[22,23] In our model, the D3Q19 LBM lattice structure, where D is the dimension and Q is the number of velocity directions, representing the three-dimensional 19 velocity lattice, was used to solve the level set equation.[24] The evolution equation of LBM is given by

$$
f_i(\vec{r}+\vec{e}_i, t+1) - f_i(\vec{r},t) = \Omega_{BGK}, \tag{12}
$$

where f_i is the function for the particle distribution, $\vec{e}_i$ is the directional unit vector, and Ω_{BGK} is the BGK model, as given below,

$$\Omega_{\text{BGK}} = \frac{1}{\tau}\left[f_i^{eq}(\vec{r},t) - f_i(\vec{r},t)\right] + \frac{2\times\tau-1}{2\times\tau}\times\vec{F}\times\vec{e}_i, \qquad (13)$$

where τ is the relaxation time and f_i^{eq} is the local equilibrium particle distribution, which is defined as follows when modeling the diffusion phenomenon:

$$f_i^{eq}(\rho) = \rho A_i, \qquad (14)$$

where $\rho = \sum f_i$, and A_i is the weighting factor. The body force $\vec{F}$ in Eq. (13) can be defined as

$$\begin{aligned}
\vec{F} \triangleq \lambda\frac{\partial\phi}{\partial t} = &-\lambda\delta(\phi)\left((I_F-\mu_F)^2-(I_F-\upsilon_F)^2\right)\\
&-\lambda\delta(\phi)\left((I_B-\mu_B)^2-(I_B-\upsilon_B)^2\right)\\
&-\lambda\delta(\phi)\left((I_G-\mu_G)^2-(I_G-\upsilon_G)^2\right),
\end{aligned} \qquad (15)$$

where λ is the parameter used to modulate the strength of $\vec{F}$ on the active contour motion. The larger the value of λ, the larger the segmentation contours. In this study, λ was set as 1 for all studies acquired on different PET scanners and using different protocols to test the generalness and robustness of the MASAC algorithm. To achieve better results in specific setting, the parameter λ in MASAC could be used to calibrate the algorithm for a specific scanner/data acquisition or reconstruction protocols by trial-and-error experiments using phantom studies. The proposed level set equation can thus be solved to obtain the final active contour segmentation results using the following lattice Boltzmann evolution equation:

$$\begin{aligned}
f_i(\vec{r}+\vec{e}_i,t+1) = &f_i(\vec{r},t) + \frac{1}{\tau}\left[f_i^{eq}(\vec{r},t)-f_i(\vec{r},t)\right]\\
&-\lambda\delta(\phi)\frac{2\times\tau-1}{2\times\tau}\left((I_F-\mu_F)^2-(I_F-\upsilon_F)^2\right)\\
&-\lambda\delta(\phi)\frac{2\times\tau-1}{2\times\tau}\left((I_B-\mu_B)^2-(I_B-\upsilon_B)^2\right)\\
&-\lambda\delta(\phi)\frac{2\times\tau-1}{2\times\tau}\left((I_G-\mu_G)^2-(I_G-\upsilon_G)^2\right).
\end{aligned} \qquad (16)$$

2.B. Phantom and clinical studies

To evaluate the proposed method's performance, the algorithm was evaluated using phantom and clinical studies. Phantom studies were performed using the 4D extended cardiac-torso (XCAT) phantom, a realistic whole-body computer model based on nonuniform rational B-spline surfaces representing the human anatomy and physiology.[25] Realistic lesion shapes with three different levels of tracer uptake, heterogeneous within the tumor, were modeled based on clinical PET images and incorporated within the XCAT phantom according to the procedures proposed by Le Maitre et al.[26] In our study, thresholding was used to delineate different irregularly shaped heterogeneities within tumors extracted from clinical images. Thereafter, the 3D mesh of segmented structures was reconstructed using AMIRA software (FEI, USA) and converted into nonuniform rational B-spline (NURBS)

surfaces using RHINOCEROS software (CADLINK, France). Finally, different NURBS surfaces, each representing a specific activity level, were created to model the heterogeneities within the tumor. In our study, three levels of activity concentration were modeled as it was deemed to represent more realistically tumors than two levels of heterogeneities adopted by Le Maitre et al.[26] The levels of the uptake ratio within the lesions were 10:9:8, with the background outside the lesions set to 1. These ratios were observed in many clinical studies and are also in agreement with contrast levels adopted by other groups.[27–30] Ten different lesions were simulated in two different lung locations, without or with high activity in the nearby myocardium (volume range 3.26–41.44 cm^3). Respiration motion was not considered for the simulated studies in this work. The attenuation maps generated at 511 keV were used for attenuation correction during image reconstruction. The matrix size of all images (both activity and attenuation maps) was 288×288 with a voxel size of $2.13\times2.13\times3.27$ mm. For PET simulation, we used an analytical projector and ordered subsets expectation maximization reconstruction algorithm implemented within the software for tomographic image reconstruction (STIR) modeling a GE Discovery RX PET/CT scanner.[31] The 511-keV attenuation maps were applied to model attenuation during the simulation process.

The clinical studies include six cases of pharyngo-laryngeal squamous cell carcinoma (PH), volume range (4.03–32.92 cm^3) and ten cases of nonsmall cell lung cancer (LU), volume range (1.12–46.28 cm^3), where the ground truth 3D contours were derived from the macroscopic specimen obtained after surgery.[32,33] The mean tumor to background ratio in the clinical studies is 7.98 ± 5.83 (range 2.20–21.10). Gated image acquisition was performed for the LU datasets. Two patients in LU had lesions connected with a high activity area nearby, which are difficult to distinguish and separate. These two lesions were carefully cropped to exclude the nearby high activity region as performed in our previous work.[34]

2.C. Comparative study and evaluation metrics

This proposed method was compared with the contourlet-based active contour algorithm (CAC)[34] and Schaefer's thresholding method (ST).[35] Widely adopted image segmentation metrics including the relative volume error (RE), Dice similarity coefficient (DSC), and classification error (CE) were used to evaluate quantitatively the performance of these PET image segmentation algorithms. They are defined as

$$\text{RE} = \frac{(\text{volume}(\text{SM})-\text{volume}(\text{GT}))}{\text{volume}(\text{GT})}\times100\%, \qquad (17)$$

$$\text{DSC} = \frac{2\times|\text{SM}\cap\text{GT}|}{|\text{SM}|+|\text{GT}|}, \qquad (18)$$

$$\text{CE} = \frac{\text{FN}+\text{FP}}{\text{TP}}\times100\%, \qquad (19)$$

where SM and GT represent the output of PET image segmentation method and the ground truth, respectively, FN and FP are the false negative and false positive errors, respectively, and TP is total number of voxels within the

ground truth volume.[36–38] An additional metric, K, reporting the ratio of FN/TP is also added to split CE into contributions from FN and FP,

$$K = \frac{FN}{TP} \times 100\%. \tag{20}$$

The results are expressed as means with standard deviations, and presented as box-and-whisker plots as appropriate, providing lower to upper quartile (25–75 percentile, central box), the median (middle line of the box), and the minimum to maximum value.

3. RESULTS

Figure 2 summarizes the performance metrics of the segmentations for the phantom studies. The means and standard deviations of the metrics for all phantom studies are presented in Table I. For the simulated XCAT phantom studies (PSs), MASAC and CAC provide similar segmentations of the different lesions, while ST fails to achieve reliable results. For MASAC, it can be seen that there is no major difference between the mean values of the metrics between phantom studies without high nearby activity (PS1) and phantom studies with nearby high background activity (PS2), indicating the robustness of MASAC in different clinical scenarios. Figures 3 and 4 show representative contours on realistic anthropomorphic phantom studies without and with nearby high background activity, respectively.

Figure 5 shows box-and-whisker plots for all metrics across the clinical studies. For patients involved in the study

TABLE I. Mean and standard deviation of the metrics used for assessment of the performance of the proposed PET image segmentation algorithm for all phantom studies. The abbreviations used are as defined in the captions of the previous figures.

Metrics	MASAC	CAC	ST
PS1			
RE(%)	−12.60 ± 11.66	14.78 ± 8.40	−25.10 ± 13.78
DSC	0.90 ± 0.04	0.92 ± 0.03	0.85 ± 0.08
CE(%)	18.23 ± 7.24	17.35 ± 7.92	25.60 ± 13.00
K(%)	15.41 ± 9.21	1.28 ± 1.88	25.35 ± 13.39
PS2			
RE(%)	−14.10 ± 12.65	19.97 ± 14.24	8.69 ± 28.06
DSC	0.89 ± 0.06	0.90 ± 0.06	0.67 ± 0.12
CE(%)	20.15 ± 8.78	22.51 ± 16.00	68.87 ± 29.76
K(%)	17.13 ± 10.47	1.27 ± 1.88	30.09 ± 15.54
PSs			
RE(%)	−13.35 ± 11.87	17.37 ± 11.69	−8.21 ± 27.63
DSC	0.89 ± 0.05	0.91 ± 0.05	0.76 ± 0.14
CE(%)	19.19 ± 7.89	19.93 ± 12.57	47.24 ± 31.50
K(%)	16.27 ± 9.64	1.28 ± 1.83	27.72 ± 14.32

protocol with those presenting with connected high-uptake regions excluded (CSs), the mean DSC for MASAC was found to be larger than 0.70, with a DSC larger than 0.60 (13 of 14 patients) in most of the cases, indicating a good quality of segmentation of the clinical studies. Table II lists the mean and standard deviations of the metrics for the clinical studies. It can be seen that CAC provides for the lowest mean RE (−8.38% ± 27.49%), while MASAC
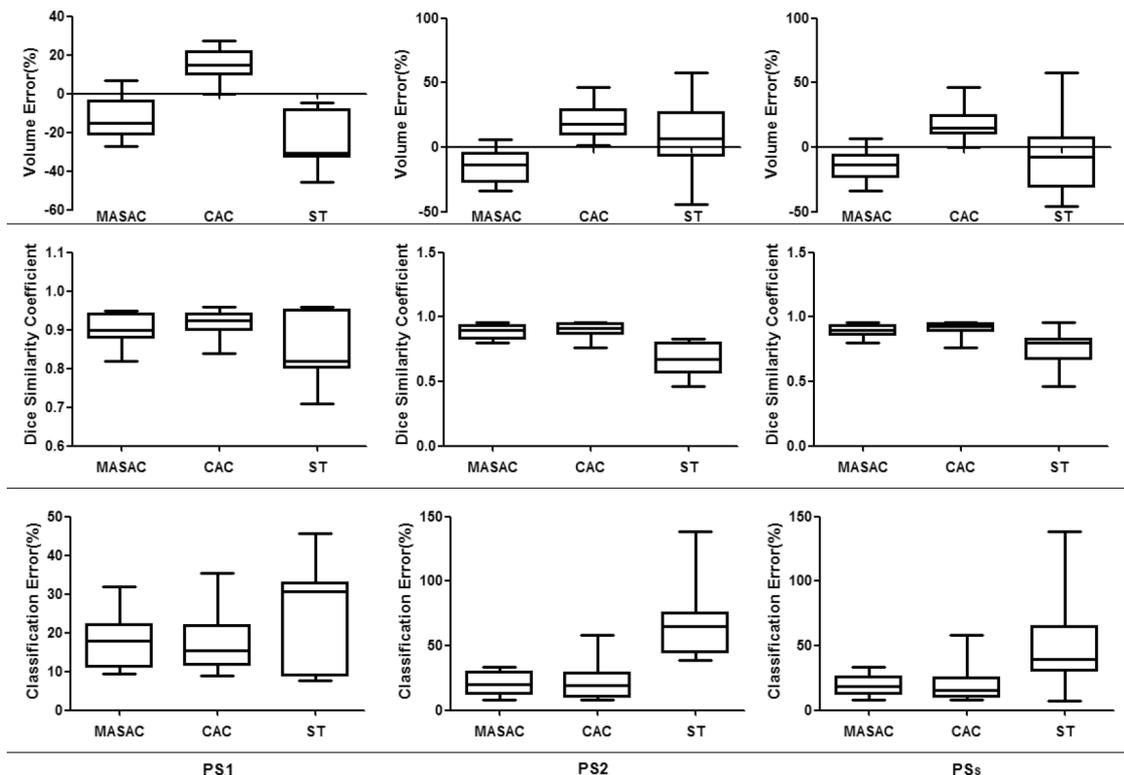

FIG. 2. Box-and-whisker plots of each metric across the phantom studies. The presented results represent phantom studies without nearby high activity (PS1), phantom studies with nearby high activity (PS2), and all phantom studies (PSs).
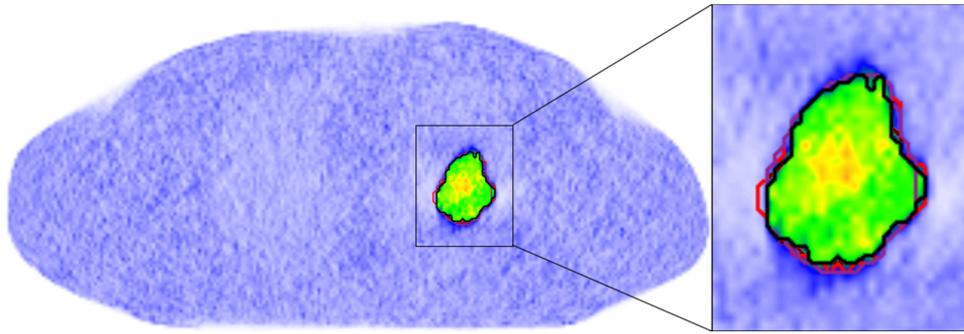
FIG. 3. Representative segmentation result of a realistic anthropomorphic phantom study (without high nearby background activity) showing contours extracted by MASAC (purple), CAC (red), and ST (black), compared to the ground truth (brown). The initial cropping area has been shown in the rectangle area. (See color online version.)

achieves the best mean DSC (0.71 ± 0.09) and mean CE ($53.92\% \pm 12.65\%$), respectively. Considering the variability in tumor characteristics and data acquisition/reconstruction protocols, the clinical studies show larger standard deviations for the metrics used compared to the phantom studies (Tables I and II). It should be emphasized that in some cases, the lesion contours defined on the macroscopic specimen do not perfectly coincide with the visible high metabolic volume. This has already been observed by other groups using the same clinical datasets.

For the pharyngolaryngeal carcinoma clinical studies (Fig. 6), it can be observed that both MASAC and CAC can properly delineate the tumors from PET images. The contours extracted by the MASAC algorithm of a representative patient from the lung cancer studies are presented in Fig. 7. It can be seen that the MASAC algorithm still performs well in the case of heterogeneous tumors with necrotic regions. The contours extracted by MASAC fit well the ground truth defined on the macroscopic specimen. As can be seen in Table II the mean DSCs of MASAC for PH and LU8 datasets are almost similar. However, the mean CE for the PH datasets was slightly larger than that in LU8.

4. DISCUSSION

Molecular PET images provide relevant information that can be used for radiation therapy treatment planning, therapy monitoring, and prognosis and the detection of recurrences or metastatic disease. For many of these clinical applications, accurate delineation of the metabolic tumor volume from PET is of high importance. In the present situation, automated PET image segmentation algorithms are commonly viewed as efficient tools for the clinicians, providing them with a good starting point for an initial review and adjustments if needed of the delineations. A number of automated PET image segmentation approaches have been proposed during the last few years. The majority of these approaches have been tested on either simplistic phantom studies or patient studies where the ground truth is in most of the cases defined manually by experts since it is practically unknown. Currently, there is still no consensus on the most accurate and robust algorithm that should be adopted for reliable routine clinical PET image segmentation.

With the introduction of fuzziness for the membership of image pixels, FCM types of algorithms have proven their effectiveness for image segmentation. However, FCM is very sensitive to noise, because it does not consider the spatial information. Some studies reported on the modification of the objective function to incorporate local spatial information into the FCM algorithm to improve its performance.[39,40] These algorithms incorporate a parameter α which controls the trade-off between the original images and their corresponding averaged images. To keep a reasonable balance between noise-immunity and detail-preservation, the selection of α
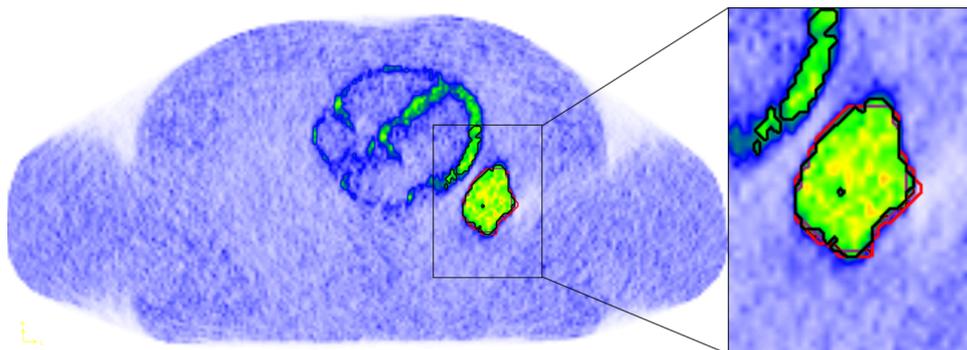

FIG. 4. Representative segmentation result of a realistic anthropomorphic phantom study (with nearby high activity in the myocardium) showing contours extracted by MASAC (purple), CAC (red), and ST (black), compared to the ground truth (brown). The initial cropping area has been shown in the rectangle area. (See color online version.)
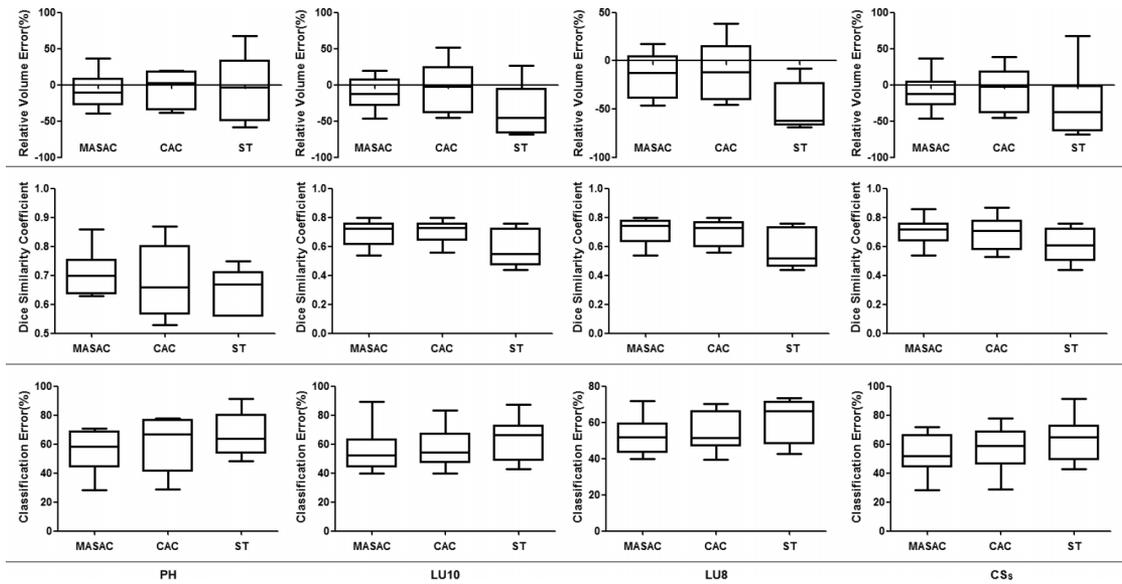
FIG. 5. Box-and-whisker plots of each metric across the clinical studies. The presented results represent six Louvain pharyngolaryngeal carcinoma studies (PH), ten Louvain lung cancer studies (LU10), Louvain lung cancer studies with two studies presenting with connected high-uptake regions excluded (LU8), and all clinical studies with two studies presenting with connected high-uptake regions excluded (CSs).

is difficult and as such, its selection remain heuristic and user-dependent.

The active contour model is one of the most commonly used approaches for image segmentation. We recently reported on a novel contourlet-based active contour algorithm for PET image segmentation.[34] However, the main shortcoming of this technique is that it largely depends on user-defined parameters, which could substantially influence the segmentation results. Consistent selection and adjustment of these parameters are

TABLE II. Mean and standard deviation of the metrics used for assessment of the performance of the proposed PET image segmentation algorithm for all clinical studies. The abbreviations used are as defined in the captions of the previous figures.

Metrics	MASAC	CAC	ST
PH			
RE(%)	−7.95 ± 25.98	−4.66 ± 25.29	−3.54 ± 48.10
DSC	0.71 ± 0.08	0.68 ± 0.13	0.65 ± 0.08
CE(%)	55.64 ± 16.00	60.31 ± 19.28	66.62 ± 15.32
K(%)	31.80 ± 13.70	32.51 ± 19.93	35.09 ± 20.39
LU10			
RE(%)	−11.06 ± 23.34	−1.92 ± 34.10	−35.54 ± 34.68
DSC	0.70 ± 0.09	0.70 ± 0.08	0.59 ± 0.13
CE(%)	56.21 ± 14.86	57.83 ± 13.34	63.99 ± 13.64
K(%)	33.65 ± 13.89	29.85 ± 16.64	49.76 ± 19.80
LU8			
RE(%)	−13.56 ± 23.28	−11.18 ± 30.43	−48.11 ± 24.52
DSC	0.71 ± 0.09	0.70 ± 0.09	0.58 ± 0.14
CE(%)	52.63 ± 10.50	54.92 ± 11.12	61.05 ± 12.35
K(%)	33.09 ± 15.68	33.04 ± 17.25	54.58 ± 18.22
CSs			
RE(%)	−11.15 ± 23.66	−8.38 ± 27.49	−29.01 ± 41.68
DSC	0.71 ± 0.09	0.69 ± 0.10	0.61 ± 0.12
CE(%)	53.92 ± 12.65	57.23 ± 14.74	63.43 ± 13.44
K(%)	32.54 ± 14.32	32.81 ± 17.69	46.23 ± 20.95

difficult and could hardly be achieved by trial-and-error experiments.

To overcome the above problems, the MASAC algorithm is implemented by modifying the objective function in the conventional active contour model, consistent with approaches used elsewhere.[34,39,40] We used an active contour model framework to incorporate the histogram and localized and textural information in our approach. The FCM clustering membership was taken as one of the inputs for the energy function. It should be noted that the number of gray levels in PET images is usually much smaller than the number of pixels. Therefore, the gray level histogram of the images was used for the clustering instead of image pixels to enhance efficiency. Then the membership matrix was used as input for the MASAC algorithm, providing probabilistic fuzzy value estimates for the segmentation.

The bilateral filter was incorporated in our algorithm to generate weighted average values from neighboring pixels to cope with image noise and utilize the statistical information in local regions. To improve its robustness, the parameters, σ_s and σ_r, of the filter were predefined by the variance derived from the smoothed histogram of the images. Because of the diversity of PET images depending on clinical indications, it is impossible to use the same parameters for all images. For lesions with high variance, it tends to require lower values for σ_s and σ_r to preserve the edges in the images, while higher values are required for large lesions with low variance.

The complexity of human anatomy and physiology and neoplastic growth produce sophisticated texture patterns in PET images. For texture discrimination, we propose to use Gabor functions whose frequency and orientation representations are similar to those of the human visual system. Sandberg et al.[41] used the Gabor function to discriminate textures in a curve evolution based level set approach. It would take much time for 45 Gabor transforms
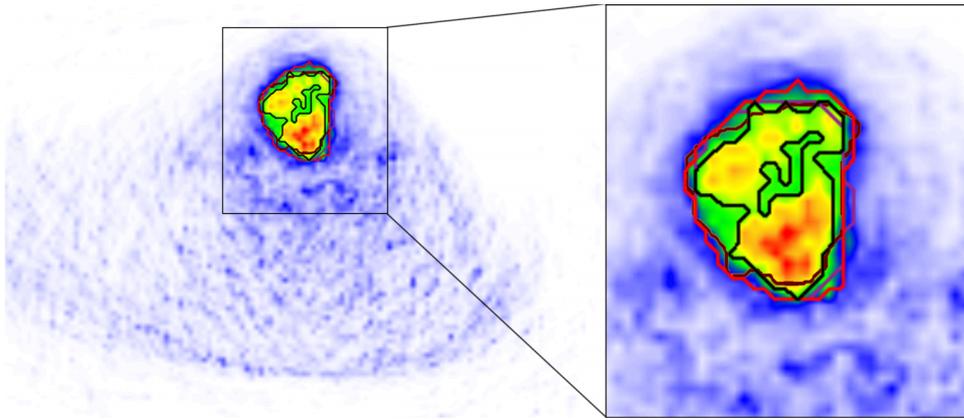
FIG. 6. Representative clinical pharyngolaryngeal carcinoma PET study showing contours extracted by MASAC (purple), CAC (red), and ST (black), compared to the ground truth (brown). The initial cropping area has been shown in the rectangle area. (See color online version.)

when dealing with 3D images. Xu et al.[42] focused on texture analysis based on Gabor filters to extract feature vectors and follow a probabilistic method to get the segmentation results. However, it is assumed that a large database is available for training to get feature vectors. In this work, the convolutions of Gabor functions with the original images were used to obtain 16 different channels, to keep a reasonable balance between efficiency and effectiveness. The average of these channels would be the input of the multichannel active contour algorithm.

By combining histogram clustering, bilateral filtering, and Gabor transformation introduced above, we obtain the MASAC energy function devoted to dealing with noise and intensity inhomogeneity, which are generally observed in PET images. In this work, the proposed algorithm is evaluated using both phantom and clinical studies where tumors differed in biology, phenotype, size, and location. Our results show that the application of histogram and localized and textural features endows the MASAC algorithm with an extensive capture range, exhibiting robustness to different types of tumors and clinical scenarios. Although the CAC algorithm seems to achieve more accurate results in terms of quantifying tumor volumes, the model depends highly on user-defined parameters, which require adjustments not only for different scanners or different protocols, but also for different

patients scanned on the same scanner with the same protocol, even for different lesions in the same patient. The use of different parameters might influence the segmentation results significantly, and these parameters most probably cannot be adequately parameterized experimentally. Compared with the various parameters used in CAC, the MASAC algorithm keeps the same parameters for all the datasets used in this work.

From Tables I and II, it seems that our algorithm has a systematic underestimation of the tumor for the datasets that we test. The factor λ in MASAC could be used for calibration purposes but was set to 1 for all cases in this study. This aim of this work is to develop a generic and robust algorithm for PET image segmentation suitable for adoption in clinical setting. Therefore, we did not calibrate the method for any specific scanner or protocol and evaluated its performance using different datasets with λ set as 1. The results might be improved further if phantom experiments are used to calibrate the method targeting a specific scanner/protocol.

The MASAC algorithm has shown comparable or even better performance compared with those observed in other studies involving validation against clinical datasets. Our results using the PH database (six studies) showed a mean DSC of 0.71 and mean CE of 55.64% compared with 0.54 and 50.20%, respectively, (for seven studies) reported for a previous algorithm in our previous work.[43] The clinical
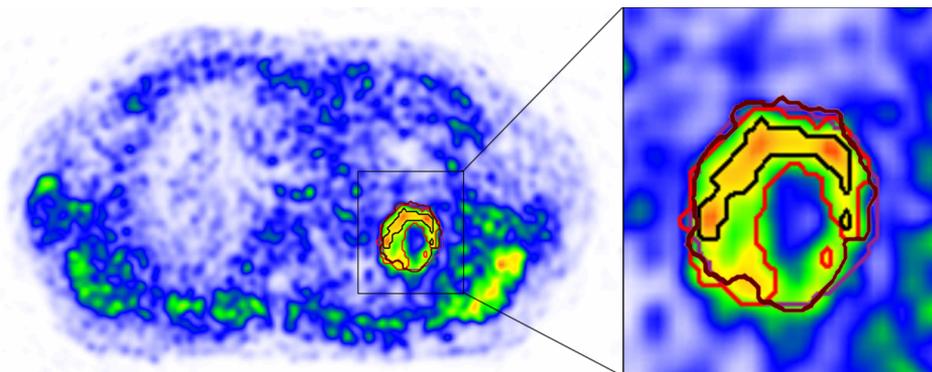

FIG. 7. Representative clinical nonsmall cell lung cancer PET study showing contours extracted by MASAC (purple), CAC (red), and ST (black), compared to the ground truth (brown). The initial cropping area has been shown in the rectangle area. (See color online version.)

assessment of the LU database (ten studies) resulted in a mean DSC of 0.70 and mean CE of 56.21%, compared with 0.41 and 112.86%[43] and 0.54 and 60.58%[34] (nine studies), respectively. Schaefer *et al.*[3] evaluated consensus contours from different PET segmentation algorithms using clinical PET images and reported a mean DSC varying from 0.59 to 0.67.

The mean DSC achieved by MASAC for the clinical studies was greater than 0.70, with a DSC larger than 0.60 (13 of 14 patients) in most of the cases. Zijdenbos *et al.*[44] suggested that a DSC > 0.7 indicates good agreement, therefore we adopted the value of 0.7 as a criterion for reasonably good agreement. Similar thresholds were also adopted in other studies. For instance, the automatic multimodality segmentation algorithm developed by Yang *et al.*[45] achieved a mean DSC of 0.74, compared with a mean DSC of 0.65 achieved when using only PET images.[46]

As can be seen in Fig. 4, the MASAC algorithm achieves good performance for the lesion with a high nearby activity. However, if the lesions are connected with the high nearby activity, the results could be debatable even if the algorithm achieves the segmentation results close to the ground truth. In this case, it is difficult to distinguish only depending on the PET images, and the clinicians need to judge by other means of data. Therefore, it seems that the results of LU8 are more objective, with two connected studies excluded.

By comparing PSs and CSs in detail in Tables I and II, it can be noted that the mean RE of MASAC in CSs is better than that in PSs while the mean DSC and CE of MASAC in PS are poorer than those in CSs. These results may be explained by the fact that the ground truth for some cases in CSs fits poorly the visible high metabolic area in PET images, and as such, poor DSC and CE and good RE results are achieved. The RE metric is used to describe the segmented results to the ground truth in terms of volume, while DSC and CE are used to quantify their spatial positions and geometrical shapes. Since each metric only reflects one aspect of the results, it is recommended to combine different quality measures for overall performance assessment.

It was found that the mean CE of MASAC in PH was slightly larger than that in LU8, as shown in Table II. This may be explained by different acquisitions taken in these datasets. In PH datasets, the matrix size was 128×128 and the patients were scanned with larger voxels, while smaller voxels with larger matrix size were used for the patients in the LU8 datasets. Therefore, in clinical routine, smaller voxels with a larger matrix size would be preferred if possible at all, especially for small lesions, where the partial volume effect renders the problem much more complex.

To some extent, the MASAC algorithm is a hybrid of three existing algorithms, which could be straightforwardly extended to multimodality segmentation by having inputs set as different modalities. The active contour model was combined and integrated with other algorithms, e.g., Otsu's threshold method used as the initial contour and the histogram FCM clustering method to achieve the probabilistic estimate membership matrix for the energy function. The concept of combining several segmentation methods has already been proposed in the literature. Schaefer *et al.*[3] used a consensus

algorithm combining the results of different PET segmentation methods to achieve robustness against the inconsistent performance of individual segmentation techniques. Likewise, McGurk *et al.*[47] demonstrated the improved robustness achieved by this approach using experimental phantom measurements by combining the results of five individual segmentation methods. It should be noted that this different approach makes use of a number of segmentation results, while MASAC uses a variety of image processing techniques to achieve accurate PET image segmentation.

Lankton *et al.*[48] and Li *et al.*[49] provided ways for localizing active contour models. The key idea behind their localized active contour model was to design the local regions along the curve evolution using the convolution with a mask function. In this work, we introduce a bilateral filter that uses the local weighted average values for the evolution to guarantee that the relevant statistics evolve together. Moreover, to speed up the segmentation process, we employ several methods, such as parallel programming, lattice Boltzmann method, and clustering on gray-level histogram to efficiently delineate structure boundaries. In this fast framework, we achieved execution times of 1–5 s for most of the studies, which would meet daily clinical routine requirements.

Despite the many promising aspects of the proposed algorithm, its limitations need to be emphasized. First, it is difficult to achieve accurate segmentation for cases with high nearby activity connected. Two cases from the LU databases were carefully cropped to exclude the high uptake region nearby. These cases were used to demonstrate the limitations of the proposed technique in specific situations (presence of nearby high physiologic uptake) and conditions that need careful consideration. Second, as shown in Figs. 3, 4, 6, and 7, the initial cropping area around the tumor to segment images was delineated to enhance efficiency. We observed that different cropping areas have little effect on the segmentation results, except in those cases where there is a high uptake region nearby the tumor or where the image characteristics are highly modified by the initial cropping area.

The datasets used in this work to test the proposed approach also have some inherent limitations. The ground truth in clinical studies is not known *a priori* and is difficult if not impossible to obtain. Most published studies used manual contouring performed by experienced physicians as the ground truth.[50] However, manual contouring is prone to large interobserver and intraobserver variabilities and, as such, is questionable as reference for assessment of image segmentation techniques. The contouring variability was assessed quantitatively in multicenter studies reporting a significant interobserver variability with a generalized DSC within the range 0.40–0.80.[51,52] In our work, we preferred to use the macroscopic specimen keeping in mind the limitations of this approach. In essence, the contour defined on the macroscopic specimen does not perfectly follow the visible high metabolic volume in some cases, meaning that some uncertainties persist because of the shrinkage and deformation of the specimen. Another possible cause may be the inconsistency between the high metabolic volume in PET images and the tumor volume in reality due to heterogeneities

within the tumor. Although no reference is perfect, we believe that the macroscopic specimen may offer a better approximate of the boundary of the tumor. A large database of test images acquired with different acquisition/reconstruction protocols on various PET scanners, presenting with various lesion sizes and contrasts located in various regions in the body, with different background activity as well as individual anatomical variations, would be required to draw reliable conclusions.

Another issue to emphasize is that all segmentations are derived from image characteristics. Therefore, provided that a reliable data acquisition/reconstruction protocol is used to produce good quality images, the proposed method is likely to produce accurate segmentation of target volumes. For example, the long acquisition time of oncological PET studies might induce motion artifacts, thus degrading PET image quality. One of the major problems is respiratory artifacts caused by patients' free breathing during PET scanning. Lesions located in the thorax or upper abdomen would spread throughout along with respiratory motion, leading to an underestimation of the tracer uptake and an overestimation of lesions' volume. A number of strategies including 4D PET/CT data acquisition techniques were devised to reduce the impact of respiratory motion.

5. CONCLUSION

We developed a novel generic and robust method for fully automated PET image segmentation using an active contour model. Besides, its performance was evaluated using phantom and clinical studies, compared with other algorithms. Our results demonstrate that the proposed method is applicable to various types of lesions assessed in our work and could generate accurate and consistent target volume delineations, potentially resulting in reduced interobserver variability of manual delineation and improved accuracy in treatment planning and outcome evaluation.

CONFLICT OF INTEREST DISCLOSURE

The authors have no COI to report.

APPENDIX A: HISTOGRAM FUZZY C-MEANS CLUSTERING

The objective function to minimize is given as

$$J = \sum_{k=1}^{c}\sum_{j=1}^{q}\text{Hist}(j)\mu_{kj}^{m}(j-\nu_k)^2 \tag{A1}$$

with

$$\sum_{k=1}^{c}\mu_{kj} = 1, \tag{A2}$$

where ν_k represents the center of the kth cluster, μ_{kj} denotes the fuzzy membership value of the gray level j as to the kth cluster, c is the number of clusters, q represents the number of the gray levels in the images, $\text{Hist}(j)$ is the number of pixels with the gray level j, and m is the weighting exponent for the membership function.

In order to minimize Eq. (A1), a new objective function could be constructed using the Lagrange multiplier method,

$$F = \sum_{k=1}^{c}\sum_{j=1}^{q}\text{Hist}(j)\mu_{kj}^{m}(j-\nu_k)^2 - \lambda\left(\sum_{k=1}^{c}\mu_{kj}-1\right). \tag{A3}$$

The minimization of the new objective function Eq. (A3) leads to

$$\frac{\partial F}{\partial \lambda} = \sum_{k=1}^{c}\mu_{kj}-1 = 0, \tag{A4}$$

$$\frac{\partial F}{\partial \mu_{kj}} = \text{Hist}(j)m\mu_{kj}^{m-1}(j-\nu_k)^2 - \lambda = 0, \tag{A5}$$

$$\frac{\partial F}{\partial \nu_i} = \sum_{j=1}^{q}\text{Hist}(j)\mu_{kj}^{m}[-2(j-\nu_k)] = 0. \tag{A6}$$

To solve Eqs. (A4)–(A6), the update equations of membership function μ_{kj} and cluster center ν_k could be obtained using the following expressions:

$$\mu_{kj} = \frac{1}{\sum_{l=1}^{c}\left(\frac{j-\nu_k}{j-\nu_l}\right)^{\frac{2}{m-1}}}, \tag{A7}$$

$$\nu_k = \frac{\sum_{j=1}^{q}j\,\text{Hist}(j)\mu_{kj}^{m}}{\sum_{j=1}^{q}\text{Hist}(j)\mu_{kj}^{m}}. \tag{A8}$$

APPENDIX B: 3D GABOR FILTER

A 3D Gabor filter is a 3D Gaussian kernel function modulated by a sinusoidal wave, which is defined as

$$G(x,y,z) = S\times\exp\left(-\frac{1}{2}\left(\left(\frac{x'}{\sigma_x}\right)^2 + \left(\frac{y'}{\sigma_y}\right)^2 + \left(\frac{z'}{\sigma_z}\right)^2\right)\right)$$
$$\times\exp(j2\pi(xu+yv+zw)) \tag{B1}$$

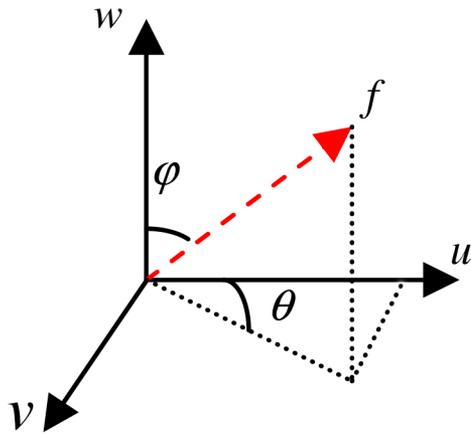FIG. 8. The frequency domain for the 3D Gabor filter.

with

$$u = f \sin\varphi \cos\theta, \tag{B2}$$

$$v = f \sin\varphi \sin\theta, \tag{B3}$$

$$w = f \cos\varphi, \tag{B4}$$

$$[x'\,y'\,z']^T = R \times [x\,y\,z]^T, \tag{B5}$$

where S is the normalization scale factor, f is the central frequency, φ and θ are the orientation angles in the 3D frequency domain (Fig. 8), and σ_x, σ_y, and σ_z define the standard deviations of the Gaussian envelope in x, y, and z axis, respectively. By changing σ, the size of the image region to be analyzed was also changed.

The real part of the Gabor function is utilized in our study to detect the tumor in images because it has been reported to be useful for target detection, which is given as

$$G(x,y,z) = S \times \exp\left(-\frac{1}{2}\left(\left(\frac{x'}{\sigma_x}\right)^2 + \left(\frac{y'}{\sigma_y}\right)^2 + \left(\frac{z'}{\sigma_z}\right)^2\right)\right)$$
$$\times \cos(2\pi(xu + yv + zw)). \tag{B6}$$

With the orientation angles φ set to 90°, θ to 0°, and $\sigma_x = \sigma_y = \sigma_z = \sigma$, the 3D Gabor function that we use could be simplified as

$$G(x,y,z) = S \times \exp\left(-\frac{1}{2\sigma^2}(x^2 + y^2 + z^2)\right) \times \cos(2\pi f x). \tag{B7}$$

[a)]Electronic mail: habib.zaidi@hcuge.ch

[1]T. Shepherd, M. Teras, R. R. Beichel, R. Boellaard, M. Bruynooghe, V. Dicken, M. J. Gooding, P. J. Julyan, J. A. Lee, S. Lefevre, M. Mix, V. Naranjo, X. Wu, H. Zaidi, Z. Zeng, and H. Minn, "Comparative study with new accuracy metrics for target volume contouring in PET image guided radiation therapy," IEEE Trans. Med. Imaging **31**(11), 2006–2024 (2012).

[2]B. Berthon, C. Marshall, M. Evans, and E. Spezi, "Evaluation of advanced automatic PET segmentation methods using nonspherical thin-wall inserts," Med. Phys. **41**(2), 022502 (9pp.) (2014).

[3]A. Schaefer, M. Vermandel, C. Baillet, A. S. Dewalle-Vignion, R. Modzelewski, P. Vera, L. Massoptier, C. Parcq, D. Gibon, T. Fechter, U. Nemer, I. Gardin, and U. Nestle, "Impact of consensus contours from multiple PET segmentation methods on the accuracy of functional volume delineation," Eur. J. Nucl. Med. Mol. Imaging **43**(5), 911–924 (2016).

[4]M. Hatt, C. Cheze le Rest, A. Turzo, C. Roux, and D. Visvikis, "A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET," IEEE Trans. Med. Imaging **28**(6), 881–893 (2009).

[5]D. Markel, H. Zaidi, and I. El Naqa, "Novel multimodality segmentation using level sets and Jensen–Rényi divergence," Med. Phys. **40**(12), 121908 (8pp.) (2013).

[6]J. Lapuyade-Lahorgue, D. Visvikis, O. Pradier, C. Cheze Le Rest, and M. Hatt, "SPEQTACLE: An automated generalized fuzzy C-means algorithm for tumor delineation in PET," Med. Phys. **42**(10), 5720–5734 (2015).

[7]H. Zaidi and I. E. Naqa, "PET-guided delineation of radiation therapy treatment volumes: A survey of image segmentation techniques," Eur. J. Nucl. Med. Mol. Imaging **37**, 2165–2187 (2010).

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

[9]J. A. Lee, "Segmentation of positron emission tomography images: Some recommendations for target delineation in radiation oncology," Radiother. Oncol. **96**(3), 302–307 (2010).

[10]N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. SMC **9**(1), 62–66 (1979).

[11]G. Liu, Y. Zhang, and A. Wang, "Incorporating adaptive local information into fuzzy clustering for image segmentation," IEEE Trans. Image Process. **24**(11), 3990–4000 (2015).

[12]D. Zhou and H. Zhou, "A modified strategy of fuzzy clustering algorithm for image segmentation," Soft Comput. **19**(11), 3261–3272 (2015).

[13]L. Szilagyi, Z. Benyo, S. M. Szilagyi, and H. S. Adam, "MR brain image segmentation using an enhanced fuzzy C-means algorithm," in *25th Annual International Conference of IEEE EMBS* (IEEE, Cancun, Mexico, 2003), Vol. 1, pp. 724–726.

[14]F. Zhao, "Fuzzy clustering algorithms with self-tuning non-local spatial information for image segmentation," Neurocomputing **106**(6), 115–125 (2013).

[15]Z. I. Botev, J. F. Grotowski, and D. P. Kroese, "Kernel density estimation via diffusion," Ann. Stat. **38**(5), 2916–2957 (2010).

[16]B. Foster, U. Bagci, Z. Xu, B. Dey, B. Luna, W. Bishai, S. Jain, and D. J. Mollura, "Segmentation of PET images for computer-aided functional quantification of tuberculosis in small animal models," IEEE Trans. Biomed. Eng. **61**(3), 711–724 (2014).

[17]C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the International Conference on Computer Vision* (IEEE, Bombay, India, 1998), pp. 839–846.

[18]M. Elad, "On the origin of the bilateral filter and ways to improve it," IEEE Trans. Image Process. **11**(10), 1141–1151 (2002).

[19]J. A. Lee, X. Geets, V. Gregoire, and A. Bol, "Edge-preserving filtering of images with low photon counts," IEEE Trans. Pattern Anal. Mach. Intell. **30**(6), 1014–1151 (2008).

[20]J. K. Kamarainen, V. Kyrki, and H. Kalviainen, "Invariance properties of Gabor filter-based features–overview and applications," IEEE Trans. Image Process. **15**(5), 1088–1099 (2006).

[21]T. F. Chan and L. A. Vese, "Active contours without edges," IEEE Trans. Image Process. **10**(2), 266–277 (2001).

[22]S. Balla-Arabe, X. Gao, and B. Wang, "GPU accelerated edge-region based level set evolution constrained by 2D gray-scale histogram," IEEE Trans. Image Process. **22**(7), 2688–2698 (2013).

[23]Y. Chen, Z. Yan, and J. Shi, "Application of lattice Boltzmann method to image segmentation," in *2007 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (IEEE, Lyon, France, 2007), pp. 6562–6565.

[24]W. Xian and A. Takayuki, "Multi-GPU performance of incompressible flow computation by lattice Boltzmann method on GPU cluster," Parallel Comput. **37**(9), 521–535 (2011).

[25]W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," Med. Phys. **37**(9), 4902–4915 (2010).

[26]A. Le Maitre, W. P. Segars, S. Marache, A. Reilhac, M. Hatt, S. Tomei, C. Lartizien, and D. Visvikis, "Incorporating patient-specific variability in the simulation of realistic whole-body F-18-FDG distributions for oncology applications," Proc. IEEE **97**(12), 2026–2038 (2010).

[27]C. Liu, L. A. Pierce, A. M. Alessio, and P. E. Kinahan, "The impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging," Phys. Med. Biol. **54**(24), 7345–7362 (2009).

[28]P. Geramifar, M. S. Zafarghandi, P. Ghafarian, A. Rahmim, and M. R. Ay, "Respiratory-induced errors in tumor quantification and delineation in CT attenuation-corrected PET images: Effects of tumor size, tumor location, and respiratory trace: A simulation study using the 4D XCAT phantom," Mol. Imaging Biol. **15**(6), 655–665 (2013).

[29]S. A. Nehmeh, H. El-Zeftawy, C. Greco, J. Schwartz, Y. E. Erdi, A. Kirov, C. R. Schmidtlein, A. B. Gyau, S. M. Larson, and J. L. Humm, "An iterative technique to segment PET lesions using a Monte Carlo based mathematical model," Med. Phys. **36**(10), 4803–4809 (2009).

[30]F. Lamare, T. Cresson, J. Savean, C. C. Le Rest, A. J. Reader, and D. Visvikis, "Respiratory motion correction for PET oncology applications using affine transformation of list mode data," Phys. Med. Biol. **52**(1), 121–140 (2007).

[31]K. Thielemans, C. Tsoumpas, S. Mustafovic, T. Beisel, P. Aguiar, N. Dikaios, and M. W. Jacobson, "STIR: Software for tomographic image reconstruction release 2," Phys. Med. Biol. **57**(4), 867–883 (2012).

[32]J. F. Daisne, V. Gregoire, T. Duprez, M. Lonneux, M. Hamoir, H. Reychler, and V. Gregoire, "Tumor volume in pharyngolaryngeal squamous cell carcinoma: Comparison at CT, MR imaging, and FDG PET and validation with surgical specimen," Radiology **233**, 93–100 (2004).

[33]M. Wanet, J. A. Lee, B. Weynand, M. De Bast, A. Poncelet, V. Lacroix, E. Coche, V. Gregoire, and X. Geets, "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).

[34]M. Abdoli, R. A. J. O. Dierckx, and H. Zaidi, "Contourlet-based active contour model for PET image segmentation," Med. Phys. **40**(8), 082507 (12pp.) (2013).

[35]A. Schaefer, S. Kremp, D. Hellwig, C. Rübe, C. M. Kirsch, and U. Nestle, "A contrast-oriented algorithm for FDG-PET-based delineation of tumour volumes for the radiotherapy of lung cancer: Derivation from phantom measurements and validation in patient data," Eur. J. Nucl. Med. Mol. Imaging **35**(11), 1989–1999 (2008).

[36]L. R. Dice, "Measures of the amount of ecologic association between species," Ecology **26**(3), 297–302 (1945).

[37]K. H. Zou, S. K. Warfield, A. Bharatha, C. M. Tempany, M. R. Kaus, S. J. Haker, W. M. Wells III, F. A. Jolesz, and R. Kikinis, "Statistical validation of image segmentation quality based on a spatial overlap index," Acad. Radiol. **11**(2), 178–189 (2004).

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

[39]S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," IEEE Trans. Syst. Man. Cybern., Part B (Cybern.) **34**(4), 1907–1916 (2004).

[40]S. Belhassen and H. Zaidi, "A novel fuzzy C-means algorithm for unsupervised heterogeneous tumor quantification in PET," Med. Phys. **37**(3), 1309–1324 (2010).

[41]B. Sandberg, T. Chan, and L. Vese, "A level-set and Gabor-based active contour algorithm for segmenting textured images," UCLA Department of Mathematics CAM Report, 2002.

[42]Z. Xu, W. M. Allen, R. B. Baucom, B. K. Poulose, and B. A. Landman, "Texture analysis improves level set segmentation of the anterior abdominal wall," Med. Phys. **40**(12), 121901 (11pp.) (2013).

[43]H. Zaidi, M. Abdoli, C. L. Fuentes, and I. M. E. Naqa, "Comparative methods for PET image segmentation in pharyngolaryngeal squamous cell carcinoma," Eur. J. Nucl. Med. Mol. Imaging **39**, 881–891 (2012).

[44]A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, and A. C. Palmer, "Morphometric analysis of white matter lesions in MR images: Method and validation," IEEE Trans. Med. Imaging **13**(4), 716–724 (1994).

[45]J. Yang, B. M. Beadle, A. S. Garden, D. L. Schwartz, and M. Aristophanous, "A multimodality segmentation framework for automatic target delineation in head and neck radiotherapy," Med. Phys. **42**(9), 5310–5320 (2015).

[46]M. Aristophanous, B. C. Penney, M. K. Martel, and C. A. Pelizzari, "A Gaussian mixture model for definition of lung tumor volumes in positron emission tomography," Med. Phys. **34**(11), 4223–4235 (2007).

[47]R. J. McGurk, J. Bowsher, J. A. Lee, and S. K. Das, "Combining multiple FDG-PET radiotherapy target segmentation methods to reduce the effect of variable performance of individual segmentation methods," Med. Phys. **40**(4), 042501 (9pp.) (2013).

[48]S. Lankton and A. Tannenbaum, "Localizing region-based active contours," IEEE Trans. Image Process. **17**(11), 2029–2039 (2008).

[49]C. Li, C. Y. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," IEEE Trans. Image Process. **17**(10), 1940–1949 (2008).

[50]M. Soufi, A. Kamali-Asl, P. Geramifar, M. Abdoli, and A. Rahmim, "Combined fuzzy logic and random walker algorithm for PET image tumor delineation," Nucl. Med. Commun. **37**(2), 171–181 (2016).

[51]C. J. Tao, J. L. Yi, N. Y. Chen, W. Ren, J. Cheng, S. Tung, L. Kong, S. J. Lin, J. J. Pan, G. S. Zhang, J. Hu, Z. Y. Qi, J. Ma, J. D. Lu, D. Yan, and Y. Sun, "Multi-subject atlas-based auto-segmentation reduces interobserver variation and improves dosimetric parameter consistency for organs at risk in nasopharyngeal carcinoma: A multi-institution clinical study," Radiother. Oncol. **115**(3), 407–411 (2015).

[52]V. Carillo, C. Cozzarini, L. Perna, M. Calandra, S. Gianolini, T. Rancati, A. E. Spinelli, V. Vavassori, S. Villa, R. Valdagni, and C. Fiorino, "Contouring variability of the penile bulb on CT images: Quantitative assessment using a generalized concordance index," Int. J. Radiat. Oncol., Biol., Phys. **84**(3), 841–846 (2012).