Published in final edited form as:

Eur J Radiol. 2012 November ; 81(11): 3535–3541. doi:10.1016/j.ejrad.2012.01.001.

Delineation of FDG-PET tumors from heterogeneous background using spectral clustering

Fei Yang^{a,*} and Perry W. Grigsby^{b,1}

^aDepartment of Radiation Oncology, Washington University School of Medicine, 4921 Parkview Place, St. Louis, MO 63110, United States

^bDepartment of Radiation Oncology, Alvin J. Siteman Cancer Center, Washington University School of Medicine, 4921 Parkview Place, St. Louis, MO 63110, United States

Abstract

This paper explored the feasibility of using spectral clustering to segment FDG-PET tumor in the presence of heterogeneous background. Spectral clustering refers to a class of clustering methods which employ the eigenstructure of a similarity matrix to partition image voxels into disjoint clusters. The similarity between two voxels was measured with the intensity distance scaled by voxel-varying factors capturing local statistics and the number of clusters was inferred based on rotating the eigenvector matrix for the maximally sparse representation. Metrics used to evaluate the segmentation accuracy included: Dice coefficient, Jaccard coefficient, false positive dice, false negative dice, symmetric mean absolute surface distance, and absolute volumetric difference. Comparison of segmentation results between the presented method and the adaptive thresholding method on the simulated PET data shows the former attains an overall better detection accuracy. Applying the presented method on patient data gave segmentation results in fairly good agreement with physician manual annotations. These results indicate that the presented method have the potential to accurately delineate complex shaped FDG-PET tumors containing inhomogeneous activities in the presence of heterogeneous background.

Keywords

Spectral clustering; Tumor segmentation; Heterogeneous background; Digital phantom; FDG-PET

1. Introduction

The use of ¹⁸F-fluoro-deoxy-2-glucose (¹⁸F-FDG) positron emission tomography (PET) has been gaining acceptance in radiation oncology, fostered by growing and emerging evidence that ¹⁸F-FDG PET offers high sensitivity and specificity for establishing the diagnosis of a number of malignancies. Although it is widely used in staging, detecting primary or recurrent cancer, and monitoring therapeutic response [1], its applications in radiation therapy are limited due to the difficulty in accurately and precisely delineating tumors because the PET tumor boundaries suffer from PET's poor spatial resolution and the uneven uptake of ¹⁸F-FDG in tumors [2].

© 2012 Elsevier Ireland Ltd. All rights reserved.

*Corresponding author. Tel.: +1 314 362 6074; fax: +1 314 362 6074. fei.yang@wustl.edu (F. Yang), pgrigsby@radonc.wustl.edu (P.W. Grigsby).

¹Tel.: +1 314 747 7236.

Conflict of interest

No conflict of interest.

To date, various techniques for ^{18}F -FDG PET-based target-volume delineation have been developed. Manual contouring is the most commonly used method. The segmented volume is sensitive to the window and level settings on the image presentation and is also subject to intraobserver and interobserver variations [3]. Thresholding using either fixed standardized uptake values [4,5] or preselected percentages of the maximum activity [6] were proposed. However, the clinical application of these methods is limited due to the complex shape and heterogeneity of actual tumors. As shown by Gregoire et al. [3] using laryngectomy data, the threshold needed to adequately fit the true tumor volume revealed on histological examination ranged from 36% to 73% and the defined target volume was unpredictable when trying to apply a fixed threshold. Thresholding that takes into account the background activity were also suggested [7–9]. However, threshold values used in methods of this type need to be calibrated for each scanner and reconstruction, and a background region needs to be manually defined when applying these methods on patient scans. A number of non-thresholding approaches [10–13] have been proposed to address PET tumor segmentation. As pointed out by Hatt et al. [14], however, most of these approaches often depend on pre- or postprocessing steps like deconvolution and denoising, and were validated on geometric phantom acquisitions. Recently, Hatt et al. [14] extended the fuzzy locally adaptive Bayesian (FLAB) method to allow for characterization of heterogeneous tumors with the use of an additional third classification class. However, using a total of three classes may limit the utility of this method from segmenting tumors in the presence of heterogeneous background, scenarios commonly encountered in clinical practice. Accurate tumor definition in PET images remains a task that needs to be improved.

This work attempted to approach this problem by using spectral clustering [15,16]. It views the image voxels as nodes of a connected graph and clusters are derived by partitioning this graph, based on its spectral decomposition, into subgraphs with nodes in the same cluster having high similarity and nodes in different clusters having low similarity. We compared spectral clustering and adaptive thresholding, one of those most commonly used in clinical practice, with respect to their segmentation accuracy on the simulated PET data. We also applied the spectral clustering method on patient data and compared its segmentation results with the manually labeled results. The remainder of this paper is organized as follows. In Section 2, we first present our implementation of the spectral clustering algorithm, then describe the adaptive thresholding method used for comparison, the generation of the simulated PET data, and the acquisition of patient scans along with details of the quantitative measures used for segmentation evaluation. The segmentation results obtained from spectral clustering are presented and compared with those offered by adaptive thresholding on the simulated data and those manually annotated for patient scans in Section 3, followed by a discussion of this work in Section 4. Finally, Section 5 concludes the paper.

2. Methods

2.1. Spectral clustering with local scaling

Instead of detecting of a lesion in the whole PET image, the goal of this study is to delineate its volume and shape. So it is assumed that a 3D “box” encompassing the tumor has been identified for each PET scan. The image voxels within the “box” are then modeled as a weighted, undirected graph $G(V, E)$ with each vertex (V) in this graph representing a voxel. Two vertices are connected by an edge (E) whose weight captures the similarity between them. In this study, the weight $w_{i,j}$ of the edge between vertex i and vertex j was defined as the brightness similarity between their corresponding voxels such as:

$$w_{ij} = \begin{cases} e^{-\frac{\|I(i)-I(j)\|_2^2}{\sigma_i \sigma_j}} & \text{if } \|X(i)-X(j)\|_2 < R \text{ and } i \neq j, \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $I(i)$ is the image brightness of the voxel corresponding to vertex i ; $X(i)$ is the spatial locations of the voxel corresponding to vertex i ; σ_i , the local scaling parameter, is defined as the absolute difference between $I(i)$ and the mean brightness of the n nearest neighbors of the voxel corresponding to vertex i . $I(j)$, $X(j)$, and σ_j in Eq. (1) were similarly defined for vertex j . For two vertices with their corresponding voxels separated by a large distance (greater than R), their similarity is set to 0 explicitly considering nearby voxels are much more likely than faraway voxels to fall into a same cluster. Thus for an image with N voxels the weight matrix W associated with its weighted graph $G(V, E)$ is an $N \times N$ symmetric matrix with $W(i, j) = w_{ij}$. Let D be a diagonal matrix with $D(i, i) = \sum_j W(i, j)$ and the normalized affinity matrix L can then be formed such that $L = D^{-1/2} W D^{-1/2}$. After finding the K top eigenvectors of L , the optimal number of clusters κ was estimated based on rotating the eigenvector matrix for the maximally sparse representation [17]. Upon determining the number of clusters, the k -means technique was initialized clustering on the rows of the matrix formed with the κ top rotated eigenvectors, thereby partitioning the original image into κ segments.

Having used a local scaling scheme for similarity measure and inferred the number of clusters automatically, we are left with three parameters: R , n , and K . R is the radius defining the extent of the neighborhood within which the pairwise similarity between the center voxel and each of its surrounding neighbors is computed. The smaller R , the less pairwise similarity information will be provided for global segmentation, while more computation resource is required for larger radii. In view of this tradeoff, the neighborhood was empirically set as a sphere with 6 voxel units in radius. The selection of n is independent of the intensity scale and its value was set for obtaining better overall results to 80 (equivalent to a sphere of radius 2.5 voxel units centered on the image voxel of interest). K , the largest possible cluster number, can take values up to N . However, large values of K can be computationally expensive in searching for eigenvectors and even wasteful given there exists limited number of clusters within an image. In this study, K was set to be 25 for all segmentation tasks, which can be considered as a sufficiently large estimate for the number of clusters.

2.2. Adaptive thresholding used for comparison

With known ground truths, we using the simulated PET data compared the segmentation performance between the adaptive thresholding method and the spectral clustering method. Adaptive thresholding may be considered the current state-of-the-art algorithm for target segmentation in PET images [14]. The threshold (I_{thres}) it uses is defined as:

$$I_{thres} = \alpha \times I_{mean} + I_{bgd}, \quad (2)$$

where I_{mean} is the mean activity of all voxels contained inside an initial threshold at 70% of the maximum, and I_{bgd} is the background activity which was computed as the mean intensity of a region of interest (ROI) manually defined on the background. I_{mean} and I_{bgd} were subsequently used to derive a first approximation of the source-to-background ratio. The parameter α was optimized using phantom acquisitions on the simulated scanner. The target was defined in a region-growing manner with the maximum intensity voxel as a seed and iteratively adding neighboring voxels if their value was above the calculated threshold.

2.3. PET simulation

As no well characterized and reliable test PET data is available at this point, in this study a digital PET phantom [18] was adopted as a source to generate PET tumor images with various shapes and activity distributions. The simulation was carried out by feeding the Zubal phantom as an anatomical input to the Sim-SET (the Simulation System for Emission Tomography) software, which employs Monte Carlo techniques to model the underlying physical processes and imaging system to produce simulated emission tomography data. The PET system modeled was a Reveal HD scanner using a continuous cylindrical BGO detector with a 40 cm radius. The radioisotope simulated was F-18. Using single slice rebinning (SSRB) the data were binned into 128×128 sinograms. Upon obtaining the projection data, the ordered subset expectation maximization (OSEM) algorithm for eight iterations with four subsets was used to reconstruct each slice separately and a post-reconstruction 5 mm full width at half maximum (FWHM) 3D Gaussian filter was applied to the resulting image slices. Validation of the simulation approach was demonstrated by the congruence between the simulated and the clinical images with respect to intensity profile, intensity appearance, and image texture [18]. A total of 30 lesions in the lung were simulated in this study. Half of them ($S1-S15$) were sphere-shaped with a homogeneous or heterogeneous activity distribution, while the other half ($S16-S30$) was irregularly shaped with intratumoral activity heterogeneity. Volumes of these lesions ranged from 5.2 mL to 345.1 mL (mean $\pm$ SD = 90.3 ± 70.9 mL).

2.4. Patient data for validation

The patient dataset consisted of 15 FDG-PET tumor images which had been labeled manually by radiation oncologists. Lesion sites ranged from neck, lung, and cervix to uterus. The images were reconstructed using OSEM (2 iterations, 8 subsets) with a voxel size of $5.31 \text{ mm} \times 5.31 \text{ mm} \times 3.38 \text{ mm}$ or OSEM (4 iterations, 8 subset) with a voxel size of $4.07 \text{ mm} \times 4.07 \text{ mm} \times 5 \text{ mm}$. The patient data analysis was performed after approval by the institutional review board.

2.5. Quantitative metrics for evaluation

A number of volumetric and distance-based metrics commonly used in evaluation of segmentation performance were computed for each segmentation result. The known ground truths and the manual annotations were used as gold standards for the simulated data and the patient data, respectively. In the following description of these metrics, the segmentation result is termed S and the gold standard T .

2.5.1. Dice coefficient—The Dice coefficient measures the extent of spatial overlap between the segmented results and the gold standard and is defined as:

$$Dice = 2 \times \frac{|S \cap T|}{|S| + |T|}. \quad (3)$$

2.5.2. Jaccard coefficient—The Jaccard coefficient is defined as the size of the overlap of S and T divided by the size of their union:

$$Jaccard = \frac{|S \cap T|}{|S \cup T|}. \quad (4)$$

In contrast to the Dice coefficient which gives twice the weight to matches, the Jaccard coefficient gives equal weight to matches and nonmatches.

2.5.3. False positive dice and false negative dice—To further characterize the segmentation results, the false positive dice (FPD) and the false negative dice (FND) were calculated as follows:

$$FPD = 2 \times \frac{|S \cap \bar{T}|}{|S| + |T|}, \quad (5)$$

$$FND = 2 \times \frac{|\bar{S} \cap T|}{|S| + |T|}, \quad (6)$$

where $\bar{S}$ and $\bar{T}$ are respectively the complements of the segmentation results and the gold standard with respect to the identified bounding “box” for a given tumor. The *FPD* gives a measure of over-segmentation and the *FND* of under-segmentation.

2.5.4. Symmetric mean absolute surface distance—The symmetric mean absolute surface distance (*SMASD*) attempts to estimate the average extent to which the surfaces of S and T differ [19] and is defined as below:

$$SMASD = \frac{1}{n_s + n_T} \left(\sum_{i=1}^{n_s} |d_i^{ST}| + \sum_{j=1}^{n_T} |d_j^{TS}| \right), \quad (7)$$

where n_s and n_T are respectively the number of the surface voxels on S and T ; d_i^{ST} is the distance to the closet voxel on T for the i th surface voxel on S and d_j^{TS} is similarly defined as the distance to the closet voxel on S for the j th surface voxel on T .

2.5.5. Absolute volumetric difference—The absolute volumetric difference (AVD) calculates the absolute difference between the segmented volume (V_S) and the gold standard volume (V_T) as a percentage of the gold standard volume:

$$AVD = \frac{|V_S - V_T|}{V_T} \times 100. \quad (8)$$

3. Results

3.1. Comparison with adaptive thresholding on simulated data

Taken as examples, the segmentations obtained for lesion *S22*, *S25*, *S29*, and *S30* by both algorithms at the central axial slice are presented in Fig. 1. The ground truth and the simulated image for each lesion are shown in the first two columns respectively. Results from spectral clustering, as presented in the third column, show segmentation using spectral clustering resulted in a partition of the original image into a set of disjoint regions, in contrast to the strictly binary segmentations obtained using adaptive thresholding as presented in the last column. Fig. 2 presents the results of applying the aforementioned accuracy evaluation metrics to the segmentation results obtained using both methods and Fig. 3 contains the mean and standard deviation of these metrics for both methods. Comparison of all metric results between adaptive thresholding and spectral clustering shows that the latter achieves an overall higher detection accuracy. In several scenarios, adaptive thresholding was not able to produce satisfactory segmentation. It tends to oversegment lesions adjacent to other high-uptake regions. As seen in Fig. 1d for *S25*, it incorporated the abutting high-uptake heart region into the segmentation map and resulted in

a *FPD* of 0.5 and a *AVD* of 0.6 for this case. For some lesions with strong intratumoral heterogeneity, segmentation obtained using adaptive thresholding shows a tendency towards underestimation (e.g., *S22* as shown in Fig. 1d, row 1) because only the high-activity region is retained using adaptive thresholding. While on the other hand spectral clustering was able to generate segmentation maps closer to the ground truth for all these cases.

3.2. Accuracy with respect to manual annotations on patient data

The segmentations obtained on four patient data at the central axial slice are presented in Fig. 4 as examples. PET image and the manual annotation for each tumor are shown in the first two columns respectively. Segmentation results from the spectral clustering method, presented in the last column, show the presented method is capable of delineating complex-shaped clinical lesions from highly heterogeneous background. Fig. 5 presents the mean and standard deviation of the accuracy evaluation metrics when applied to the segment results obtained from spectral clustering using the manual labels as gold standards. *Dice* with a mean of 0.85 and a standard deviation of 0.05 indicates that there is fairly good spatial overlap between the manual annotations and the segmentation results obtained using spectral clustering, judging by the criterion that in image validation a good overlap occurs when *Dice* reaches 0.75 [20]. *FPD* with a mean of 0.26 and standard deviation of 0.08 implies the lesions are oversegmented, with respect to physician annotations, by spectral clustering. Overall, compared with the performance on the simulated data, spectral clustering exhibits greater performance variation on patient data. There are various factors that may attribute to this augmented variation. In addition to the difference in reconstruction algorithm, voxel size, tumor site and other factors concomitant with imaging, variability of the manual labeling, especially between raters, may play a large part in elevating the performance variation of spectral clustering appearing in the patient data.

4. Discussion

Largely attributed to PET's unique ability to provide functional information, incorporation of PET as an adjuvant image modality to complement the structural image modalities has been continuously gaining ground in radiation therapy. With this essential characteristic, PET may have the potential to assist in planning more effective radiotherapy treatments. Nevertheless, the difficulties associated with accurate target volume delineation within PET pose a significant obstacle to the implementation of this initiative. Although a number of automated algorithms have been proposed, the target delineation of PET images in clinical practice relies upon either the manual contouring by radiation oncologists or threshold-based segmentation such as fixed thresholding or adaptive thresholding. Besides time consuming and laborious, the former is subject to inter-rater and intra-rater variation as mentioned previously. As for threshold-based methods, it has shown that they are able to accurately define target volumes on geometric phantom acquisition. As evidenced by histological examination, however, no universal agreement exists concerning the appropriate threshold level (or range of levels) required to accurately outline PET target volumes. Accurate radiation therapy target definition within PET remains a task that needs to be improved. The objective of this study was to investigate the possibility of using spectral clustering for PET tumor segmentation, especially in the presence of heterogeneous background.

Spectral clustering was compared with adaptive thresholding with respect to their segmentation accuracy on the simulated PET data. Assessing the resultant segmentations relative to the known ground truth with various metrics measuring accuracy from different aspects clearly shows spectral clustering gives better performance, especially in cases of complex shaped lesions with heterogeneous activity in presence of highly nonuniform background. Intrinsic limitations of the adaptive thresholding algorithm are responsible for its inability to accurately segment lesions of these scenarios. In cases of lesions neighboring

other high uptake regions, it is prone to count those regions or part of those regions in the lesion volumes, thereby leading to oversegmentation. For lesions with heterogeneous activity, thresholding at 70% of the maximum for the initial estimation of the source-to-background contrast may preserve only the high uptake region, thus resulting in an overestimation of the uptake. But on the other hand for lesions with a small contrast, the 70% threshold may give rise to an initial overestimation of the lesion volume, thus inducing an underestimation of the uptake. Consequently, incorrect estimations of the initial contrast, for which the following adaptive thresholding may not be able to compensate, is likely to occur in these cases. Moreover, the ROI used to estimate of the background activity is manually specified. So the variability in ROI selections may also has implications for the estimation of the initial contrast, particularly in cases where lesions are situated in highly heterogeneous background. In contrast, spectral clustering performed well in all these cases, probably largely because it quantifies the intuitive notion that voxels in a cluster are similar and voxels in different clusters are dissimilar by employing a similarity function between pairs of voxels. In doing so it warrants the total intracluster similarity and total intercluster dissimilarity are maximized in the final segmentation. For example in the cases of lesions neighboring other high uptake regions, instead of blindly incorporating all voxels connected to the seed point with intensity above the threshold as adaptive thresholding does, it would form a separate cluster for voxels in a high uptake region provided that doing so yields greater global intra-cluster similarity and intercluster dissimilarity than incorporating them into the cluster corresponding the lesion. In addition, spectral clustering as an automated method does not require estimating the initial contrast and selecting the seed points and is therefore exempt from what inherently restricts adaptive thresholding.

Objective evaluation is crucially important for an automated tumor segmentation method in gaining acceptance in clinical practice. One of the key challenges for the evaluation is the lack of a gold standard against which to compare the obtained segmentation. Given this, a digital phantom was employed in this study to generate images of PET tumors. While retaining access to the ground truth for evaluation, the digital phantom was able to produce PET images much closer to realistic clinical images than are the physical phantoms with geometric volumes and homogeneous activity such as those used by a number of previous studies [10–13]. For the purpose of having the presented method exposed to more challenging segmentation tasks, the simulation was carried out with synthetic lesions placed in the lung region as lesion segmentation in the lung region is complicated by the respiratory-induced motion and the inhomogeneous FDG uptake of the lungs. For the patient data, expert manual segmentations were regarded as the gold standard. Another important facet for segmentation evaluation is the selection of accuracy metrics. In this study, we chose multiple metrics with each emphasizing different aspects of the segmentation accuracy and attempted to provide complementary information on the accuracy and precision of the segmentation result. Dice coefficient and Jaccard coefficient along with false positive dice and false negative dice were used to quantify spatial overlap between the gold standard and the segmented regions. Symmetric mean absolute surface distance was used in measuring the surface difference. For the volumetric metrics, absolute volume difference was adopted. The results obtained by different characterization metrics clearly shows that spectral clustering accurately defines the lesion volumes in the simulated PET data and leads to formation of segmentation map with fairly good agreement to manual annotations on patient data.

Results obtained using the presented method to delineate lesions in presence of heterogeneous backgrounds show improvements over the work of adaptive thresholding and it also has several advantages over the existing approaches. As opposed to the thresholding methods such as adaptive thresholding and iterative thresholding, the presented method does not require any predetermined relationship between threshold and volume or source-to-

background contrast; neither does it require the estimation of the initial contrast and the selection of the seed points. In most of the existing methods, the resultant segmentation is binary or ternary; hence those methods may have limited use in cases of lesions in presence of highly heterogeneous background. In contrast, the presented method is able to automatically derive the number of the clusters and group voxels into distinct clusters in terms of pairwise similarity, thus enabling accurate segmentation of lesions under more complex scenarios. Intensity distance between voxels was used for measuring the pairwise similarity in our implementation of the spectral clustering algorithm, thereby limiting the utility of this method for PET images where lesions and adjacent high uptake tissues exhibit very similar intensity. Possible solution for scenarios of this kind would be to incorporate the structure anatomic information from CT or MRI data into the similarity measure, which is currently under study.

5. Conclusion

In this work, we employed an unsupervised spectral clustering method for FDG-PET tumor segmentation. Its accuracy has been assessed on both simulated and clinical data of complex shapes containing inhomogeneous activities in the presence of highly heterogeneous background. The results shows the feasibility of the presented method to accurately define such lesions.

Acknowledgments

This work was supported in part by National Cancer Institute Grant R01 CA136931. We would like to thank Dr. Michalis Aristophanous for his assistance in PET image simulations and Ms. Karma Kerns for her proofreading the manuscript.

References

1. Paulino A, Koshy M, Howell R, Schuster D, Davis L. Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. *Int J Radiat Oncol Biol Phys.* 2005; 61(5):1385–92. [PubMed: 15817341]
2. Schinagl D, Vogel W, Hoffmann A, Van Dalen J, Oyen W, Kaanders J. Comparison of five segmentation tools for ^{18}F -fluoro-deoxyglucose-positron emission tomography-based target volume definition in head and neck cancer. *Int J Radiat Oncol Biol Phys.* 2007; 69:1282–9. [PubMed: 17967318]
3. Gregoire V, Bol A, Geets X, Lee J. Is PET-based treatment planning the new standard in modern radiotherapy? The head and neck paradigm. *Semin Radiat Oncol.* 2006; 16(4):232–8. [PubMed: 17010906]
4. Nestle U, Kremp S, Schaefer-Schuler A, et al. Comparison of different methods for delineation of F-18-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med.* 2005; 46(8):1342–8. [PubMed: 16085592]
5. Hong R, Halama J, Bova D, Sethi A, Emami B. Correlation of PET standard uptake value and CT window-level thresholds for target delineation in CT-based radiation treatment planning. *Int J Radiat Oncol Biol Phys.* 2007; 67(3):720–6. [PubMed: 17293230]
6. Miller T, Grigsby P. Measurement of tumor volume by PET to evaluate prognosis in patients with advanced cervical cancer treated by radiation therapy. *Int J Radiat Oncol Biol Phys.* 2002; 53(2): 353–9. [PubMed: 12023139]
7. Daisne J, Sibomana M, Bol A, Doumont T, Lonneux M, Gregoire V. Tridimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Radiother Oncol.* 2003; 69(3):247–50. [PubMed: 14644483]
8. Black Q, Grills I, Kestin L, et al. Defining a radiotherapy target with positron emission tomography. *Int J Radiat Oncol Biol Phys.* 2004; 60(4):1272–82. [PubMed: 15519800]

9. Jentzen W, Freudenberg L, Eising E, Heinze M, Brandau W, Bockisch A. Segmentation of PET volumes by iterative image thresholding. *J Nucl Med*. 2007; 48(1):108–14. [PubMed: 17204706]
10. Geets X, Lee J, Bol A, Lonnew M, Gregoire V. A gradient-based method for segmenting FDG-PET images: methodology and validation. *Eur J Nucl Med Mol Imaging*. 2007; 34(9):1427–38. [PubMed: 17431616]
11. Montgomery D, Amira A, Zaidi H. Fully automated segmentation of oncological PET volumes using a combined multiscale and statistical model. *Med Phys*. 2007; 34(2):722–36. [PubMed: 17388190]
12. Yu H, Caldwell C, Mah K, Mozeg D. Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. *IEEE Trans Med Imaging*. 2009; 28(3): 374–83. [PubMed: 19244009]
13. Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. *IEEE Trans Med Imaging*. 2009; 28(6): 881–93. [PubMed: 19150782]
14. Hatt M, le Rest CC, Descourt P, et al. Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications. *Int J Radiat Oncol Biol Phys*. 2010; 77:301–8. [PubMed: 20116934]
15. Shi J, Malik J. Normalized cuts and image segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2000; 22(8):888–905.
16. Chen GL, Lerman G. Spectral curvature clustering (SCC). *Int J Comput Vis*. 2009; 81(3):317–30.
17. Zelnik-Manor L, Perona P. Self-tuning spectral clustering. *Adv Neural Inf Process Syst*. 2004; 17:1601–8.
18. Aristophanous M, Penney BC, Pelizzari CA. The development and testing of a digital PET phantom for the evaluation of tumor volume segmentation techniques. *Med Phys*. 2008; 35(7): 3331–42. [PubMed: 18697557]
19. Babalola KO, Patenaude B, Aljabar P, et al. An evaluation of four automatic methods of segmenting the subcortical structures in the brain. *Neuroimage*. 2009; 47(4):1435–47. [PubMed: 19463960]
20. Zijdenbos A, Dawant B, Margolin R, Palmer A. Morphometric analysis of white-matter lesions in mr-images – method and validation. *IEEE Trans Med Imaging*. 1994; 13(4):716–24. [PubMed: 18218550]

Fig. 1. Axial views of the segmentation results for lesion $S22$, $S25$, $S29$, and $S30$. (a) Ground truth; (b) simulated PET image; segmentations result obtained from (c) spectral clustering, (d) adaptive thresholding.

Fig. 2.

Comparison of segmentation accuracy between the spectral clustering method (SC) and the adaptive thresholding method (AdptThres) when applied on the simulated PET data. *Dice*: Dice coefficient; *Jaccard*: Jaccard coefficient; *FPD*: false positive dice; *FND*: false negative dice; *SMASD*: symmetric mean absolute surface distance; *AVD*: absolute volumetric difference. Dashed lines are indications of ideal values for each measure.

Fig. 3.
Mean and standard deviation of the segmentation accuracy metrics for spectral clustering (SC) and adaptive thresholding (AdptThres) when applied on the simulated PET data.

Fig. 4. Axial views of the segmentation results of four patient scans. (a) PET image; segmentations result from (b) physician annotation, and (c) spectral clustering.

Fig. 5.
Mean and standard deviation of the segmentation accuracy metrics for spectral clustering when applied on patient data.