Unsupervised Clustering of Dynamic PET Images on the Projection Domain

Mustafa E. Kamasak[a] and Bulent Bayraktar[b]

[a]Istanbul Technical University, Istanbul, Turkey
[b]Purdue University, West Lafayette, IN

ABSTRACT

Segmentation of dynamic PET images is an important preprocessing step for kinetic parameter estimation. A single time activity curve (TAC) is extracted for each segmented region. This TAC is then used to estimate the kinetic parameters of the segmented region. Current methods perform this task in two independent steps; first dynamic positron emission tomography (PET) images are reconstructed from the projection data using conventional tomographic reconstruction methods, then the time activity curves (TAC) of the pixels are clustered into a predetermined number of clusters. In this paper, we propose to cluster the regions of dynamic PET images directly on the projection data and simultaneously estimate the TAC of each cluster. This method does not require an intermediate step of tomographic reconstruction for each time frame. Therefore the dimensionality of the estimation problem is reduced. We compare the proposed method with weighted least squares (WLS) and expectation maximization with Gaussian mixtures methods (GMM-EM). Filtered backprojection is used to reconstruct the emission images required by these methods. Our simulation results show that the proposed method can substantially decrease the number of mislabeled pixels and reduce the root mean squared error (RMSE) of the cluster TACs.

Keywords: clustering, dynamic PET, kinetic models, projection domain, regularization

1. INTRODUCTION

Positron emission tomography (PET) images generally have low signal to noise ratio (SNR) and time activity curve (TAC) extracted from a single pixel may be very noisy. To improve the SNR, the TACs obtained from the physiologically similar pixels are averaged and a single TAC is obtained for each group of pixels. Therefore clustering physiologically similar pixels is an important preprocessing step. However this is not a trivial task because of the low SNR and the partial volume effect of the PET images. In many PET studies clustering is performed manually by an operator. Manual clustering is an operator dependent and time consuming process. For improved reproducibility and faster clustering various automatic clustering algorithms are developed.

Ashburner *et al.*[1] proposed a modified mixture model algorithm. This algorithm computes the likelihood of each pixel TAC being in a cluster and iteratively maximizes this likelihood. Wong *et al.*[2] proposed a distance based clustering algorithm. Weighted distance between the pixel TACs within each cluster is minimized. This algorithm is further described in section 3.1. Chen *et al.*[3] used an expectation maximization (EM) based clustering algorithm with Markov random field (MRF) models. Brankov *et al.*[4] proposed a new distance metric between the pixel TACs and iteratively minimizes this distance within the pixel TACs of each cluster. Guo *et al.*[5,6] proposed a hierarchical linkage based algorithm for clustering pixels. Automatic clustering can also be integrated into kinetic parameters estimation algorithms.[7] In some studies, segmentation is used to estimate the plasma input function from the PET images without arterial sampling.[8,9]

These clustering algorithms generally use pixel TACs as their feature vectors. Therefore these algorithms require reconstructed dynamic PET images. Sinogram data acquired with PET scanners are reconstructed using conventional tomographic reconstruction algorithms and TACs are extracted from these reconstructed images. In this paper, we propose a new algorithm which clusters the pixels in the projection domain. Therefore it does

Further author information:
Send correspondence to Mustafa E. Kamasak
E-mail: kamasak@purdue.edu

not require tomographic reconstruction of dynamic PET images. A maximum *a priori* (MAP) based estimation framework is used for clustering pixels and for computing the TAC of each cluster. A similar algorithm was used by Frese *et al.* for discrete tomographic reconstruction of PET images.[10] We extended this algorithm for the unsupervised clustering of dynamic PET pixels directly in the projection domain.

This paper is organized as follows; Section 2 introduces the proposed method that clusters dynamic PET images directly on the projection domain. Section 3 briefly describes the conventional image domain clustering algorithms. The simulation results are given in section 4.

2. UNSUPERVISED CLUSTERING ON PROJECTION DOMAIN

This section describes the unsupervised clustering algorithm on the projection domain. First we will introduce some notation and then give some brief information about the scanner model and describe our MAP framework for clustering.

Assume that, the data is collected at K time frames, and there are L clusters in the image. Each cluster has an associated time activity curve, and a set of pixels that belongs to this cluster. For cluster l, let $\mu_l = [\mu_{l0}, \cdots, \mu_{l(K-1)}]$ denote the time activity curve that represents the cluster, and let $\mathcal{C}_l$ denote the set of pixels that belongs to this cluster. Let μ denote $L \times K$ matrix formed as $\mu = [\mu_0, \mu_1, \cdots, \mu_{L-1}]^T$ where superscript T denotes the matrix transpose. Let $\mathcal{C}$ denote the label image, ie. $\mathcal{C} = \{\mathcal{C}_0, \cdots, \mathcal{C}_{L-1}\}$.

Given the sinogram measurements, denoted by Y, the MAP estimates of μ and $\mathcal{C}$ are

$$(\mu, \mathcal{C}) \leftarrow \arg\max p(\mu, \mathcal{C}|Y) \ , \tag{1}$$

where $p(\cdot)$ denotes the probability.

In the following sections we are going to formulate $p(\mu, \mathcal{C}|Y)$ and then we are going to describe how to estimate $(\mu, \mathcal{C})$ iteratively and efficiently.

2.1. Scanner Model

Let Y_{mk} denote the sinogram measurement for projection $0 \leq m < M$ and time frame $0 \leq k < K$, and let Y be the $M \times K$ matrix of independently distributed Poisson random variables that form the sinogram measurements. Furthermore, let A be the forward projection matrix, with elements A_{ms}. Then the expected number of counts for each measurement at a given time, t_k is given by

$$E[Y_{mk}|\mu, \mathcal{C}] = \sum_{l=0}^{L-1} \sum_{s \in \mathcal{C}_l} A_{ms}\mu_l \ . \tag{2}$$

For simplicity of notation let's define

$$Q_{ml}(\mathcal{C}) \triangleq \sum_{s \in \mathcal{C}_l} A_{ms} \ ,$$

$$Q_m(\mathcal{C}) \triangleq [Q_{m0}, \cdots, Q_{m(L-1)}] \ ,$$

and

$$Q(\mathcal{C}) \triangleq \begin{bmatrix} Q_0 \\ \vdots \\ Q_{M-1} \end{bmatrix}$$

Then equation (2) can be compactly expressed in the matrix notation as

$$E[Y|\mu, \mathcal{C}] = Q(\mathcal{C})\mu \ .$$

Using these notation we can show that the probability density function for the measured sinogram is[11]

$$p(Y|\mu, \mathcal{C}) = \prod_{k=0}^{K-1} \prod_{m=0}^{M-1} \frac{(Q_m(\mathcal{C})\mu_{*k})^{Y_{mk}} e^{-(Q_m(\mathcal{C})\mu_{*k})}}{Y_{mk}!} \tag{3}$$

where μ_{*k} is the k^{th} column of μ. The log likelihood of the sinogram matrix is then given by

$$LL(Y|\mu,\mathcal{C}) = \sum_{k=0}^{K-1}\sum_{m=0}^{M-1} Y_{mk}\log(Q_m(\mathcal{C})\mu_{*k}) - (Q_m(\mathcal{C})\mu_{*k}) - \log(Y_{mk}!) \ . \tag{4}$$

2.2. Estimation Framework

A cost function can be formed by negating the log likelihood given in equation (4) and adding a stabilizing function, $S(\mathcal{C})$.

$$C(Y|\mu,\mathcal{C}) = -LL(Y|\mu,\mathcal{C}) + S(\mathcal{C}) \tag{5}$$

The stabilizing function penalizes the local label changes and therefore it controls the spatial continuity of pixel labels. This type of stabilizing function was used by Besag[12] for image clustering.

The stabilizing function can be obtained from an assumed prior distribution of the label image. In this work, we model the label image as a Markov random field (MRF) with a Gibbs distribution. The likelihood of a particular label image, $\mathcal{C}$ is then

$$p(\mathcal{C}) = \frac{1}{Z}\exp\left\{-\beta\sum_{s,r\in\mathcal{N}} g_{s-r}(1-\delta(c_s,c_r))\right\} \ , \tag{6}$$

where Z is the normalization constant, $\mathcal{N}$ is the set of all spatially neighboring pixel pairs in $\mathcal{C}$, g_{s-r} is the coefficient linking pixels s and r, β is a constant that controls the spatial smoothness of the label image, and $\delta(\cdot,\cdot)$ denotes the Kronicker delta function.

In this paper, $\mathcal{N}$ is formed by 8-point spatial neighborhood. We choose the negative logarithm of (6) as our stabilizing function, ie.

$$S(\mathcal{C}) = \beta\sum_{s,r\in\mathcal{N}} g_{s-r}(1-\delta(c_s,c_r)) \ . \tag{7}$$

Note that with this stabilizing function, high values of β will correspond to spatially smoother label images.

We can similarly add another stabilizing function for the temporal smoothness of the cluster TACs.

2.3. Clustering with Iterative Coordinate Descent Clustering (CICD)

There is no closed form expression for the minimization of the cost function given in (5). Therefore we used an iterative minimization technique that we named clustering with iterative coordinate descent (CICD). It is a modified version of iterative coordinate descent (ICD) algorithm which is commonly used in conventional PET image reconstruction.[11]

A CICD iteration has two steps; first the cluster TACs are fixed and pixel labels are sequentially updated to minimize the cost function. When all pixel labels are updated, the cluster TACs are updated to minimize the cost function. Therefore with each CICD iteration, the cost function given in (5) monotonically decreases.

2.3.1. Pixel Label Update

Assume that we know all cluster TACs and we fix them during the update of pixel labels. Let c_s denote the current label of pixel s, and we want to change it to be $\tilde{c}_s$ in this iteration so that the change in the cost function is minimized. If we change the label of pixel s from c_s to $\tilde{c}_s$, the change in the cost function is

$$\Delta C(Y|c_s,\tilde{c}_s) \triangleq C(Y|c_s) - C(Y|\tilde{c}_s) \ .$$

The evaluation of cost function requires re-computation of log likelihood, which leads to prohibitive computational complexity. Instead of computing the whole log likelihood, we can only compute the change in the log likelihood.

$$\begin{aligned}\Delta LL(Y|c_s,\tilde{c}_s) &\triangleq LL(Y|\tilde{c}_s) - LL(Y|c_s) \\ &= \sum_{k=0}^{K-1}\sum_{m=0}^{M-1}\left\{A_{ms}(\mu_{\tilde{c}_s k} - \mu_{c_s k}) - Y_{mk} log\left(A_{ms}(\mu_{\tilde{c}_s k} - \mu_{c_s k}) + \sum_{l=0}^{L-1} Q_{ml}\mu_{lk}\right)\right\}\end{aligned} \tag{8}$$

Using the changes in the stabilizing function that only depend on the current pixel, the change in the cost function can be written as

$$\Delta C(Y|c_s,\tilde{c}_s) = \Delta LL(Y|c_s,\tilde{c}_s) + \beta \sum_{r\in\partial s} g_{s-r}(1-\delta(\tilde{c}_s,c_r)) \ , \tag{9}$$

where ∂s denotes the set of pixels that are neighbors of pixel s. Then the label of each pixel is updated as

$$\tilde{c}_s \leftarrow \arg\min \Delta C(Y|c_s,\tilde{c}_s) \tag{10}$$

For efficient implementation, $\{Q_{ml}\}_{l=0}^{L-1}$ can be stored in the memory. Whenever a pixel label is updated $\{Q_{ml}\}_{l=0}^{L-1}$ are also be updated as follows

$$\begin{aligned} Q_{mc_s} &\leftarrow Q_{mc_s} - A_{ms} \\ Q_{m\tilde{c}_s} &\leftarrow Q_{m\tilde{c}_s} + A_{ms} \quad \text{for} \quad m = 0\cdots M-1 \ . \end{aligned}$$

2.3.2. Cluster TAC update

Once all the pixel labels are updated, we can update the cluster TACs. For this purpose we compute the first and second derivative of the log likelihood function at each time point. The first and second derivatives of log likelihood with respect to μ_l at time frame k are

$$\theta_{1k} \leftarrow \sum_{m=0}^{M-1} \left\{ Q_{ml} \left(1 - \frac{Y_{mk}}{\sum_{l=0}^{L-1} Q_{ml}\mu_{lk}} \right) \right\} \tag{11}$$

$$\theta_{2kk} \leftarrow \sum_{m=0}^{M-1} Y_{mk} \left(\frac{Q_{ml}}{\sum_{l=0}^{L-1} Q_{ml}\mu_{lk}} \right)^2 \tag{12}$$

Let $\theta_1 \triangleq [\theta_1 0,\cdots,\theta_{1(K-1)}]^T$ and $\theta_2 \triangleq diag\{\theta_{2kk}\}_{k=0}^{K-1}$ Then μ_l can be updated as

$$\tilde{\mu}_l \leftarrow \arg\min \left\{ \theta_1^T(\tilde{\mu}_l - \mu_l) + \frac{1}{2}(\tilde{\mu}_l - \mu_l)^T\theta_2(\tilde{\mu}_l - \mu_l) \right\} \tag{13}$$

There is a closed form expression for $\tilde{\mu}_l$, ie. $\tilde{\mu}_l = \mu_l - \theta_2^{-1}\theta_1$.

3. IMAGE DOMAIN CLUSTERING ALGORITHMS

Image domain clustering algorithms use TACs extracted from emission images. The emission images are reconstructed using conventional PET reconstruction algorithms. Let x_{sk} be the reconstructed emission rate for pixel s at time frame k, and $x_s = [x_{s0},\cdots,x_{s(K-1)}]$ be the reconstructed time response of pixel s.

3.1. Weighted Least Squares Clustering (WLS)

This algorithm minimizes the weighted square distance between the pixel TACs and the cluster TACs, ie.

$$(\mu,\mathcal{C}) \leftarrow \arg\min_{\mu,\mathcal{C}} \sum_{l=0}^{L-1} \sum_{s\in\mathcal{C}_l} \|x_s - \mu_l\|_W^2 \ , \tag{14}$$

where W is a weight matrix, and $\|x\|_W^2$ denotes x^TWx. In this work we used a diagonal weighting matrix formed as $W = diag\{\Delta t_k\}_{k=0}^{K-1}$ where Δt_k is the duration of k^{th} time frame.

This algorithm also iteratively updates the pixel labels and cluster TACs. Each iteration consists of two steps. In the first step, labels of pixels are sequentially updated. The label of a pixel is updated as follows

$$\tilde{c}_s \leftarrow \arg\min_l \|x_s - \mu_l\|_W^2 \tag{15}$$

After all pixel labels are updated, the cluster TACs are updated as follows to decrease the weighted distance given in (14).

$$\mu_l = \frac{1}{|\mathcal{C}_l|} \sum_{s \in \mathcal{C}_l} x_s \ , \tag{16}$$

where $|\mathcal{C}_l|$ denotes the number of pixels that are labeled as l. Each CICD iteration monotonically decreases the cost function, and iterations are repeated until the stopping (convergence) criteria is reached.

3.2. Gaussian Mixture Model with Expectation Maximization (GMM-EM)

It can be assumed that the pixel TACs are Gaussian distributed around the cluster TACs. Similar to other clustering methods pixel labels and cluster TACs can be updated iteratively.

Let R_l denote the covariance matrix of cluster l, and π_l denote the probability of cluster l. The posterior probability of a pixel being in cluster l, given its time response is

$$p(c_s = l | x_s, \mu_l) = \frac{\pi_l}{(2\pi)^{K/2}} |R_l|^{-1/2} \exp\left\{ -\frac{1}{2} (x_s - \mu_l)^T R_l^{-1} (x_s - \mu_l) \right\} \tag{17}$$

If the TACs and covariance matrices of the clusters are known, we can assign pixel labels to maximize the posterior, ie.

$$c_s \leftarrow \arg\min_l \left\{ \frac{1}{2} (x_s - \mu_l)^T R_l^{-1} (x_s - \mu_l) + \frac{1}{2} \log |R_l| - \log(\pi_l) \right\} \tag{18}$$

Once the labels are assigned the cluster TACs and covariance matrices can be updated using the EM algorithm.[13]

3.3. Initialization Clustering Algorithms

All clustering algorithms described above require initial cluster TACs, pixel labels, or both. It is possible to initialize these algorithms with randomly chosen initial labels and cluster TACs. To avoid local minima in these iterative algorithms, these algorithms should be executed multiple times with different initial points. The set of initial points that results in the lowest final cost should be used.

It is also possible to start them with user selected seed points. The best candidate from each cluster can be manually selected and their corresponding TACs can be used to initialize these algorithms.

4. RESULTS

4.1. Simulations

Our simulation experiments are based on a phantom of a rat's head. The phantom and kinetic parameters for the regions in this phantom are taken from Kamasak *et al.*[14] Figure 1 shows a schematic representation of the phantom and its regions. The phantom has 7 regions including the background. The regions and their corresponding parameters are given in Table 1, and their TACs are shown in Fig. 2. For further details about the phantom see Kamasak *et al.*[14] Time frames of emission images are generated using these parameter images and the 2-tissue compartment model equations, and the plasma function, $C_P(t)$, is generated using the second model in Wong *et al.*[15] The blood contribution to the PET activity is assumed to be zero, and the tracer is assumed to be raclopride with ^{11}C, which has a decay constant of $\lambda = 0.034$ min^{-1}. Total scan time is 60 min., divided into 18 time frames with 4×0.5 min, 4×2 min, and 10×5 min. The phantom resolution is 128×128 with each pixel having dimensions of (1.2 mm)^2.

The rat phantom image at each time frame is forward projected into sinograms using a Poisson model for the detected counts. Each sinogram consists of 180 angles and 200 radial bins per angle.

The emission images required by image domain clustering algorithms are generated using filtered backprojection algorithm (FBP). The initial cluster TACs are chosen manually. Same initial points are used for all clustering algorithms.

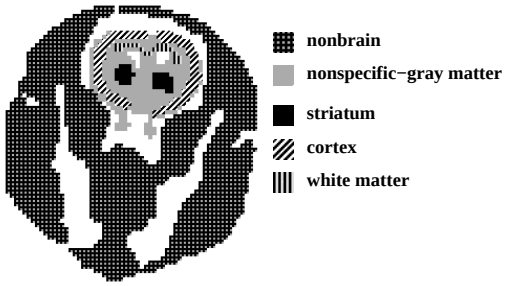


Figure 1. Single-slice rat phantom. Regions of the rat phantom were derived from a segmented MR image. Different fill patterns indicate kinetically distinct tissue regions. Striatum is a region containing specific receptors for the tracer. Nonspecific-gray matter is tissue containing no specific binding sites for tracer but comparable blood flow parameters (k1, k2) to striatal area; cortex is modeled as containing low concentration of binding sites; white matter in our dynamic phantom contains no specific binding sites and low flow;non-brain, which comprises much of the slice has fast influx and efflux of tracer. Solid white areas in figure represent a mixture of background regions that do not contain any activity over time. The small white areas dorsal to (above) the striatum are ventricles that contain cerebral spinal fluid and no tracer. White areas surrounding brain correspond to skull which does not take up appreciable amounts of tracer.

Region	k_1 min^{-1}	k_2 min^{-1}	k_3 min^{-1}	k_4 min^{-1}	a min^{-1}	b min^{-1}	c min^{-1}	d min^{-1}
Background	0	0	0	0	0	0	0	0
CSF	0	0	0	0	0	0	0	0
Nonbrain	.1836	.8968	0	0	.1836	0	.8968	0
Nonspecific-gray matter	.0918	.4484	0	0	.0918	0	.4484	0
Striatum	.0918	.4484	1.2408	.1363	.02164	.07016	1.7914	.0312
Cortex	.0918	.4484	.141	.1363	.0607	.0311	.628	.09725
White matter	.02295	.4484	0	0	.02295	0	.4484	0

Table 1. Kinetic parameters used in the simulations for distinct tissue regions of the rat head.

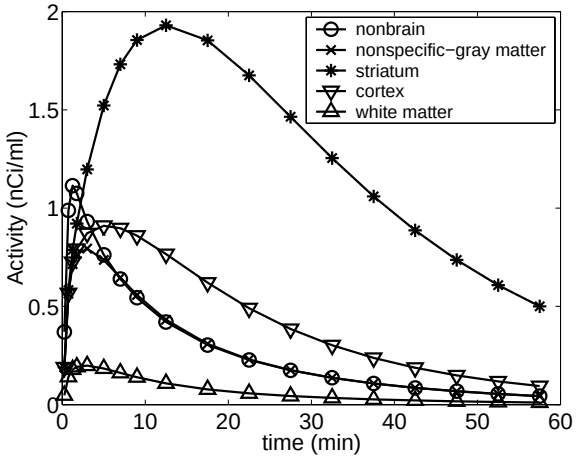
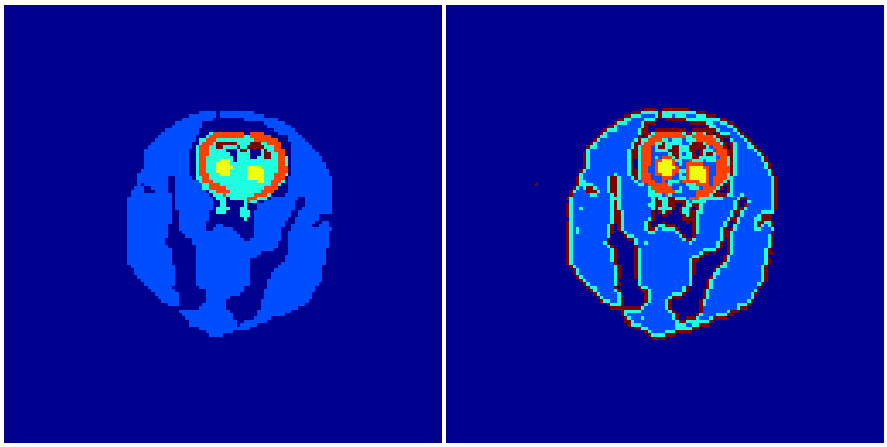
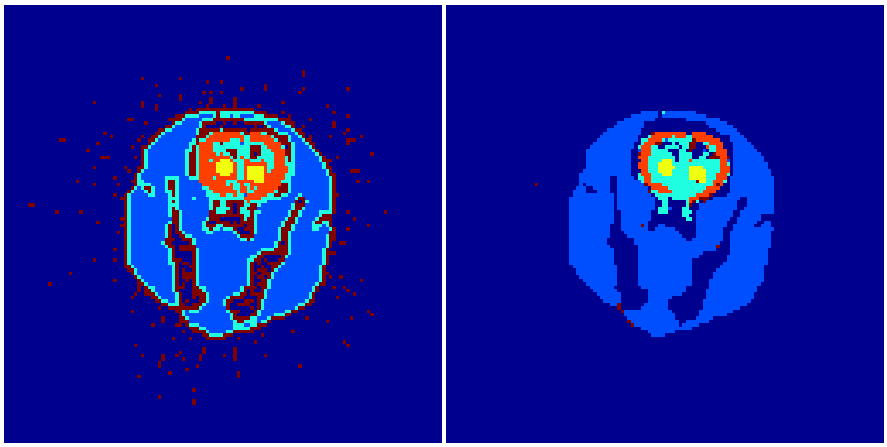


Figure 2. Time-activity curves for 5 distinct tissue regions in rat brain phantom.



(a) Original (b) FBP+WLS



(c) FBP+GMM-EM (e) CICD

Figure 3. Pixel labels assigned by the clustering algorithms.

4.1.1. Simulation Results

The pixel labels assigned by the image domain algorithms and the proposed method, CICD, are shown in figure 3. Visually it can be seen that CICD algorithm results have less mislabeled pixels than image domain clustering methods. The percentage of mislabeled pixels for these algorithms are given in table 2. From this table, it can be seen that the proposed clustering algorithm has the lowest mislabeled pixel percentage.

The cluster TACs estimated by the clustering algorithms are shown in figure 4. The root mean squared error for the cluster TACs are given in table 3. This table shows that for all the regions except the white matter, the proposed algorithm have produced the lowest RMSE between the estimated cluster TACs and the actual cluster TACs.

The success of the proposed CICD algorithm is due to the reduction in the number of estimated parameters. CICD algorithm assigns N labels and estimates $L \times K$ time points for cluster TACs. In addition to these, image domain clustering algorithms require the estimation of $N \times K$ emission rates for reconstructed emission images.

5. CONCLUSION

We proposed a new clustering algorithm that we call clustering with iterative coordinate descent (CICD). CICD clusters the dynamic PET images directly on the projection domain and it does not require the reconstruction of emission images. The results of CICD algorithm are substantially better than the conventional image domain

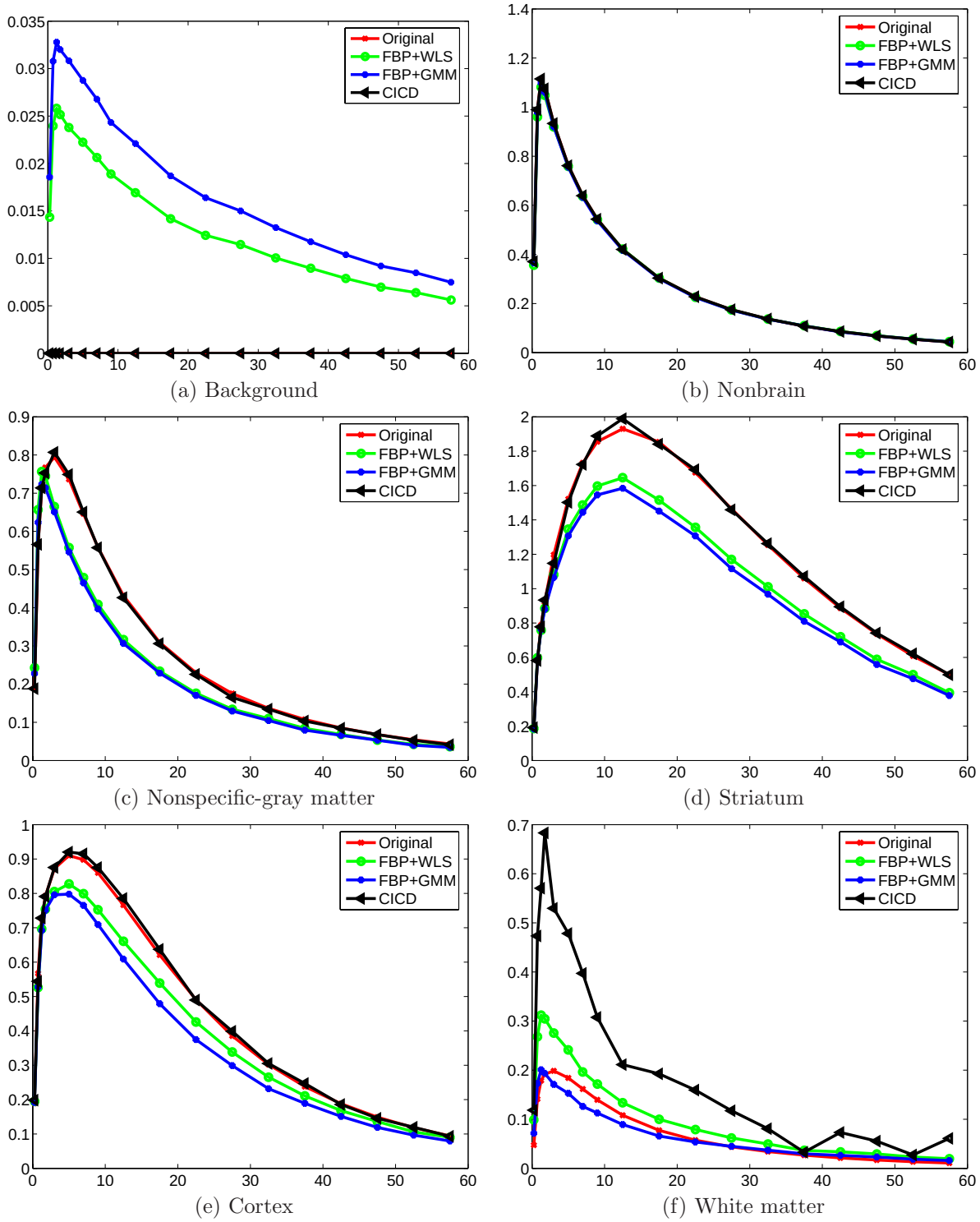


Figure 4. Cluster TACs estimated by the clustering algorithms for each region in the rat head.

method	misclassification (%)
FBP + WLS	6.951904
FBP + GMM	9.136963
CICD	0.439453

Table 2. Percentage of mislabeled pixels for the clustering algorithms.

region	WLS	GMM-EM	CICD
background	0.017	0.022	0.000
nonbrain	0.013	0.007	0.0004
nonspecific-gray matter	0.088	0.092	0.0063
striatum	0.207	0.239	0.0219
cortex	0.059	0.088	0.0110
white matter	0.059	0.019	0.216

Table 3. RMSE of the cluster TACs for each region in the rat's head.

clustering algorithms. It produces less mislabeled pixels and estimates cluster TACs with lower RMSE than the image domain clustering algorithms.

Therefore the proposed CICD algorithm is quite promising for the region of interest analysis before the kinetic parameter estimation. We believe that more accurate estimates for the kinetic parameters can be obtained using this algorithm.

However more tests on real dynamic PET data are required for further analysis of this algorithm. More flexible regularization strategies can also be integrated into this algorithm which may result in better clustering.

REFERENCES

1. J. Ashburner, J. Haslam, V. J. C. C. Taylor, and T. Jones, "A cluster analysis approach for the characterization of dynamic PET data," in *Quantification of Brain Function Using PET*, R. Myers, V. Cunningham, D. Bailey, and T. Jones, eds., pp. 301–306, Academic Press, San Diego, 1996.
2. K.-P. Wong, D. Feng, S. R. Meikle, and M. J. Fulham, "Segmentation of dynamic PET images using cluster analysis," *IEEE Trans. on Nuclear Science* **49**(1), pp. 200–207, 2002.
3. J. L. Chen, S. R. Gunn, M. S. Nixon, and R. N. Gunn, "Markov random field models for segmentation of PET images," in *Proceedings of Information Processing in Medical Imaging 2082*, pp. 468–474, 2001.
4. J. G. Brankov, N. P. Galatsanos, Y. Yang, and M. N. Wernick, "Segmentation of dynamic PET or fMRI images based on a similarity metric," *IEEE Trans. on Nuclear Science* **50**(5), pp. 1410–1414, 2003.
5. H. Guo, R. Renaut, and K. Chen, "Clustering for three dimensional kinetic PET data," in *Proceedings of IEEE International Conference on Data Mining, Clustering Large Data Sets, Workshop notes*, pp. 43–48, (Melbourne, Florida), 2003.
6. H. Guo, R. Renaut, K. Chen, and E. Reiman, "Clustering huge data sets for parametric PET imaging," *Biosystems* **71**(1-2), pp. 81–92, 2003.
7. M. S. Y. Kimura and N. Alpert, "Fast formation of statistically reliable FDG parametric images based on clustering and principal components," *Phys. Med. Biol.* **47**, pp. 455–468, August 2002.
8. M. Liptrot, K. H. Adams, L. Martiny, L. H. Pinborg, M. N. Lonsdale, N. V. Olsen, S. Holm, C. Svarer, and G. M. Knudsen, "Cluster analysis in kinetic modelling of the brain: a noninvasive alternative to arterial sampling," *NeuroImage* **21**(2), pp. 483–493, 2004.
9. K.-P. Wong, D. Feng, S. R. Meikle, and M. J. Fulham, "Non-invasive extraction of physiological parameters in quantitative PET studies using simultaneous estimation and cluster analysis," in *Proceedings of IEEE Medical Imaging Conference*, pp. 141–145, (Lyon, France), October 2000.

10. T. Frese, C. A. Bouman, and K. Sauer, "Multiscale Bayesian methods for discrete tomography," in *Discrete Tomography: Foundations, Algorithms and Applications*, G. T. Herman and A. Kuba, eds., pp. 237–261, Birkhauser Boston, Cambridge, MA, 1999.
11. C. A. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. on Image Processing* **5**, pp. 480–492, March 1996.
12. J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society B* **48**(3), pp. 259–302, 1986.
13. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society B* **39**(1), pp. 1–38, 1977.
14. M. E. Kamasak, C. A. Bouman, E. D. Morris, and K. Sauer, "Direct reconstruction of kinetic parameter images from dynamic PET data," *IEEE Trans. on Medical Imaging* **24**, pp. 636–650, May 2005.
15. K.-P. Wong, D. Feng, S. R. Meikle, and M. J. Fulham, "Simultaneous estimation of physiological parameters and the input function - in *vivo* PET data," *IEEE Transactions on Information Technology in Biomedicine* **5**, pp. 67–76, March 2001.