33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

# The Impact of Reconstruction Algorithms on Semi-automatic Small Lesion Segmentation for PET: A Phantom Study

Cherry Ballangan, Chung Chan, Xiuying Wang, David Dagan Feng, *Fellow, IEEE*

*Abstract*— **A robust lesion segmentation method is critical for quantification of lesion activity in positron emission tomography (PET), especially for the cases where lesion boundary is not discernible in the corresponding computed tomography (CT). However, lesion delineation in PET is a challenging task, especially for small lesions, due to the low intrinsic resolution, image noise and partial volume effect. The combinations of different reconstruction methods and post-reconstruction smoothing on PET images also affect the segmentation result significantly which has always been overlooked. Therefore, the aim of this study was to investigate the impact of different reconstruction methods on semi-automated small lesion segmentation for PET images. Four conventional segmentation methods were evaluated including region growing technique based on maximum intensity (RGmax) and mean intensity (RGmean) thresholds, Fuzzy c-mean (FCM) and watershed (WS) technique. All these methods were evaluated on a physical phantom scan which was reconstructed with Ordered Subset Expectation Maximization (OSEM) with Gaussian post-smoothing and Maximum *a Posteriori* (MAP) with quadratic prior respectively. The results demonstrate that: 1) the performance of all the segmentation methods subject to the smoothness constraint applied on the reconstructed images; 2) FCM method applied on MAP reconstructed images yielded overall superior performance than other evaluated combinations.**

## I. Introduction

TUMOR segmentation in positron emission tomography (PET) plays an important role in oncology diagnosis, therapy planning and disease monitoring. However, defining an accurate boundary for lesions in PET is a challenging task due to the low intrinsic resolution, noise corruption and partial volume effect. This situation is even more challenging for small lesions which very often do not have matched boundaries in the co-registered anatomical images, such as computed tomography (CT).

This work was supported by ARC and PolyU grants.

C. Ballangan and X. Wang are with the Biomedical and Multimedia Information Technology (BMIT) Research Group, School of Information Technologies, The University of Sydney, Australia. (e-mail: cherry@it.usyd.edu.au, xiuying@it.usyd.edu.au)

C. Chan was with the Faculty of Health Sciences, Brain and Mind Research Institute, and BMIT research group, University of Sydney, Australia. He is now with the Department of Diagnostic Radiology, Yale University, USA. (e-mail: janc@it.usyd.edu.au)

David Dagan Feng is with the BMIT Research Group, School of IT, University of Sydney, Australia; the Center for Multimedia Signal Processing (CMSP), Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong; and the Med-X Research Institute, Shanghai Jiao Tong University, China (e-mail: feng@it.usyd.edu.au).

Currently, the most commonly used method is based on a thresholding approach that can be implemented using a region growing technique. The threshold values are usually obtained by setting a certain percentage of the maximum intensity [1], or can also be obtained iteratively based on the mean intensity of the initial lesion region [2]. Other lesion segmentation methods include gradient-based approaches such as watershed technique [3] and statistical approaches such as Fuzzy c-means (FCM) clustering [4].

Although some studies have compared different segmentation approaches [5-7], no segmentation methods have been demonstrated to yield robust and consistent performance on PET. Most of these algorithms are system specific and their application to different system settings can result in significantly different segmented volumes [5]. Therefore, the variability of reconstruction methods and post-smoothing filters on PET images may affect the performance of the segmentation methods.

In this paper, we investigate the effect of different reconstruction methods and post-reconstruction smoothing on the lesion segmentation for PET images. We focus on small lesion especially as it is more susceptible to the partial volume effect and its segmentation results were often disappointing [6-8].

## II. Materials and Methods

### A. Phantom Datasets

The methods were evaluated using a physical torso phantom (Anthropomorphic Torso phantom™, Data Spectrum Inc, Hillsborough NC) study (Fig 1a). CT and fluorodeoxyglucose ($^{18}$FDG) PET scans of the torso phantom were acquired on a Siemens Biograph Truepoint PET/CT scanner (Siemens Healthcare Molecular Imaging, Knoxville TN). The fillable volumes of the phantom and the lung inserts, which contained Styrofoam beads, were 10L and 0.7L, respectively. Three 4.5mL plastic tubes approximately 10 mm in diameter and 50 mm in length were placed in locations corresponding to the lesions in the computer simulation. The phantom contained a total of 213 MBq of $^{18}$FDG. The $^{18}$FDG concentrations in the lesions and lung tissue were 4 and 0.5 times the soft tissue concentration, respectively. Note the lesions residing in the soft tissue region were not detected in the CT image (Fig. 1(c)), where only lesion 1 can be observed in the lung region. To reveal the true lesion locations, the 10th plane of the physical

phantom is shown in Fig 1(b) where the caps of the tubes used to simulate lesions can be observed. The emission data were acquired for 4 minutes. The sinograms comprised 168 radial samples and 168 angular samples. For the PET scan, the true and random event rates were approximately 480 kcps and 175 kcps, respectively. The acquired sinograms were normalized, corrected for attenuation and scatter, and rebinned into 2D parallel projections using Fourier rebinning.

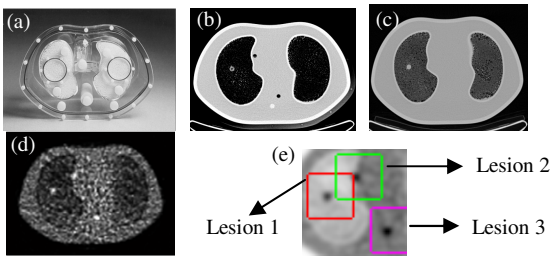



Fig. 1. (a) The Anthropomorphic Torso Phantom™; (b) The 10th plane is shown here to reveal the exact lesion locations. The caps of the containers that were used to simulate lesions are shown in this image. (c) The 26th slice of the CT image of the physical phantom; (d) The corresponding PET image reconstructed with OSEM at 25 iterations and 4 level of subsets; (e) Three volumes of interest denoted by red square for lesion 1, green square for lesion 2 and magenta square for lesion 3 drawn on a post-smoothed image (inverse gray scale).

Two reconstruction algorithms were evaluated, the Ordered Subset Expectation Maximization (OSEM) approach with 25 iterations and 4 subsets [9]-[10] and Maximum A Posteriori (MAP) [11] with quadratic prior. All the images reconstructed with OSEM were post-filtered with a 3D Gaussian filter to suppress noise.

### B. Segmentation Methods

We applied four segmentation methods, denoted RGmax, RGmean, FCM and WS to a manually defined volume of interest that covers each lesion and provides sufficient background voxels (Fig. 1(e)). We optimized the parameters used in each approach to reduce bias caused by the parameter settings.

RGmax is a threshold approach with the threshold value $T$ calculated based on a fixed percentage $\alpha$ of the maximum activity $I_{max}$ within the lesion:

$$T = \alpha \times I_{\max} \tag{1}$$

We implemented this threshold approach using a region growing technique with a seed point located within the lesion. We optimized the fixed percentage value $\alpha$ using a leave-one-out procedure as in [7].

RGmean is an adaptive threshold approach which relies only on the mean intensity of the lesion [2]. The threshold value is obtained iteratively using:

$$Th_n = \beta \times M(Th_{n-1}) + \gamma \tag{2}$$

where $Th_n$ is the threshold value at iteration $n$, and $M(Th_{n-1})$ is the mean activity value of the lesion volume obtained with threshold $Th_{n-1}$. We optimized the parameters $\beta$ and $\gamma$ by using regression analysis and reproducing the experiments in [2].

FCM is a Fuzzy c-means clustering method which is based on the minimization of an objective function $O$:

$$O = \sum_{i=1}^{D} \sum_{j=1}^{C} m_{ij}^{n} \left| x_i - c_j \right| \tag{3}$$

where $m_{ij}$ is the degree of membership of data $x_i$ in a cluster $j$, $n$ is usually set as 2, $c_j$ is the center of cluster $j$, $D$ is the number of data to be clustered and $C$ is the number of clusters. In our implementation, we optimized the number of clusters by maximizing the Bayesian Information Criterion (BIC) [12]:

$$BIC(C_k) = \left( \sum_{i=1}^{k} -0.5\, n_i \log |\Sigma_i| \right) - \lambda P \log N \tag{4}$$

where $k$ is the number of cluster, $\Sigma_i$ is the covariance matrix of the samples in cluster $c$ with $n_i$ number of samples, $P$ is the number of parameters of each cluster and $N$ is the total number of samples to be clustered. BIC has been used to optimize the number of clusters for PET segmentation in [13].

WS is a watershed technique which regards the intensity of an image as the height of the topology. It is usually applied on an image gradient such that the homogeneous regions have similar height, separated by the higher gradient magnitude set as the watershed. We used marker-based watershed to avoid over-segmentation usually found in watershed approach. In our implementation, we used the gradient magnitude of the PET image and located a marker inside a lesion (maximum intensity) and another marker in the background region (minimum intensity) and grew a region for each marker. The lesion boundary was then formed in the meeting line of the two markers' regions.

### C. Similarity Measure

For quantitative analysis, we used the normalized volume error (NVE) as the similarity measure, which is defined as:

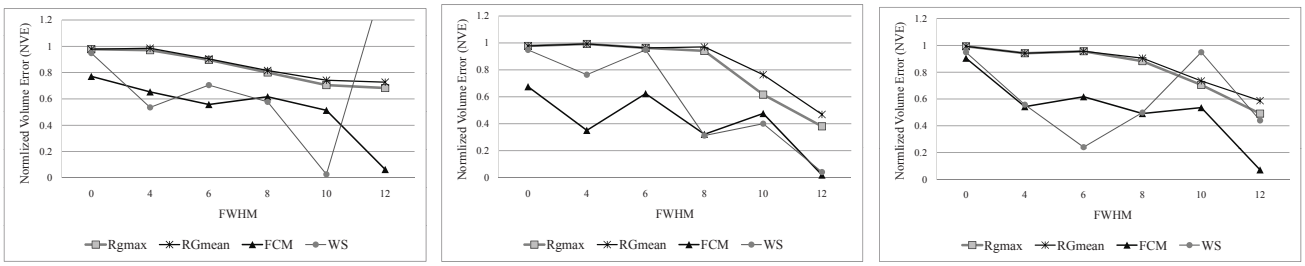
$$NVE = \frac{|V_{SEG} - V_{GT}|}{V_{GT}}. \tag{5}$$

where $V_{SEG}$ is the volume of the lesion obtained from each of the segmentation approach and $V_{GT}$ is the volume of the ground truth.

## III. Results and Discussion

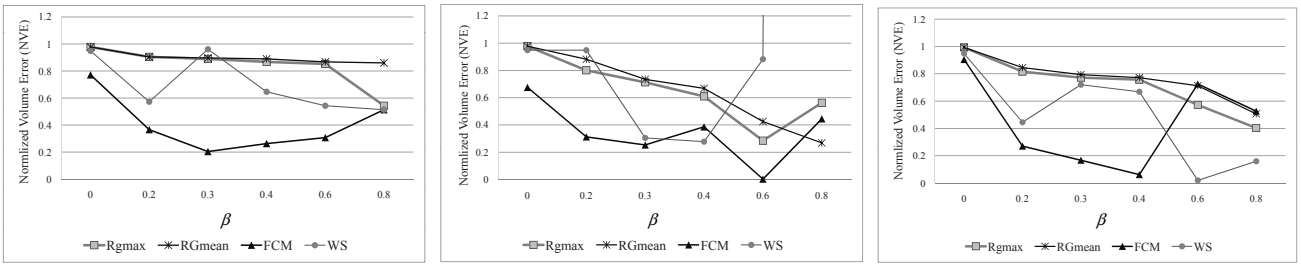
Fig. 2 shows the NVEs of the three lesions (i-iii) using different segmentation methods on two reconstruction approaches with varying smoothing parameters. Fig. 3(a-b) shows the average NVEs of all lesions for each reconstruction algorithm with varying smoothing parameters. In Fig 3(c), the total NVE of all lesion locations and all smoothing parameters for each reconstruction method is presented, which shows the performance of these segmentation algorithms on different reconstruction methods.

In general RGmax and RGmean produced relatively consistent segmented volumes for all lesions (different background activities) and across different reconstruction methods. RGmax produced an NVE of 0.78 ± 0.20 and RGmean produced a slightly higher NVE of 0.82 ± 0.18. In average, both segmentation methods showed the same trends of decreasing NVE when more smoothing strength was applied in both OSEM and MAP reconstructions (Fig. 3(a-b)). Both methods produced relatively smaller segmented volumes compared to the ground truth volumes (4.5 mL), and when stronger smoothness were applied, the size of the segmented volumes increased and approached to the size of the ground truth and thus the NVEs decreased. This trend was consistent in every lesion and reconstruction methods, except for lesion 2 with MAP reconstruction (Fig. 2(b)(ii)), where the segmented volume was close to 4.5 mL when $\beta$ = 0.6 (and thus the NVE was low) and increased to more than 7 mL when $\beta$ = 0.8 (and thus the NVE increased) due to the reduced lesion-background contrast.




(i) Lesion 1 (ii) Lesion 2 (iii) Lesion 3

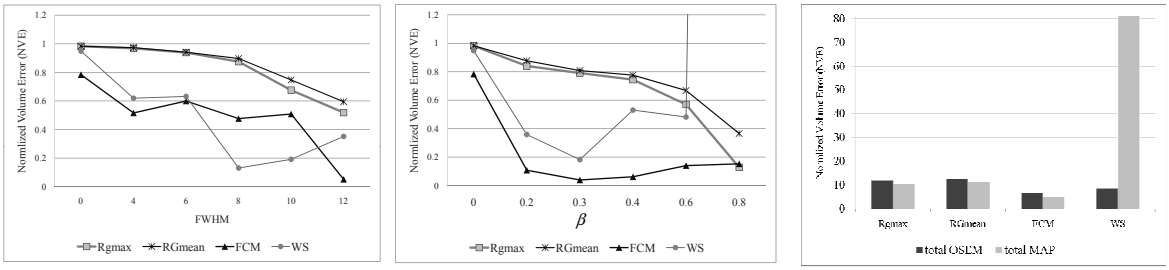
(a) OSEM reconstruction with varying FWHMs




(i) Lesion 1 (ii) Lesion 2 (iii) Lesion 3

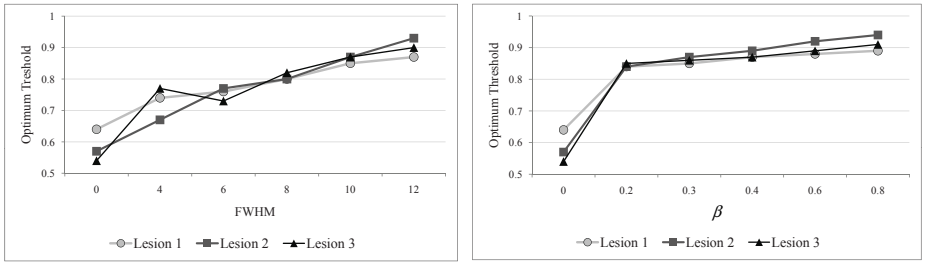
(b) MAP reconstruction with varying $\beta$s

Fig. 2. NVE of four segmentation methods applied on the images reconstructed with (a) OSEM and (b) MAP for all the lesions.




(a) OSEM with varying FWHMs (b) MAP with varying $\beta$s (c) Total NVE of OSEM and MAP

Fig.3. Average NVE of the three lesions using four different segmentation methods over different post-smoothing parameters in OSEM (a) and MAP (b) reconstructions; and total NVEs of the three lesions and of all smoothing parameters in OSEM and MAP (c).




(a) OSEM with varying FWHMs (b) MAP with varying $\beta$s

Fig.4. Increasing trend of optimum threshold (as a fraction of maximum intensity) over different smoothing parameters in OSEM (a) and MAP (b)

However, both region growing methods were sensitive to segmentation parameters. As described in Section II, we have optimized these parameters using the training datasets. When these parameters were changed, the volumes may change significantly. For example, Fig. 4 shows that the smoother the image, the higher the optimum threshold. Thus, applying lower threshold value to a smooth image may result in leakage. To avoid this leakage, the optimized parameter over all training data for RGmax was relatively high and therefore it resulted in smaller volumes.

FCM yielded the least NVE of 0.44 ± 0.25 among all the methods. However, it performed differently with the reconstruction methods. It can be observed that for the post-smoothed OSEM images, the best result was obtained on the images post-filtered with the strongest smoothness such as FWHM=12mm. In contrast, the NVE reached a minimum on the images reconstructed with MAP at moderate smoothing values and then increased again.

WS produced the highest NVE of 2.66 ± 12.23 among all the methods because of one leakage case in MAP reconstruction at $\beta$ = 0.8 for lesion 2. When this case was removed, the NVE was 0.63 ± 0.34. The leakage occurred as the lesion was located on the boundary of the lung and the mediastinum. Due to the smoothing effect, the gradient magnitude of the lesion was overshadowed by the gradient magnitude of the lung boundary (Fig. 5(b)). This caused the WS adapted to the more pronounced gradient of the lung boundary instead (Fig. 6(d)). Unlike the other segmentation methods which generally have decreasing NVE with increasing smoothness, the NVE of WS is very inconsistent.

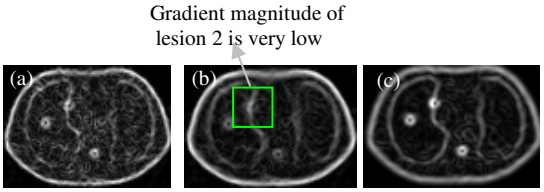



Fig. 5. Gradient image of MAP reconstruction with $\beta$ = 0.2 (a) and $\beta$ = 0.8 (b) and OSEM reconstruction with FWHM 12mm (c) to show that compared to other reconstruction settings, MAP reconstruction with $\beta$ = 0.8 resulted in low gradient magnitude of lesion 2.

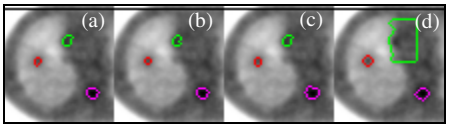



Fig. 6. Screen shots of the segmentation results in MAP reconstruction, $\beta$=0.8, using: (a) RGmax, (b) RGmean, (c) FCM, and (d) WS. Lesion 1, lesion 2 and lesion 3 are delineated with red, green and magenta color respectively.

In general, MAP reconstructions yielded segmented volumes with less NVE compared to OSEM reconstructions for all segmentation methods, except the WS (Fig. 3(c)) due to the low gradient magnitude produced by MAP.

## IV. Conclusion

In this study, the accuracy of four different segmentation methods: RGmax, RGmean, FCM and WS applied on the images reconstructed with two different reconstructions were analyzed. The evaluations on a physical phantom study demonstrate that RGmax and RGmean are relatively insensitive to the reconstruction methods, although the parameters should be carefully optimized according to the training datasets. FCM yielded the best performance among all the methods, especially on MAP reconstructions. The conventional WS may be unsuitable for PET images due to the low contrast subject and high correlation between pixels introduced by smoothing. We plan to evaluate more reconstruction methods including MAP with edge preserving priors and anatomical priors in future.

## References


[1] K. J. Biehl, F.-M. Kong, F. Dehdashti, J.-Y. Jin, S. Mutic, I. El Naqa, B. A. Siegel, and J. D. Bradley, "18F-FDG PET Definition of Gross Tumor Volume for Radiotherapy of Non-Small Cell Lung Cancer: Is a Single Standardized Uptake Value Threshold Approach Appropriate?," *J Nucl Med,* vol. 47, pp. 1808-1812, November 2006.

[2] Q. C. Black, I. S. Grills, L. L. Kestin, C.-Y. O. Wong, J. W. Wong, A. A. Martinez, and D. Yan, "Defining A Radiotherapy Target with Positron Emission Tomography," *Int. J. Radiation Oncology Biol. Phys.,* vol. 60, pp. 1272-1282, November 2004.

[3] X. Geets, J. Lee, A. Bol, M. Lonneux, and V. Grégoire, "A gradient-based method for segmenting FDG-PET images: methodology and validation," *Eur J Nucl Med Mol Imaging,* vol. 34, pp. 1427-1438, September 2007.

[4] S. Belhassen and H. Zaidi, "A novel fuzzy C-means algorithm for unsupervised heterogeneous tumor quantification in PET," *Med. Phys.,* vol. 37, pp. 1309-1324, March 2010 2010.

[5] U. Nestle, S. Kremp, A. Schaefer-Schuler, C. Sebastian-Welsch, D. Hellwig, C. Rube, and C.-M. Kirsch, "Comparison of Different Methods for Delineation of F-FDG PET-Positive Tissue for Target Volume Definition in Radiotherapy of Patients with Non-Small Cell Lung Cancer," *J Nucl Med,* vol. 46, pp. 1342-1348, August 2005.

[6] L. Drever, W. Roa, A. McEwan, and D. Robinson, "Comparison of three image segmentation techniques for target volume delineation in positron emission tomography," *J Appl Clin Med Phys,* vol. 8, pp. 93-109, 2007.

[7] P. Tylski, S. Stute, N. Grotus, K. Doyeux, S. Hapdey, I. Gardin, B. Vanderlinden, and I. Buvat, "Comparative Assessment of Methods for Estimating Tumor Volume and Standardized Uptake Value in 18F-FDG PET," *J Nucl Med,* vol. 51, pp. 268-276, February 1, 2010 2010.

[8] M. Soret, S. L. Bacharach, and I. Buvat, "Partial-Volume Effect in PET Tumor Imaging," *J Nucl Med,* p. jnumed.106.035774, May 15, 2007 2007.

[9] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imag.,* vol. 13, pp. 601-609, 1994.

[10] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imaging,* vol. MI-1, pp. 113-122, 1982.

[11] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imag.,* vol. 9, pp. 84-93, 1990.

[12] S. Chen and P. S. Gopalakrishnan, "Speaker, Environment and Channel Change Detection and Clustering via the Bayesian Information Criterion," in *Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop*, 1998.

[13] D. W. G. Montgomery, A. Amira, and H. Zaidi, "Fully automated segmentation of oncological PET volumes using a combined multiscale and statistical model," *Med. Phys.,* vol. 34, pp. 722-736, February 2007.