Quantitative Cardiac Dynamic Imaging of Small Animal PET Images Using Cluster Analysis

S Domenichelli[1,2], D D'Ambrosio[1,2], S Trespidi[3], C Nanni[3], V Ambrosini[3], S Boschi[3], R Franchi[3], M Marengo[2], AE Spinelli[1,2]

[1]Scuola di Specializzazione in Fisica Sanitaria, Università di Bologna, Italy
[2]Servizio di Fisica Sanitaria, Policlinico S Orsola - Malpighi, Bologna, Italy
[3]Servizio di Medicina Nucleare, Policlinico S Orsola - Malpighi, Bologna, Italy

Abstract

Quantitative PET imaging requires a dynamic scan in order to measure the arterial input function and the tissue time-activity curves. By combining these two curves with adequate mathematical models it is possible to obtain useful physiological information such as the metabolic rate, perfusion, receptors density etc. Cluster Analysis (CA) allows to group pixels having the same kinetic. In this work the performance of two clustering algorithms were assessed. The user must supply a set of images acquired at different time points and the number of clusters. The choice of the correct number of clusters was performed by using a parsimony criteria. In order to test the CA method real dynamic small animal PET data were acquired. Image derived arterial input function and myocardial FDG uptake were measured. Results showed that CA allow us to obtain accurate tissue time activity curves without the need of manual region on interest delineation.

1. Introduction

Positron emission tomography (PET) is an imaging technique that allows us to obtain *in vivo* measurement of useful physiological parameters such as perfusion, glucose metabolic rate, receptors density etc. Quantitative measurements can be obtained by combining the arterial input function and tissue time activity curves (TAC) with adequate mathematical models [1-3].

In order to obtain the arterial input function and the tissue time activity curves region of interest (ROI) are drawn on the left ventricle and the organ of interest. This approach is time consuming and operator dependent. In addition, ROI drawing can be difficult when applied to small structures, as in small animal PET studies.

Cluster analysis (CA) is a multivariate data analysis technique that provides an automated ROI delineation method [4-6]. The main goal of CA is to split a large data set into a smaller number of clusters, such that points belonging to the same cluster have similar characteristics. If we assume that dynamic PET data can be represented by a finite number of kinetics and that every point belonging to the same tissue has the same kinetic, CA permits to group automatically pixels with the same kinetic.

In this work we chose to consider two different clustering algorithms: K-means and Fuzzy C-means. Both are iterative procedures that minimize the distance between data set objects and segment the image into a fixed number of clusters. K-means algorithm creates a small number of clusters that are mutually exclusive and exhaustive, so that objects into the same cluster are similar to each other while objects of different clusters are dissimilar. Instead, Fuzzy logic defines a membership degree for every point to every cluster [7, 8].

In this study, we apply CA to segment automatically dynamic rat cardiac PET images obtained using ^{18}fluorine-fluoro-2-deoxy-D-glucose (FDG). This procedure allow us to extract the arterial input function and the myocardium time activity curves in order to obtain for example *in vivo* measurement of the myocardium glucose metabolic rate.

2. Methods

2.1. Cluster analysis

A clustering algorithm analyses a multidimensional data set and splits data into a small number of distinct classes, minimizing the distance between points in the same cluster. Any cluster defines a region in the image where points have a similar kinetic. Each cluster is represented by a centroid that can be considered as the average of TACs in the cluster.

For simplicity we can assume that exist K tissues and that every tissue has a different kinetic described by a characteristic curve. PET data can be stored into a 4-D matrix **M**, which contains information about the radiotracer concentration of each voxels at

different times. As already mentioned in the introduction we chose to consider two different clustering algorithms: K-means and Fuzzy C-means. The main goal of each algorithm is to minimize an objective function D representing pixel within-cluster distance. Both K-means and Fuzzy C-means segment the image into a fixed number of clusters. K-means assigns every point to only one cluster, while Fuzzy C-means defines a membership degree $u_{i,k}$ for the ith point to the jth cluster. Fuzzy C-Means seems to be suited in medical image processing because biological tissues have overlapping grey-scale intensity distributions due to imperfect image uniformity, noise and partial volume effects. The objective function D can be defined as follow:

$$D_m = \sum_{j=1}^{K} \sum_{i=1}^{N} \left(u_{i,j} \right)^m \cdot \left(d_{i,j} \right)^a \text{ , and } \sum_{j=1}^{K} u_{i,j} = 1$$

where $d_{i,j}$ is the distance between the ith tissue TAC inside the data matrix **M** and the jth centroid. $m \in [1, +\infty[$ is a weighting exponent called the fuzzifier. For the K-Means algorithm $u_{i,j} = 0,1$ and $m = 1$, while for the Fuzzy C-Means algorithm $0 \le u_{i,j} \le 1$ and m is typically equal to 2. The distance $d_{i,j}$ can be evaluated using different metrics: in this study we used K-means algorithm with Manhattan ($a = 1$) and Euclidean ($a = 2$) metrics and Fuzzy C-Means with Euclidean metric. All the code was implemented using Interactive Data Language (IDL) 6.4.

These two clustering algorithms allow us to automatically group pixels in an unsupervised manner, only by trying to consider similarity into data set objects, without using any kind of information. The user must define the number of clusters K and to choose initial cluster centroids in order to start the algorithm. In our case they are initialized at random. Every point of the image is iteratively associated to one cluster and the objective function D is evaluated. The process is repeated until D converges to a minimum. In this case points are no more moved from one cluster to another. Since the original data set belong to a high dimensional space the convergence to a global minimum is not guaranteed because of the presence of several local minima in the solution space. In order to avoid local minima, it is necessary to restart the algorithm by changing the initial cluster centroids.

In this work, the clustering was performed independently on each slice, however the proposed method could work also on fully 4-D data set. Obviously the use of the whole volume can be computationally demanding. In this case in order to reduce the computational time it is possible to define a threshold or to perform a preclustering [9].

As mentioned earlier, the user must specify the number of cluster K in which the algorithm partitions the data. Usually the correct number of clusters is not known *a priori* and typically the elbow criterion is used to determine it [10]. This criterion consists into assessing the validity of clustering by plotting the average Mean-Squared Error (MSE) across clusters. It is based on the observation that the within-cluster dispersion, defined as the sum of distance between any two data points in the same clusters (see the following equation), is reduced when number of clusters increases. On the other hand, the reduction in the within-cluster dispersion usually decreases significantly when the number of clusters exceeds the correct number.

$$MSE = \frac{1}{N} \sum_{j=1}^{K} \sum_{i=1}^{N} \left(d_{i,j} \right)^2$$

In this work, we considered that the optimum number for clustering our data is when the within-cluster dispersion decreases less than 33%.

2.2. Rat cardiac dynamic PET image acquisition

In order to validate *in vivo* the CA method described in section 2.1 a rat cardiac study was performed. More precisely a dynamic PET scan was performed by injecting in the tail a 400 g rat with 50 MBq of FDG. Image acquisition was performed using a dedicated small animal PET scanner (GE eXplore Vista) [11].

List mode data were acquired using a 100-700 keV energy window for 60 minutes and then reformatted into 30 sequential frames. No corrections for randoms or scatter were performed. Image were reconstructed using an iterative 2D-OSEM algorithm after Fourier (FORE) rebinnig. The size of the reconstructed dynamic image (the data matrix **M**) were equal to: 175x175x61x30 and the corresponding voxel size was equal to: 0.38x0.38x0.77 mm^3.

3. Results

The results of CA performed on a dynamic cardiac FDG small animal PET scan are shown in figure 1. Clusters are represented by different grey levels. The image clearly show that the clustering algorithm groups correctly regions with similar kinetics.

As one can see the myocardium and left ventricle are well delineated permitting to extract accurate tissue time activity curves shown in figures 2 and 3.

For cardiac FDG studies it is possible for example to use these curves to calculate the glucose metabolic rate k_m using graphical analysis with the Patlak plot as shown in figure 4 [12].

Fig 1. The image on top show the last frame of a cardiac dynamic PET scan obtained by injecting 50 MBq of FDG in a 400 g rat. The bottom image shows the results of cluster analysis, as one can see the myocardium and left ventricle are well delineated.

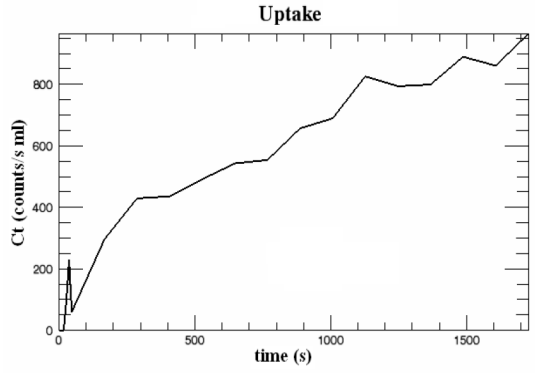

Fig 2. The plot show the myocardium tissue time activity curve obtained using the CA.

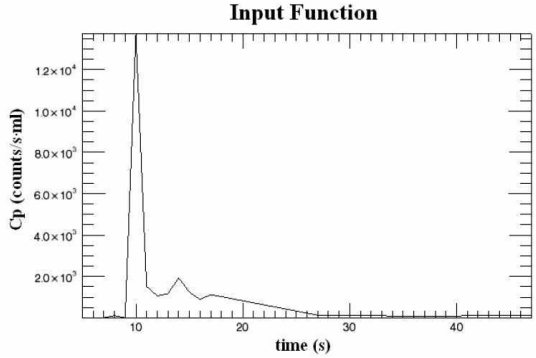

Fig 3. The plot show the arterial input function from the left ventricle cluster centroid shown in figure 1.

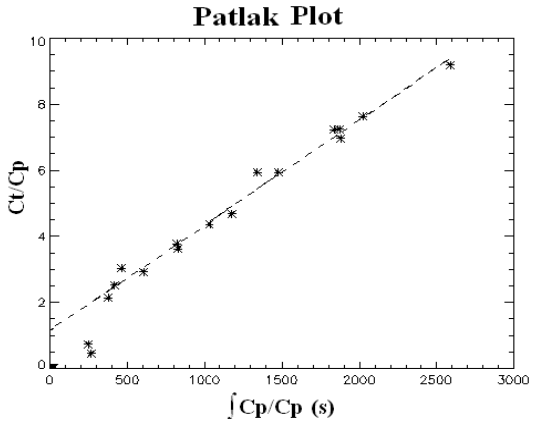

Fig 4. Patlak Plot obtained starting from input function and myocardium uptake curves shown in figures 2 and 3. The slope is equal to the glucose metabolic rate k_m.

4. Discussion and conclusions

In this paper we proposed a CA method for image segmentation and automatic tissue TAC extraction for dynamic small animal PET images. The use of hand drawn ROI is commonly used in the clinical and research practice to analyse PET datasets. Such approach is operator dependent and time-consuming. The proposed method based on CA has the advantage of being more fast and reliable by reducing the dependence from the operator and improving at the same time the reproducibility of the results.

Consequently the physiological parameters obtained from a dynamic scan like for example the metabolic rate have lower error with respect to the values calculated with the manual ROI-based method (data not shown).

The number of clusters K suitable to represent the considered data set is not known *a priori*. However we can reasonably assume that the PET data matrix **M** contain a small finite number of clusters (typically less than 5). By using the elbow criterion, one can determine the optimum number of clusters for the given dataset. In the literature there several statistical criteria than can be used to determine the optimal number of clusters like for example the Akaike's Information Criterion [13-15].

One interesting aspect of the CA approach described in this paper is that the proposed method partitions the PET data without considering voxel location, but only valuating temporal informations represented by voxel TAC.

One of the main limitation of CA is the ability to distinguish between anatomical structures that have similar kinetic. The result of CA may also depend on the chosen algorithm and on the metric used to valuate the objective function D. In this study, we showed that K-means Euclidean algorithm is successful in correctly recognizing different functional areas in real small animal PET images.

Acknowledgements

The authors would like to acknowledge the Fondazione del Monte di Bologna for the support.

References

[1] Meikle SR, Eberl S, Iida H, Instrumentation and Methodology for Quantitative Pre-Clinical Imaging Studies, Current Pharmaceutical Design, 2001, 7, 1945-1966.

[2] Maguire RP, Leenders KL, PET Pharmacokinetic Course Manual, Montreal Neurological Institute, Montreal, Canada, 2003.

[3] Gunn RN, Gunn SR, Cunningham VJ, Positron Emission Tomography Compartmental Models, Journal of Cerebral Blood Flow and Metabolism, 2001, 21, 635–652.

[4] Liptrot M, Adams KH, Martiny L, Pinborg LH, Lonsdale MN, Olsen NV, Holm S, Svarer C, Knudsena GM, Cluster analysis in kinetic modelling of the brain: a non-invasive alternative to arterial sampling, NeuroImage, 2004, 21, 483–493.

[5] Wong KP, Feng D, Meikle SR, Fulham MJ, Non-invasive Estimation of Cerebral Metabolic Rate of Glucose Using Simultaneous Estimation and Cluster Analysis: A Feasibility Study, Conferences in Research and Practice in Information Technology, 2000, Vol.2.

[6] Wu HM, Kreissl MC, Schelbert HR, Ladno W, Prins M, Shoghi-Jadid K, Chatziioannou A, Phelps ME, Huang SC, First-Pass Angiography in Mice Using FDG-PET: A Simple Method of Deriving the Cardiovascular Transit Time Without the Need of Region-of-Interest Drawing, IEEE Tans. Nucl. Sci., 2005, Vol.52, 5, 1311-1315.

[7] Bezdek JC, Hathaway RJ, Recent convergence results for the fuzzy c-means clustering algorithms, J. of Classification, 1988, 5, 237-247.

[8] Zaidi H, Diaz-Gomez M, Boudraa A, Slosman DO, Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging, Physics in Medicine and Biology, 2002, 47, 1143-1160.

[9] Guo H, Renaut R, Chen K, Reiman E, Clustering huge data sets for parametric PET imaging, BioSystems, 2003, 71, 81-92.

[10] Chen F, El-Dawlatly S, Jin R, Oweiss K, Identifying and Tracking the number of independent clusters of functionally interdependent neurons from biophysical models of population activity, Proc. of 3rd Int. IEEE EMBS Conf. on Neural Engineering, pp. 542-545, 2007.

[11] Spinelli AE, D'Ambrosio D, Pettinato C, Trespidi S, Nanni C, Ambrosini V, Baldazzi G, Bergamini C, Marengo M, Performance evaluation of a Small Animal PET Scanner. Spatial resolution characterization using 18-F and 11-C Nucl Instr and Meth in Phys Res. 2007 571, 1-2, 215-218

[12] Spinelli AE, Quantitative Dynamic Imaging Using Positron Emission Tomography, PhD. Thesis, London University, 2003.

[13] Akaike H, A New Look at the Statistical Model Identification, IEEE Transaction on Automatic Control, 1974, Vol.19, 6, 716-723.

[14] Kullback S, Leibler RA, On information and sufficiency, The Annals of Mathematical Statistics, 1951, Vol.22, 1, 79-86.

[15] Wong KP, Feng D, Meikle SR, Fulham MJ, Segmentation of Dynamic PET Images Using Cluster Analysis, IEEE Tansaction on Nuclear Science, 2002, Vol.49, 1, 200-207.

Address for correspondence:
Antonello E Spinelli, Ph.D.
Servizio di Fisica Sanitaria,
Policlinico S. Orsola – Malpighi,
Via Massarenti, N. 9, 40138, Bologna, Italy
E-mail address: antonello.spinelli@unibo.it