Recovery of Dynamic PET Regions via Simultaneous Segmentation and Deconvolution

Benjamin Smith[1,2], Ahmed Saad[1,2], Ghassan Hamarneh[1], and Torsten Möller[2]

[1] Medical Image Analysis Lab
[2] Graphics, Usability and Visualization Lab,
School of Computing Science, Simon Fraser University, Canada,
{brsmith, aasaad, hamarneh, torsten }@cs.sfu.ca

Abstract. We derive a new model for simultaneous segmentation and deconvolution of dynamic PET images. By incorporating the PSF of the imaging system into our segmentation model, we simultaneously estimate region boundaries, and correct tissue activities for the partial volume effect. We show improved segmentation results, and outperform two state-of-the-art dynamic PET segmentation methods.

1 Introduction

Dynamic Positron Emission Tomography (dPET) is a functional imaging modality that allows observation of the metabolic activity of biological tissue in vivo. By injecting radioactive tracers into a subject, a 3D+time distribution of tracer uptake can be reconstructed from the resulting photon emissions.

The reconstructed distribution can be considered a 3D volume, where each location is described by a Time Activity Curve (TAC): a vector of tracer concentration measurements. By fitting a compartmental model to the observed TAC, kinetic parameters are recovered which describe the properties of the tissue. These parameters are important for applications in tracer evaluation, clinical investigation, and drug design. Unfortunately, accurate quantification of PET is a difficult problem, and is complicated by both physical factors (photon attenuation and scattering) and characteristics of the scanning hardware itself [1]. Therefore, the images produced by current PET systems exhibit low resolution and high noise. To overcome these limitations in practice, quantification of PET data requires delineation of regions of interest (ROIs) that exhibit homogeneous TACs, and therefore physiological behavior. Accumulated TAC statistics from these regions can be used to compensate for the effects of noise on TAC estimation and the resulting kinetic parameters.

The response of an imaging system to an infinitesimally small point source is known as the point-spread function (PSF). The PSF is large in dPET, and while spatially variant, it can be approximately described by a spatially invariant Gaussian filter with a full width half maximum (FWHM) on the order of $4 - 8mm$ [1]. As a consequence of this large PSF, TACs of neighboring tissue structures are corrupted along their boundaries, and regions blur together. This

is known as "spill–over" and creates a misleading or implausible TAC shape. This effect, especially pronounced in smaller structures, must be corrected in order to accurately quantify true activities and boundaries.

Several approaches have been applied to delineation of ROIs in PET data. Manual delineation by an expert is possible, but is difficult and time consuming, especially for 3D+time data. Automated methods are preferable, and can provide faster results with decreased variability. Early automated methods employed factor analysis or principal component analysis. However, the resulting components do not necessarily have biological significance, and are susceptible to artifacts. Clustering methods have also been applied, most recently by Saad et al. [2] where variants of k–Means were extended with kinetic regularization. All of these methods are susceptible to misclassification errors (Fig. 1, left) because they do not explicitly correct for the effect of the PSF.

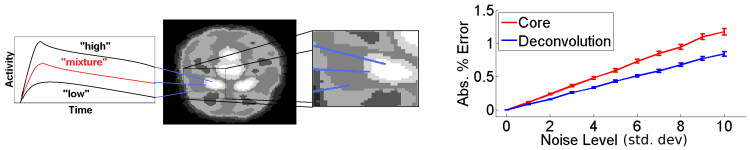

Fig. 1. (Left) TACs in transition regions exhibit a mixture of true activities. Without proper consideration, segmentation methods such as k–Means classify these TACs as belonging to incorrect regions. (Right) Empirical comparison of L_2 recovery error using the core method [3], and our model (Sec. 2.3). As noise increases, the error of the core model increases.

The effect of the PSF introduces a tradeoff. On the one hand, excluding corrupted transition regions from a segmentation will improve parameter recovery, since it will be based only on TACs uncorrupted by the PSF. On the other hand, excluding transition regions results in an incorrect segmentation with a smaller ROI and fewer TACs are available to correct for the effects of noise. In very recent work, Maroy et. al [3] follow the former approach and extract region cores for kinetic parameter estimation. These estimates are based on fewer TAC samples, and are thus more susceptible to noise. We postulate that more robust estimates and improved delineations are obtained by considering the entire region, and *correcting* for the PSF. Fig. 1, right, gives a simple demonstration via a numerical experiment: for a fixed, 1D, piecewise constant signal of width >FWHM of the PSF, increasing levels of noise were added and the activity estimated using two methods. The mean of the samples in the unaffected interval (the "core" region) provided one estimator, while the estimator that corrects for the PSF, developed in Section 2.3, provided another. The developed estimator, using all region statistics, demonstrates a more robust estimate.

Direct correction of the effect of the PSF in dPET data has been investigated previously. Given a complete spatial description of region boundaries, and a model of the PET scanner, the GTM method [1] corrects data for PSF effects in either sinogram or image space. Relevant scanner characteristics can be measured, but knowledge of region boundaries necessitates additional structural information such as CT or MRI. Chiao [4] et al is the only previous work that attempts to estimate both region boundaries and activities, and avoids the cost of acquiring structural information. Chiao et al. use an explicit contour model, and a detailed description of the scanning hardware. However, their method suffers from the known shortcomings of discrete contour models, and is developed for a single ROI of fixed topology, restricting its applicability. In the 2D image processing domain, Bar et. al [5] developed a method to simultaneously perform segmentation and semi–blind restoration of non–medical images. However their model used an edge map instead of an explicit region model, which prevented modeling of important region properties. Many other PET segmentation and correction approaches have been developed. Due to space restrictions, we refer the reader to the references in [1] or [2].

We propose a new method (Sec. 2) for simultaneous segmentation and correction of dPET data. To the best of our knowledge, we are the first to apply a multi–phase level-set Mumford–Shah model, incorporating and accounting for the PSF, to segmentation of homogeneous physiological regions. We address the inaccuracies of state of the art methods, and demonstrate more accurate segmentation and signal recovery (Sec. 3 and 4). Our method is general, requiring no fixed topology, and is therefore not restricted to any specific anatomy. Our region estimates employ all image statistics, and are more resilient to noise than methods which use only core statistics. We make the practical assumptions that the number of regions are known, and the PSF has been measured beforehand. Measurement of the PSF avoids errors associated with PSF estimation.

2 Method

2.1 Model Formulation

At some spatial location $x \in \Omega \subset R^3$ and time $t \in [0, T]$, we model the formation of an observed dPET image, $I(\mathbf{x}, t)$ as:

$$I(\mathbf{x}, t) = \left[I_{true}(t) * h\right](\mathbf{x}) + \eta(\mathbf{x}, t). \tag{1}$$

Here, I_{true} is the unknown spatio-temporal distribution. It is convolved by the PSF of the imaging system, denoted h, which is assumed to be a spatially and temporally invariant Gaussian. Raw dPET data exhibits Poisson noise characteristics, but the noise becomes approximately Gaussian after image reconstruction. Therefore, we assume the observed image is corrupted by a zero–centered, spatially and temporally independent Gaussian noise process, denoted η.

A common assumption in kinetic analysis is that healthy tissue with identical metabolic behavior exhibits identical functional behavior. Therefore, we assume

that the true image can be well approximated by a known number, R, of disjoint, piecewise–constant *functional regions*, each with a characteristic TAC:

$$I_{true}(\mathbf{x},t) \approx \left(\sum_{i=1}^{R} \chi_i(\mathbf{x}) \cdot c_i(t) \right). \tag{2}$$

Here, χ_i denotes the characteristic, or labelling, function that defines the i^{th} region with TAC c_i. The goal is to find the closest estimate of I_{true}, by recovering the characteristic functions χ_i and associated TACs c_i for each region.

We use the approach of Mansouri et al. [6] and represent R regions as combinations of the zero level sets of $R-1$ level set functions $\Phi_k, 1 \leq k \leq R-1$. The zero level set of Φ_k defines a closed area where $\Phi_k > 0$ for points inside the k^{th} contour, and vice versa. Set operations on these areas define the ROIs. The k^{th} region is the region inside Φ_k, and outside all previous $\Phi_i, i \neq k$. The region outside all Φ_k represents the final R^{th} region. Note that the regions are mutually exclusive and cover the entire domain, eliminating the issues of overlap and vacuum [6]. Formally, χ_i for R regions can be expressed using $R-1$ functions, Φ_i, and the Heaviside step, $H(\cdot)$, and dirac delta, $\delta(\cdot)$, functions as:

$$\chi_i = H(\Phi_i)^{1-\delta(R-i)} \left[\prod_{k=1}^{i-1} (1 - H(\Phi_k)) \right]. \tag{3}$$

2.2 Segmenting Functional Regions

Using this representation we seek the time–varying image estimate, $\tilde{I}(\mathbf{x},t)$, formed by convolving the approximation of I_{true} (2) by the imaging PSF, h. This is accomplished by minimizing the following energy function, w.r.t χ_i and c_i:

$$E = \int_{\Omega} \left(\int_{t=0}^{T} \left(I(\mathbf{x},t) - \tilde{I}(\mathbf{x},t) \right)^2 dt + \sum_{i=1}^{R-1} \left[\mu |\nabla \Phi_i|(\mathbf{x}) + \frac{1}{2}(|\nabla \Phi_i|(\mathbf{x}) - 1)^2 \right] \right) d\mathbf{x}. \tag{4}$$

The first term describes the spatio–temporal fit of the estimate $\tilde{I}$ to the observed data. The second and third terms, typical in active contour models, regularize the level sets, encourage smooth region boundaries [7], and maintain the level sets as signed distance functions. The parameter μ controls the influence of the smoothness regularizer. Note that the energy in (4) is a generalization of the vector–valued Active Contours Without Edges (ACWOE) multiphase energy [7]: as the PSF $h(\cdot)$ converges to $\delta(\cdot)$, (4) converges to the ACWOE energy.

Minimizing equation (4) w.r.t the unknown regions χ_i we derive the level set update equations, with artificial step parameter $\tilde{t}$:

$$\begin{aligned} \frac{\partial \Phi_j}{\partial \tilde{t}}(\mathbf{x},\tilde{t}) = 2 \int_{t=0}^{T} \left(\tilde{I}(\mathbf{x},t) - I(\mathbf{x},t) \right) \cdot \left(\left[\left(\sum_{i=1}^{R} \frac{\partial \chi_i}{\partial \Phi_j} \cdot c_i(t) \right) * h \right](\mathbf{x}) \right) dt \\ + (\mu - 1) \nabla \cdot \left(\frac{\nabla \Phi_j}{|\nabla \Phi_j|} \right)(\mathbf{x}) + \Delta \Phi_j(\mathbf{x}). \end{aligned} \tag{5}$$

$$\frac{\partial \chi_i}{\partial \Phi_j} = \begin{cases} \left(-H\left(\Phi_i\right)\right)^{1-\delta(i-j)} \delta\left(\Phi_j\right) \prod_{k=1}^{i-1} \left(1-H\left(\Phi_k\right)\right)^{1-\delta(k-j)} \; ; \; i \neq R, j \leq i \\ -\delta\left(\Phi_j\right) \prod_{k=1}^{i-1} \left(1-H\left(\Phi_k\right)\right)^{1-\delta(k-j)} \; ; \; i = R, j \leq i \\ 0 \; ; \; \forall j > i \end{cases} \tag{6}$$

Note that $H(\cdot)$ and $\delta(\cdot)$ are discontinuous functions, and must be regularized [7]. This regularization must have a localized behavior, or $\tilde{I}$ will suffer from artificial "spill-over" across regions. Therefore we choose:

$$H_\varepsilon(z) = abs(z<1) \cdot \frac{1}{2}\left(1 - \frac{z}{\varepsilon} + \frac{1}{\pi} sin\left(\frac{-\pi z}{\varepsilon}\right)\right), \quad \delta_\varepsilon(z) = -\frac{1}{\varepsilon} cos\left(\frac{z\pi}{\varepsilon}\right). \tag{7}$$

with $\varepsilon = 1$. This restricts the effect of regularization to 1 voxel from the level set boundary. This models the behavior of boundaries that pass through the center of a voxel (i.e.: $H(0) = 0.5$, an equal mixture of two regions).

In dPET, temporal activity is measured by accumulating activity over a small number of time intervals. These intervals are non–uniform, and denoted $\triangle_t$. Discretizing the integral (5) yields the evolution equation to iteratively update Φ_j and consequently $\chi_i \; \forall i$.

$$\begin{aligned} \frac{\partial \Phi_j}{\partial \tilde{t}}(\mathbf{x}, \tilde{t}) \approx 2 \sum_{t=0}^{T} \triangle_t \Big(\tilde{I}(\mathbf{x},t) - I(\mathbf{x},t)\Big) \cdot \left[\left(\sum_{i=1}^{R} \frac{\partial \chi_i}{\partial \Phi_j} \cdot c_i(t)\right) * h\right](\mathbf{x}) \\ + (\mu - 1)\nabla \cdot \left(\frac{\nabla \Phi_i}{|\nabla \Phi_i|}\right)(\mathbf{x}) + \Delta \Phi_i(\mathbf{x}). \end{aligned} \tag{8}$$

As Φ_j and the functional regions are updated, the estimated region TAC c_i must also be updated accordingly. This is described below (Sec. 2.3).

2.3 Calculating Region TAC

To determine the corrected TAC estimates $c_i(t)$, we use the first order optimality conditions from equation (4) to derive the optimal c_i, for a given estimate of region boundaries:

$$\frac{\partial E}{\partial c_m}(t) = -2 \int_\Omega \Big(I(\mathbf{x},t) - \tilde{I}(\mathbf{x},t)\Big) \cdot (\chi_m * h)(\mathbf{x}) d\mathbf{x}. \tag{9}$$

Equation (9) is rewritten as a system of linear equations, for $m = 1..R$, and each time step t. We obtain estimates of $c_i(t)$ as the solution of $Ac(t) = b(t)$, where:

$$A_{ij} = \int_\Omega \left[(\chi_i * h)(\mathbf{x})\right]\left[(\chi_j * h)(\mathbf{x})\right] d\mathbf{x}, \quad b_i = \int_\Omega I(\mathbf{x},t)\left[(\chi_i * h)(\mathbf{x})\right] d\mathbf{x}. \tag{10}$$

3 Validation

In order to perform a quantitative evaluation of accuracy, experiments were performed using synthetic and simulated data (Fig. 2), with a known ground truth. Accuracy of our method (PSF-SEG) was compared to traditional vector–valued ACWOE (ACWOE), and the recently proposed kinetically regularized versions of k–Means (KM-KM) and MRF k–means (KM-MRF) [2]. Parameters were chosen to give the best results for each method, after manual exploration of the parameter space. Following segmentation, the DICE similarity index, a measure of region overlap, was computed for each region to determine segmentation accuracy. A DICE value of 1 indicates perfect segmentation. Recovery of regional kinetic parameters was evaluated by computing a normalized L_2 difference, denoted k–error, between the recovered and true parameters. A k–error of 0 indicates perfect kinetic recovery.

First, synthetic dPET data was created using 2D slices from 12 labeled MRI brain scans. Realistic tissue parameters were used to generate TACs using the COMKAT kinetic modelling tool [8] and FDG model. TACs were assigned to specific brain structures and a Gaussian filter with a FWHM of 6mm [1] was applied to simulate PET imaging hardware. Finally, normally distributed noise was added. The regularization parameter μ for PSF-SEG and ACWOE was fixed for both methods, and chosen proportional to the noise in the image: $\mu = \{5, 25, 75, 125, 125, 150\}$ for $\{0, 1, 3, 5, 8, 10\}$ standard deviations respectively.

Next, 12 Monte Carlo simulated Raclopride dPET volumes [9] were used to compare the methods in 3D under realistic noise and hardware conditions. The PSF of the PET simulator has previously been measured as approximately Gaussian with FWHM of $6.67mm$ in plane and $7.06mm$ axially. The parameter μ was empirically set to 1000 for PSF-SEG and ACWOE. In both datasets, the tissues of interest are: scalp SP, gray matter GM, white matter WM, cerebellum CB and putamen PN.

Crude ROI, at most 2 voxelized spheres per ROI, were drawn on each dataset and used as the manual initializations for all algorithms and subjects.

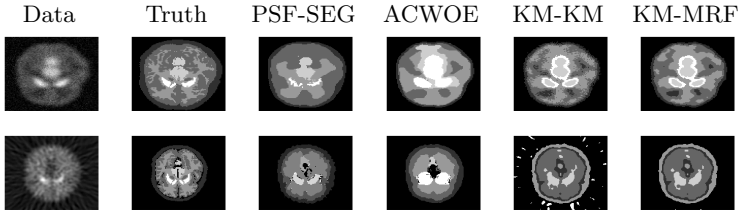

Fig. 2. Column 1: Last time step of input data. Column 2: Ground truth labels. Columns 3–6: Example segmentations. Top row: Data from the noise experiment with noise at 3 std. dev. Bottom row: Slice 40 of the simulated data.

4 Results

For each region and noise level of the synthetic data, the mean DICE index and k–Error were computed with a 95% confidence interval (Fig. 3).

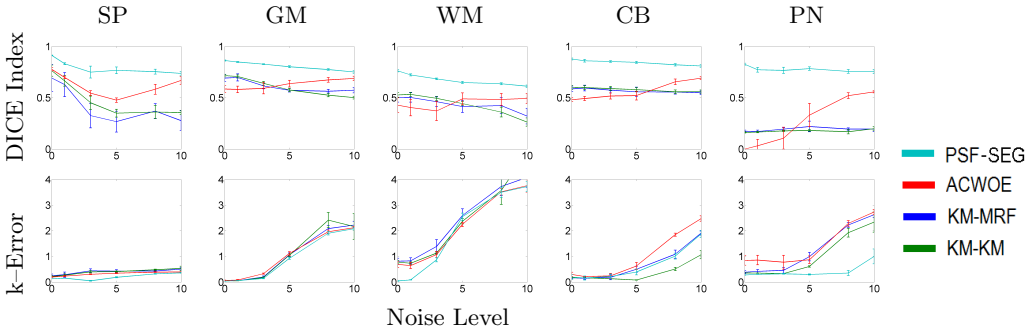

Fig. 3. DICE index (top) and k–error (bottom) for each region in the noise experiment. Error bars denote 95% confidence interval.

The DICE index shows significant and consistent performance across all ROI and noise levels. In general, PSF-SEG realizes an improvement in accuracy between $10-35\%$. Most noteable is the improvement in accuracy in the PN (top–right). KM-MRF, KM-KM and ACWOE erroneously mislabel the putamen and cerebellum as identical regions in favor of giving a unique label to transition regions. The PSF-SEG method successfully deals with these transition regions, and labels the regions correctly. Examining the k–Error, PSF-SEG outperforms for the putamen, due to better segmentation results, and performs similarly to other methods in the remaining regions.

Examining the DICE index for the realistic simulated data (Fig. 4), PSF-SEG significantly outperforms the state of the art algorithms on the simulated WM, CB and PN regions. These regions all contain thinner regions which are affected by the PSF, and subject to misclassification by KM-MRF and KM. PSF-SEG also exhibits low variability, indicating a more reliable, and consistent behavior. PSF-SEG appears to perform slightly worse than the other two methods for GM, but within the variability of these methods. ACWOE also demonstrates improved results. ACWOE's regularized characteristic function simulates the PSF, and improves its performance. For smaller regions, ACWOE's performance decreases as the error in PSF approximation becomes more significant.

5 Discussion and Future Work

We developed a novel model for segmentation of dPET data, based on incorporating the known PSF of the imaging hardware. Improved segmentation results

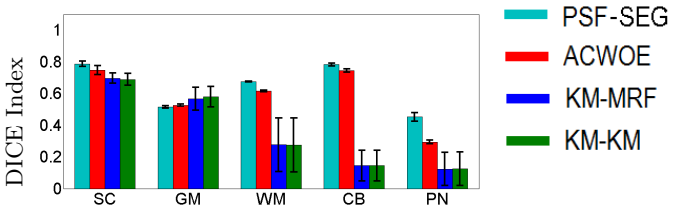

Fig. 4. Segmentation results for realistic simulated data

were demonstrated, and our method outperformed two state-of-the art techniques. In addition, we argue that TAC estimates obtained via this method are more robust under noise than the mean of region cores. The next step is validation on real dPET data, adapting the model to more general, non-piecewise constant models and leveraging the underlying kinetic model of TAC.

References

1. Rousset, O., Zaid, H.: Correction for partial volume effects in emission tomography. In: Quantitative Analysis in Nuclear Medicine Imaging. Springer (2005) 236–271
2. Saad, A., Smith, B., Hamarneh, G., Möller, T.: Simultaneous segmentation, kinetic parameter estimation, and uncertainty visualization of dynamic PET images. In: MICCAI (2). (2007) 726–733
3. Maroy, R., Boisgard, R., Comtat, C., Frouin, V., Cathier, P., Duchesnay, E., Nielsen, P., Trebossen, R., Tavitian, B.: Segmentation of rodent whole-body dynamic PET images: An unsupervised method based on voxel dynamics. IEEE Trans. Medical Imaging **27**(3) (2008) 342–354
4. Chiao, P.C., Rogers, L., Clinthorne, N.H., Fessler, J.A., Hero, A.O.: Model–based estimation for dynamic cardiac studies using ECT. IEEE Trans. Medical Imaging **13**(2) (1994) 217–226
5. Bar, L., Sochen, N., Kiryati, N.: Semi-blind image restoration via mumford-shah regularization. IEEE Trans. Image Processing **15**(2) (2006) 483–493
6. Mansouri, A.R., Mitiche, A., Vasquez, C.: Multiregion competition: a level set extension of region competition to multiple region image partitioning. Computer Vision and Image Understanding **101**(3) (2006) 137–150
7. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Processing **10**(2) (2001) 266–277
8. Muzic, R., Cornelius, S.: Comkat: compartment model kinetic analysis tool. J. Nucl. Medicine **42**(4) (2001) 636–645
9. Reilhac, A., Lartizien, C., Costes, N., Sans, S., Comtat, C., Gunn, R., Evans, A.: PET-SORTEO: A Monte Carlo-based simulator with high count rate capabilities. IEEE Trans. Nucl. Science **51**(1) (2004) 46–52