Chapter 3

Medical Image Segmentation: Methods and Applications in Functional Imaging

Koon-Pong Wong

3.1 Introduction

Detection, localization, diagnosis, staging, and monitoring treatment responses are the most important aspects and crucial procedures in diagnostic medicine and clinical oncology. Early detection and localization of the diseases and accurate disease staging can improve the survival and change management in patients prior to planned surgery or therapy. Therefore, current medical practice has been directed toward early but efficient localization and staging of diseases, while ensuring that patients would receive the most effective treatment.

Current disease management is based on the international standard of cancer staging using TNM classification, viz. size, location, and degree of invasion of the primary tumor (T), status of regional lymph node (N), and whether there is any distant metastasis (M). Over the century, histopathology retains its main role as the primary means to characterization of suspicious lesions and confirmation of malignancy. However, the pathologic interpretation is heavily dependent on the experience of the pathologist, and sampling errors may mean that there are insufficient amounts of tissue in the specimens, or the excised tissue does not accurately reflect tumor aggressivity. In addition, some lesions may return nondiagnostic information from the specimens, or they are difficult or too

[1]Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

dangerous to biopsy. As a result, more invasive and unpleasant diagnostic procedures are sought.

The last few decades of the twentieth century have witnessed significant advances in medical imaging and computer-aided medical image analysis. The revolutionary capabilities of new multidimensional medical imaging modalities and computing power have opened a new window for medical research and clinical diagnosis. Medical imaging modalities are used to acquire the data from which qualitative and quantitative information of the underlying pathophysiological basis of diseases are extracted for visualization and characterization, thus helping the clinicians to accurately formulate the most effective therapy for the patients by integrating the information with those obtained from some other possibly morbid and invasive diagnostic procedures. It is important to realize that medical imaging is a tool that is complementary to but not compete with the conventional diagnostic methods. Indeed, medical imaging provides additional information about the disease that is not available with the conventional diagnostic methods, and paves a way to improve our understanding of disease processes from different angles.

Modern medical imaging modalities can be broadly classified into two major categories: *structural* and *functional*. Examples of major *structural imaging* modalities include X-ray computed tomography (CT), magnetic resonance imaging (MRI), echocardiography, mammography, and ultrasonography. These imaging modalities allow us to visualize anatomic structure and pathology of internal organs. In contrast, *functional imaging* refers to a set of imaging techniques that are able to derive images reflecting biochemical, electrical, mechanical, or physiological properties of the living systems. Major functional imaging modalities include positron emission tomography (PET), single-photon emission computed tomography (SPECT), fluorescence imaging, and dynamic magnetic resonance imaging such as functional magnetic resonance imaging (fMRI) and magnetic resonance spectroscopy (MRS). Fundamentally, all these imaging techniques deal with reconstructing a three-dimensional image from a series of two-dimensional images (projections) taken at various angles around the body. In CT, the X-ray attenuation coefficient within the body is reconstructed, while in PET and SPECT our interest is in reconstructing the time-varying distribution of a labeled compound in the body in absolute units of radioactivity concentration.

Despite the differences between the actual physical measurements among different imaging modalities, the goal of acquiring the images in clinical environment is virtually the same—to extract the patient-specific clinical information

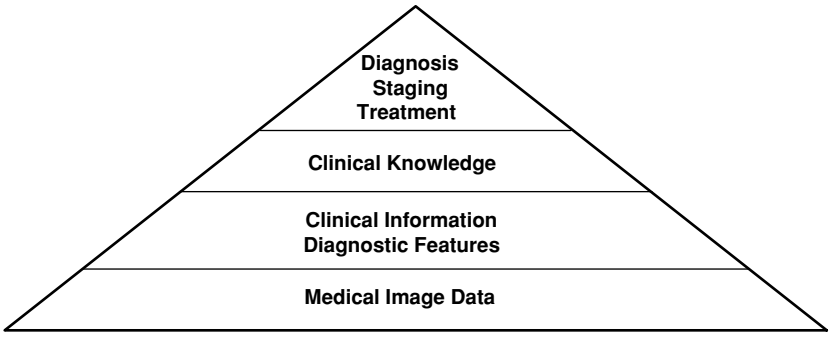

Figure 3.1: The steps and the ultimate goal of medical image analysis in a clinical environment.

and their diagnostic features embedded within the multidimensional image data that can guide and monitor interventions after the disease has been detected and localized, and ultimately leading to knowledge for clinical diagnosis, staging, and treatment of disease. These processes can be represented diagrammatically as a pyramid, as illustrated in Fig. 3.1. Starting from the bottom level of the pyramid is the medical image data obtained from a specific imaging modality, the ultimate goal (the top level of the pyramid) is to make use of the extracted information to form a set of clinical knowledge that can lead to clinical diagnosis and treatment of a specific disease. Now the question is how to reach the goal. It is obvious that the goal of the imaging study is very clear, but the solution is not. At each level of the pyramid, specific techniques are required to process the data, extract the information, label, and represent the information in a high level of abstraction for knowledge mining or to form clinical knowledge from which medical diagnosis and decision can be made. Huge amounts of multidimensional datasets, ranging from a few megabytes to several gigabytes, remain a formidable barrier to our capability in manipulating, visualizing, understanding, and analyzing the data. Effective management, processing, visualization, and analysis of these datasets cannot be accomplished without high-performance computing infrastructure composed of high-speed processors, storage, network, image display unit, as well as software programs. Recent advances in computing technology such as development of application-specific parallel processing architecture and dedicated image processing hardware have partially resolved most of the limiting factors. Yet, extraction of useful information and features from the multidimensional data is still a formidable task that requires specialized and sophisticated techniques. Development and implementation of these techniques requires

thorough understanding of the underlying problems and knowledge about the acquired data, for instance, the nature of the data, the goal of the study, and the scientific or medical interest, etc. Different assumptions about the data and different goals of the study will lead to the use of different methodologies. Therefore, continuing advances in exploitation and development of new conceptual approaches for effective extraction of all information and features contained in different types of multidimensional images are of increasingly importance in this regard.

Image segmentation plays a crucial role in extraction of useful information and attributes from images for all medical imaging applications. It is one of the important steps leading to image understanding, analysis, and interpretation. The principal goal of image segmentation is to partition an image into regions (or classes) that are homogeneous with respect to one or more characteristics or features under certain criteria [1]. Each of the regions can be separately processed for information extraction. The most obvious application of segmentation in medical imaging is anatomical localization, or in a generic term, region of interest delineation whose main aim is to outline anatomic structures and (pathologic) regions that are "of interest." Segmentation can be accomplished by identifying all pixels or voxels that belong to the same structure/region or based on some other attributes associated with each pixel or voxel. Image segmentation is not only important for feature extraction and visualization but also for image measurements and compression. It has found widespread applications in medical science, for example, localization of tumors and microcalcifications, delineation of blood cells, surgical planning, atlas matching, image registration, tissue classification, and tumor volume estimation [2–7], to name just a few.

Owing to issues such as poor spatial resolution, ill-defined boundaries, measurement noise, variability of object shapes, and some other acquisition artifacts in the acquired data, image segmentation still remains a difficult task. Segmentation of data obtained with functional imaging modalities is far more difficult than that of anatomical/structural imaging modalities, mainly because of the increased data dimensionality and the physical limitations of the imaging techniques. Notwithstanding these issues, there have been some significant progresses in this area, partly due to continuing advances in instrumentation and computer technology. It is in this context that an overview of the technical aspects and methodologies of image segmentation will be presented. As image

segmentation is a broad field and because the goal of segmentation varies according to the aim of the study and the type of the image data, it is impossible to develop only one standard method that suits all imaging applications. This chapter focuses on the segmentation of data obtained from functional imaging modalities such as PET, SPECT, and fMRI. In particular, segmentation based on cluster analysis, which has great potential in classification of functional imaging data, will be discussed in great detail. Techniques for segmentation of data obtained with structural imaging modalities have been covered in depth by other chapters of this book, and therefore, they will only be described briefly in this chapter for the purpose of completeness.

3.2 Manual Versus Automated Segmentation

As mentioned at the beginning of this chapter, detection, localization, diagnosis. staging, and monitoring treatment responses are crucial procedures in clinical medicine and oncology. Early detection and localization of the diseases and accurate disease staging could lead to changes in patient management that will impact on health outcomes. Noninvasive functional imaging is playing a key role in these issues. Accurate quantification of regional physiology depends on accurate delineation (or segmentation) of the structure or region of interest (ROI) in the images. The fundamental roles of ROI are to (1) permit quantitation, (2) reduce the dataset by focusing the quantitative analysis on the extracted regions that are of interest, and (3) establish structural correspondences for the physiological data sampled within the regions.

The most straightforward segmentation approach is to outline the ROIs manually. If certain areas in the images are of interest, the underlying tissue time–activity curve (TAC) can be extracted by putting ROIs manually around those areas. Approaches based on published anatomic atlases are also used to define ROIs. The average counts sampled over voxels in the region at different sampling intervals are then computed to compose the TAC for that region. The extracted tissue TACs are then used for subsequent kinetic analysis (Chapter 2 of *Handbook of Biomedical Image: Segmentation*, *Volume I*).

In practice, selection of ROI is tedious and time-consuming because the operator has to go through the dataset slice by slice (or even frame by frame) and choose the most representative ones from which 10–40 regions are carefully

delineated for each imaging study. Needless to say, manual ROI delineation is also operator dependent and the selected regions are subject to large intra- and interrater variability [8, 9]. Because of scatter and partial volume effects (PVEs) [10], the position, size, and shape of the ROI need careful consideration. Quantitative measurement inaccuracies exhibited by small positional differences are expected to be more pronounced for ROI delineation in the brain, which is a very heterogeneous organ and contains many small structures of irregular shape that lie adjacent to other small structures of markedly differing physiology [11]. Small positional differences can also confound the model fitting results [12, 13]. To minimize errors due to PVEs, the size of the ROI should be chosen as small as possible, but the trade-off is the increase in noise levels within the ROI, which maybe more susceptible to anatomical imprecision. On the other hand, a larger region offers a better signal-to-noise ratio but changes that occurred only within a small portion of the region maybe obscured, and the extracted TAC does not represent the temporal behavior of the ROI but a mixture of activities with adjacent overlapping tissue structures. Likewise, an irregular ROI that conforms to the shape of the structure/region where abnormality has occurred will be able to detect this change with much higher sensitivity than any other geometrically regular ROI that may not conform well. In addition, manual ROI delineation requires software tools with sophisticated graphical user interfaces to facilitate drawing ROIs and image display. Methodologies that can permit semiautomated or ideally, fully automated ROI segmentation will present obvious advantages over the manual ROI delineation.

Semiautomated or fully automated segmentation in anatomical imaging such as CT and MR is very successful, especially in the brain, as there are many well-developed schemes proposed in the literature (see surveys in [14]). This may be because these imaging modalities provide very high resolution images in which tiny structures are visible even in the presence of noise, and that four general tissue classes, gray matter, white matter, cerebrospinal fluid (CSF), and extracranial tissues such as fat, skin, and muscles, can be easily classified with different contrast measures. For instance, the T1- and T2-weighted MR images provide good contrast between gray matter and CSF, while T1 and proton density (PD) weighted MR images provide good contrast between gray matter and white matter. In contrast to CT and MRI, PET and SPECT images lack the ability to yield accurate anatomical information. The segmentation task is further complicated by poor spatial resolution and counting statistics, and patient motion

during scanning. Therefore, segmentation in PET and SPECT does not attract much interest over the last two decades, even though there has been remarkable progress in image segmentation during the same period of time. It still remains a normal practice to define ROIs manually.

Although the rationale for applying automatic segmentation to dynamic PET and SPECT images is questionable due to the above difficulties, the application of automatic segmentation as an alternative to manual ROI delineation has attracted interest recently with the improved spatial resolution of PET and SPECT systems. Automatic segmentation has advantages in that the subjectivity can be reduced and that there is saving in time for manual ROI delineation. Therefore, it may provide more consistent and reproducible results as less human intervention is involved, while the overall time for data analysis can be shortened and thereby the efficiency can be improved, which is particularly important in busy clinical settings.

3.3 Optimal Criteria for Functional Image Segmentation

Medical image segmentation is a very complicated process and the degree of complexity varies under different situations. Based on the results of a survey conducted among all centers performing emission tomographic studies and a series of international workshops to assess the goals and obstacles of data acquisition and analysis from emission tomography, Mazziotta *et al.* [15, 16] proposed a series of optimal criteria to standardize and optimize PET data acquisition and analysis:

- Reproducible
- Accurate
- Independent of tracer employed
- Independent of instrument spatial resolution
- Independent of ancillary imaging techniques
- Minimizes subjectivity and investigator bias

- Fixed assumptions about normal anatomy not required
- Acceptable to subjects' level of tolerance
- Performs well in serial studies of the same patient and individual study of separate patients in a population
- Capable of evolving toward greater accuracy as information and instruments improve
- Reasonable in cost
- Equally applicable in both clinical and research settings
- Time efficient for both data acquisition and analysis

These criteria are not specific to the functional analysis of the brain, and they are equally applicable to other organs and imaging applications upon minor modifications, in spite of the fundamentally differences between imaging modalities.

3.4 Segmentation Techniques

A large number of segmentation techniques have been proposed and implemented (see [14, 17–19]) but there is still no *gold standard* approach that satisfies all of the aforementioned criteria. In general, segmentation techniques can be divided into four major classes:

- Thresholding
- Edge-based segmentation
- Region-based segmentation
- Pixel classification

These techniques are commonly employed in two-dimensional image segmentation [1, 17–21]. A brief review of these techniques will be given in this section. More advanced techniques such as model-based approaches, multimodal approaches, and multivariate approaches, and their applications will be introduced and discussed latter in this chapter.

3.4.1 Thresholding

Semiautomatic methods can partially remove the subjectivity in defining ROIs by human operators. The most commonly used method is by means of thresholding because of its simplicity in implementation and intuitive properties. In this technique, a predefined value (threshold) is selected and an image is divided into groups of pixels having values greater than or equal to the threshold and groups of pixels with values less than the threshold. The most intuitive approach is global thresholding, which is best suited for bimodal image. When only a single threshold is selected for a given image, the thresholding is *global* to the entire image. For example, let $f(x, y)$ be an image with maximum pixel value $I_{\max}$, and suppose Γ denotes the percent threshold of the maximum pixel value above which the pixels will be selected, then pixels with value ρ given by

$$\frac{\Gamma}{100} I_{\max} \leq \rho \leq I_{\max} \tag{3.1}$$

can be grouped and a binary image $g(x, y)$ is formed:

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) \geq \rho \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

in which pixels with value of 1 correspond to the ROI, while pixels with value 0 correspond to the background.

Global thresholding is simple and computationally fast. It performs well if the images contain objects with homogeneous intensity or the contrast between the objects and the background is high. However, it may not lead itself fully automated and may fail when two or more tissue structures have overlapping intensity levels. The accuracy of the ROI is also questionable because it is separated from the data based on a single threshold value which may be subject to very large statistical fluctuations. With the increasing number of regions or noise levels, or when the contrast of the image is low, threshold selection will become more difficult.

Apart from global thresholding, there are several thresholding methods which can be classified as local thresholding and dynamic thresholding. These techniques maybe useful when a thresholding value cannot be determined from a histogram for the entire image or a single threshold cannot give good segmentation results. Local threshold can be determined by using the local statistical properties such as the mean value of the local intensity distribution or some

other statistics such as mean of the maximum or minimum values [21] or local gradient [22], or by splitting the image into subimages and calculating threshold values for the individual sub-images [23]. Some variants of the above two methods can be found in Refs. [17, 18].

3.4.2 Edge-Based Segmentation

Edge-based segmentation approaches have two major components: (1) edge detection and (2) edge linking/following to determine the edges and the regions. Loosely speaking, an edge is a collection of connected pixels that lie on the boundary between two homogeneous regions having different intensities. Therefore, edges can be defined as abrupt changes in pixel intensity that can be reflected by the gradient information. A number of edge detectors have been defined based on the first-order or second-order gradient information of the image [1, 20]. For a given image $f(x, y)$, the gradient computed at location (x, y) is defined as a vector:

$$\nabla \mathbf{f} = \begin{bmatrix} \delta f_x \\ \delta f_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix} \tag{3.3}$$

where δf_x and δf_y are gradients computed along x and y directions. The gradient vector points in the direction of maximum rate of change of f at (x, y). The magnitude and the direction of the gradient vector are given by

$$|\nabla \mathbf{f}| = \sqrt{(\delta f_x)^2 + (\delta f_y)^2} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \tag{3.4}$$

and

$$\theta = \tan^{-1}\left(\frac{\delta f_y}{\delta f_x}\right) \tag{3.5}$$

where the angle θ is measured with respect to the x axis.

In order to obtain the gradient of an image, computation of partial derivatives δf_x and δf_y at every pixel location is required. Because the images have been digitized, it is not possible to compute δf_x and δf_y using differentiation and numerical approximation of the gradient by finite difference is used instead [20]. Implementation of edge detection can be accomplished by convolving the original image with a *mask* (also called *window* or *kernel*) that runs through the

entire image. A mask is typically a 2×2 or a 3×3 matrix. For example, Roberts edge operator has two 2×2 masks:

$$\delta f_x = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \delta f_y = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

and Sobel edge operator has a pair of 3×3 masks:

$$\delta f_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \qquad \delta f_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Detailed discussion on other edge operators such as Canny, Kirsch, Prewitt, and Robinson can be found elsewhere [1, 20].

An edge magnitude image can be formed by combining the gradient components δf_x and δf_y at every pixel location using Eq. (3.4). As the computational burden required by square and square roots in Eq. (3.4) is very high, an approximation with absolute values is frequently used instead:

$$|\nabla \mathbf{f}| \approx |\delta f_x| + |\delta f_y| \tag{3.6}$$

After the edge magnitude image has been formed, a thresholding operation is then performed to determine where the edges are.

The first-order derivatives of the image $f(x, y)$ have local minima and maxima at the edges because of the large intensity variations. Accordingly, the second-order derivatives have zero crossings at the edges, which can also be used for edge detection and the Laplacian is frequently employed in practice. The Laplacian (∇^2) of a two-dimensional function $f(x, y)$ is defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{3.7}$$

There are several ways to realize the Laplacian operator in discrete-time domain. For a 3×3 region, the following two realizations are commonly used:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

It should be noted that all gradient-based edge detection methods (including the Laplacian) are very sensitive to noise because differentiation is a high pass

operation that tends to enhance image noise. In some applications, it is possible to improve the results obtained with these methods by smoothing the image prior to edge detection. For example, a smoothing filter can be applied to an image before computing the Laplacian. Marr and Hildreth [24] proposed smoothing the image with a Gaussian filter followed by the Laplacian operation to determine edge information and this operation is called Laplacian of Gaussian, which is defined as

$$\begin{aligned} h(x, y) &= \nabla^2[g(x, y) \otimes f(x, y)] \\ &= \nabla^2[g(x, y)] \otimes f(x, y) \end{aligned} \tag{3.8}$$

where $f(x, y)$ is the original image, $\otimes$ is the convolution operator, $g(x, y)$ is a Gaussian function, and $\nabla^2[g(x, y)]$ is the Laplacian of Gaussian function that is used for spatial smoothing of the original image. Edges can be determined from the resultant image, $h(x, y)$, by simply thresholding it for zero value to detect zero crossing. Equation (3.8) represents a generic operation of taking the Laplacian on the spatial smoothing filter, $g(x, y)$, which can be replaced by other filter function (e.g., directional low-pass filter [25]) to improve the performance of edge detection in a specific application. Faber *et al.* [26] applied the Laplacian edge detection technique to segment scintigraphic images and the results were promising.

In practice, edge detection techniques produce only a series of edges for the structures/areas of interest. It is not uncommon that the edge pixels do not characterize an edge and that the edges do not enclose the ROI completely because of noise and some other acquisition artifacts that caused spurious discontinuities of edges. Therefore, the second component of edge-based segmentation techniques is to track and link the edge pixels to form meaningful edges or closed regions in the edge image obtained by one of the edge detection techniques. One of the simplest approaches for linking edge pixels is to analyze local characteristics of pixels within a small block of pixels (e.g., 3×3 or 5×5) for the entire edge image and linked all edge pixels that are similar to each others according to some predefined criteria. The Hough transform [27] can also be applied to detect straight lines and parametric curves as well as arbitrarily shaped objects in the edge image. It was shown by Deans [28] that the Hough transform is a special case of the Radon transform for image projection and reconstruction [29]. Thorough discussion and comparison of different varieties of the Hough transform and their generalizations are considered beyond the

scope of this chapter and they can be found in Refs. [30, 31]. There are several more powerful edge tracking/linking techniques such as graph searching [32,33] and dynamic programming [34, 35] that perform well in the presence of noise. As might be expected, these paradigms are considerably more complicated and computationally expensive than the methods discussed so far.

3.4.3 Region-Based Segmentation

Region-based segmentation approaches examine pixels in an image and form disjoint regions by merging neighborhood pixels with homogeneity properties based on a predefined similarity criterion. Suppose that I represents an image that is segmented into N regions, each of which is denoted as R_i where $i = 1, 2, \ldots, N$, the regions must satisfy the following properties:

$$I = \bigcup_{i=1}^{N} R_i \tag{3.9}$$

$$R_i \cap R_j = 0 \;\; \forall i, j = 1, 2, \ldots, N; \;\; i \neq j \tag{3.10}$$

$$\mathcal{L}(R_i) = \text{TRUE} \quad \text{for } i = 1, 2, \ldots, N \tag{3.11}$$

$$\mathcal{L}(R_i \cup R_j) = \text{FALSE} \quad \forall i, j = 1, 2, \ldots, N; \;\; i \neq j \tag{3.12}$$

where $\mathcal{L}(\cdot)$ is a logical predicate. The original image can be exactly assembled by putting all regions together (Eq. 3.9) and there should be no overlapping between any two regions R_i and R_j for $i \neq j$ (Eq. 3.10). The logical predicate $\mathcal{L}(\cdot)$ contains a set of rules (usually a set of homogeneity criteria) that must be satisfied by all pixels within a given region (Eq. 3.11), and it fails in the union of two regions since merging two distinct regions will result in an inhomogeneous region (Eq. 3.12).

The simplest region-based segmentation technique is the *region growing*, which is used to extract a connected region of similar pixels from an image [36]. The region growing algorithm requires a similarity measure that determines the inclusion of pixels in the region and a stopping criterion that terminates the growth of the region. Typically, it starts with a pixel (or a collection of pixels) called *seed* that belongs to the target ROI. The seed can be chosen by the operator or determined by an automatic seed finding algorithm. The neighborhood of each seed is then inspected and pixels similar enough to the seed are added to the corresponding region where the seed is, and therefore, the region is growing and its shape is also changing. The growing process is repeated until no pixel

can be added to any region. It is possible that some pixels may remain unlabeled when the growing process stops.

Hebert *et al.* [37] investigated the use of region growing to automated delineation of the blood pool with computer simulations and applied the method to three gated SPECT studies using Tc-99m pertechnetate, and the results were promising. Kim *et al.* [38] also investigated an integrated approach of region growing and cluster analysis (to be described later) to segment a dynamic [^{18}F]fluorodeoxyglucose (FDG) PET dataset. Although qualitatively reasonable segmentation results were obtained, much more work is needed to overcome the difficulties in the formation of odd segments possibly due to spillover region boundaries, and evaluate the quantitative accuracy of the segmentation results using kinetic parameter estimation.

Region splitting methods take an opposite strategy to the region growing. These methods start from the entire image and examine the homogeneity criteria. If the criteria do not meet, the image (or subimage) is split into two or more subimages. The region splitting process continues until all subimages meet the homogeneity criteria. Region splitting can be implemented by quadtree partitioning. The image is partitioned into four subimages that are represented by nodes in a quadtree, which is a data structure used for efficient storage and representation. The partition procedure is applied recursively on each subimage until each and all of the subimages meet the homogeneity criteria.

The major drawback of region splitting is that the final image may contain adjacent regions R_i and R_j, which are homogeneous, i.e. $\mathcal{L}(R_i \cup R_j) = \text{TRUE}$, and ideally this region should be merged. This leads to another technique called *split-and-merge*, which includes a merging step in the splitting stage, where an inhomogeneous region is split until homogeneous regions are formed. A newly created homogeneous region is checked against its neighboring regions and merged with one or more of these regions if they possess identical properties. However, this strategy does not necessarily produce quadtree partitioning of the image. If quadtree partitioning is used, an additional step may be added to merge adjacent regions (nodes) that meet the homogeneity criterion.

3.4.4 Pixel Classification

Recall that the key step of thresholding techniques described in section 3.4.1 is the choice of thresholds that is determined either manually or in a

semiautomatic manner based on the local statistics such as mean, maximum, or minimum of the given image (or subimages). The basic concept of threshold selection can be generalized, leading to a data-driven paradigm, which determines the threshold automatically based on clustering techniques or artificial neural networks.

Pixel classification methods that use histogram statistics to define single or multiple thresholds to classify an image can be regarded as a generalization of thresholding techniques. It is particularly useful when the pixels have multiple features, which can be expressed in terms of a vector in multidimensional feature space. For instance, the feature vector may consist of gray level, local texture, and color components for each pixel in the image. In the case of single-channel (or single-frame) image, pixel classification is typically based on gray level and image segmentation can be performed in a one-dimensional feature space. Segmentation can be performed in multidimensional feature space through clustering of all features of interest for multichannel (multiple-frame) images or multispectral (multimodality) images.

Clustering, or cluster analysis, has been widely applied in anthropology, archaeology, psychiatry, and zoology, etc, for many years. An example of clustering is the taxonomy of animals and plants whose names have to be the same between different people for effective communication, although it is not necessary that the naming scheme is the best [39]. Clustering is the process of grouping of similar objects into a single cluster, while objects with dissimilar features are grouped into different clusters based on some similarity criteria. The similarity is quantified in terms of an appropriate distance measure. An obvious measure of the similarity is the distance between two vectors in the feature space which can be expressed in terms of L_p norm as

$$d\{\mathbf{x}_i, \mathbf{x}_j\} = \left(\sum_{k=1}^{n} \| \mathbf{x}_i - \mathbf{x}_j \|^p \right)^{\frac{1}{p}} \tag{3.13}$$

where $\mathbf{x}_i \in \mathbb{R}^n$ and $\mathbf{x}_j \in \mathbb{R}^n$ are the two vectors in the feature space. It can be seen that the above measure corresponds to Euclidean distance when $p = 2$ and Mahalanobis distance when $p = 1$. Another commonly used distance measure is the normalized inner product between two vectors given by

$$d\{\mathbf{x}_i, \mathbf{x}_j\} = \frac{\mathbf{x}_i^T \mathbf{x}_j}{\|\mathbf{x}_i\| \cdot \|\mathbf{x}_j\|} \tag{3.14}$$

where T denotes the transpose operation. The above distance measure is simply the angle between vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ in the feature space.

Each cluster is represented by its centroid (or mean) and variance, which indicates the compactness of the objects within the cluster, and the formation of clusters is optimized according to a cost function that typically takes the similarity within individual cluster and dissimilarity between clusters into account. There are many clustering techniques proposed in the literature (see Ref. [39]). The most famous clustering techniques are K-means [40], fuzzy c-means [41], ISODATA [42], hierarchical clustering with average linkage method [43], and Gaussian mixture approach [44].

As we will see later in this chapter, the idea of pixel classification in two-dimensional image segmentation using clustering techniques can be extended to multidimensional domain where the images convey not only spatial information of the imaged structures but also their temporal variations, for which clustering plays a pivotal role in identification of different temporal kinetics present in the data, extraction of blood and tissue TACs, ROI delineation, localization of abnormality, kinetic modeling, characterization of tissue kinetics, smoothing, and fast generation of parametric images.

3.5 Advanced Segmentation Techniques

Functional imaging with PET, SPECT, and/or dynamic MRI provides *in vivo* quantitative measurements of physiologic parameters of biochemical pathways and physiology in a noninvasive manner. A critical component is the extraction of physiological data, which requires accurate localization/segmentation of the appropriate ROIs. A common approach is to identify the anatomic structures by placing ROIs directly on the functional images, and the underlying tissue TACs are then extracted for subsequent analysis. This ROI analysis approach, although widely used in clinical and research settings, is operator-dependent and thus prone to reproducibility errors and it is also time-consuming. In addition, this approach is problematic when applied to small structures because of the PVEs due to finite spatial resolution of the imaging devices.

Methods discussed so far can be applied to almost all kinds of image segmentation problem because they do not require any model (i.e. model-free) that guides or constrains the segmentation process. However, segmenting structures

of interest from functional images is difficult because of the imprecise anatomical information, the complexity and variability of anatomy shapes and sizes within and across individuals, and acquisition artifact, such as spatial aliasing, and insufficient temporal sampling, noise, and organ/patient movements. All these factors can hamper the boundary detection process and cause discontinuous or indistinguishable boundaries. Model-free approaches usually generate ambiguous segmentation results under these circumstances, and considerable amounts of human intervention are needed to resolve the ambiguity in segmentation. In this section, some advanced segmentation approaches are introduced, including

- *model-based* segmentation techniques that use analytical models to describe the shape of the underlying ROI,
- *multimodal techniques* that integrate information available from different imaging modalities for segmentation, or the image measurements are transformed and mapped to a standard template, and
- *multivariate approaches* are data-driven techniques in which the structures are identified and extracted based on the temporal information present in the dynamic images.

3.5.1 Model-Based Segmentation

Incorporation of *a priori* knowledge of the object such as shape, location, and orientation using *deformable models* (also known as *active contour models*) is one of the possible solutions to constrain the segmentation of organ structures. The term *deformable models* was coined by Terzopoulos and his collaborators [45,46] in the 1980s, but the idea of using a deformable template for feature extraction dated back to the work of Fischler and Elschlager on spring-loaded templates [47] and the work of Widrow on rubber mask technique [48] in the early 1970s. Deformable models are analytically or parametrically defined curves or surfaces that move under the influence of forces, which have two components: *internal forces* and *external forces*. The internal forces are used to assure the smoothness of the model during deformation process and the external forces are defined to push/pull the model toward the boundaries of the structure. Parametric representations of the models allow accurate and compact description

of the object shape, while the continuity, connectivity, and smoothness of the models can compensate for the irregularities and noise in the object boundaries. Model-based approaches treat the problem of finding object boundaries as an optimization problem of searching the best fit for the image data to the model. In the case of boundary finding via optimization in image space, a fairly extensive review on various deformable model methods can be found in Ref. [49].

Mykkänen *et al.* [50] investigated automatic delineation of brain structures in FDG-PET images using generalized snakes with promising results. Chiao *et al.* [51] proposed using model-based approach for segmenting dynamic cardiac PET or SPECT data. The object model consists of two parts: a heart and the rest of the body. The heart is geometrically modeled using a polygonal model [52] and the myocardial boundaries are parameterized by the endocardial radii and a set of angular thicknesses. Kinetic parameters in the compartment model and the endocardial and epicardial radii are estimated by maximizing a joint log-likelihood function using nonlinear parameter estimation. Tissue and blood TACs are extracted simultaneously with estimated kinetic parameters. Chiao *et al.* [51] proposed that some forms of regularization can be applied, including auxiliary myocardial boundary measurements obtained by MRI or CT and registering the auxiliary measurements with the emission tomographic data, if the kinetic parameter estimation failed.

3.5.2 Multimodal Techniques

Comparisons of datasets obtained from individual subjects between imaging modalities are very important for the evaluation of the normal physiologic responses of the anatomic structure or the pathophysiological changes that accompany disease states. Likewise, it is also critical to compare data between individuals both within and across different imaging modalities. Unfortunately, many structures of interest, particularly in the brain, are often smaller than the spatial resolution of the imaging devices and corrections aided by anatomical imaging modalities such as CT and MR are often required [53, 54].

Anatomic structures, particularly those in the brain, can also be identified using a standardized reference coordinate system or functional image data can be fitted to a standard anatomical atlas (e.g., Talairach space) with the aid of anatomical landmarks or contours [55–58]. This idea is somewhat similar to the model-based approaches where analytically or parametrically defined models

are used to segment the organ boundaries. The difference lies in the definition of the model, which is described by a computerized anatomy atlas or a stereotaxic coordinate system—a reference that the functional images are mapped onto by either linear or nonlinear transformation. A number of transformation techniques have been developed for this process [59]. The ROIs defined on the template are then available to the functional image data.

Similarly, functional (PET and SPECT) images and structural (CT and MR) images obtained from individual subjects can be fused (coregistered), allowing precise anatomical localization of activity on the functional images [60, 61]. Precise alignment between the anatomic/template and PET images is necessary for these methods. Importantly, methods that use registration to a standard coordinate system are problematic when patients with pathological processes (e.g., tumors, infarction, and atrophy) are studied.

3.5.3 Multivariate Segmentation

The main aim of dynamic imaging is to study the physiology (function) of the organ *in vivo*. Typically the image sequence has constant morphologic structures of the imaged organs but the regional voxel intensity varies from one frame to another, depending on the local tissue response to the administered contrast agent or radiopharmaceutical. In the past, analysis of such dynamic images involved only visual analysis of differences between the early and delayed images from which qualitative information about the organ, for instance, regional myocardial blood flow and distribution volume are obtained. However, the sequence of dynamic images also contain spatially varying quantitative information about the organ which is difficult to extract solely based on visual analysis. This led to the method of parametric imaging where dynamic curves in the image sequence are fit to a mathematical model on a pixel-wise basis. Parametric images whose pixels define individual kinetic parameters or physiologic parameters that describe the complex biochemical pathways and physiologic/pharmacokinetic processes occurred within the tissue/organ can then be constructed. This approach is categorized as *model-led* technique that utilizes knowledge and *a priori* assumptions of the processes under investigation, and represents the kinetics of the measured data by an analytical (or parametric) model.

At the opposite end of the spectrum of model-led techniques are *data-driven* techniques, which are based on the framework of multivariate data analysis.

These paradigms minimize the *a priori* assumptions of the underlying processes whose characteristics are interrogated from the measured data, independent of any kinetic model. Multivariate approaches have been explored and successfully applied in a number of functional imaging studies. The aims borne in mind when applying these approaches are to (1) classify structures present in the images, (2) extract information from the images, (3) reduce the dimensionality of data, (4) visualize the data, and (5) model the data, all of which are crucial in data analysis, medical research, data reduction, and treatment planning. In general, the underlying structures are identified and extracted based on the temporal information present in the sequence of the medical images. The implicit assumptions for the validity of applying these approaches are that the statistical noise present in the images is uncorrelated (independent) between different frames and that there is a high degree of correlation (similarity) between tissue TACs if they are originated from similar structures. In this section, we focus our attention on four techniques among many different multivariate analysis approaches and their applications in dynamic, functional imaging are discussed.

3.5.3.1 Similarity Mapping

In this section, we introduce an intuitive temporal segmentation technique called *similarity mapping* (or *correlation mapping*), which was proposed by Rogowska [62]. This approach identifies regions according to their temporal similarity or dissimilarity with respect to a dynamic curve obtained from a reference region. Consider a sequence of N spatially registered time-varying images $\mathbf{X}$ of size $M \times N$, with M being the number of pixels in one image and N the number of frames. Then each row of $\mathbf{X}$ represents a pixel vector, i.e., a time-intensity curve as stated in Rogowska's paper [62] (also called a *dixel* [63] or a tissue TAC in PET/SPECT or fMRI studies—it is just a matter of nomenclature!) which is a time series

$$\mathbf{x}_i = [X_i(t_1), X_i(t_2), \ldots, X_i(t_N)]^T \tag{3.15}$$

where t_j ($j = 1, 2, \ldots, N$) represents the time instant at which the jth frame is acquired, $X_i(t_j)$ is the pixel value of the ith element evaluated in the jth frame of $\mathbf{X}$, for $j = 1, 2, \ldots, N$, and T denotes the transpose operation.

Similar to the pixel classification technique described earlier in section 3.4.4, some quantitative index is necessary to measure the similarity between

time-intensity curves for different pixels or the mean of the pixel values averaged over a selected ROI. Suppose a ROI is drawn on a reference region in the dynamic sequence of images and its time course is extracted

$$\mathbf{r} = [r(t_1), r(t_2), \ldots, r(t_N)]^T \tag{3.16}$$

The similarity between the reference time-intensity curve $\mathbf{r}$ and the time-intensity curves for all pixels can then be calculated. And a similarity map, which is a image where the value of each pixel shows the temporal similarity to the reference curve, can be constructed.

Since the time instants do not affect the computation of cross correlation between two time-intensity curves as pixel intensity values in one frame are measured at the same time, $\mathbf{x}_i$ in Eq. (3.15) and $\mathbf{r}$ in Eq. (3.16) can be rewritten in a time-independent form as

$$\mathbf{x}_i = [X_{i,1}, X_{i,2}, \ldots, X_{i,N}]^T \tag{3.17}$$

where $X_{i,j} \equiv X_i(t_j)$ is the pixel value of the ith element evaluated in the jth frame of $\mathbf{X}$, and

$$\mathbf{r} = [r_1, r_2, \ldots, r_N]^T \tag{3.18}$$

whose mean intensity value is given by

$$\overline{r} = \frac{1}{N}\sum_{j=1}^{N} r_j \tag{3.19}$$

The similarity map $\mathbf{R}$ based on normalized cross correlation can be defined for each pixel i as

$$R_i = \frac{\sum_{j=1}^{N}\left(X_{i,j} - \overline{X}_i\right)\left(r_j - \overline{r}\right)}{\sqrt{\sum_{j=1}^{N}\left(X_{i,j} - \overline{X}_i\right)^2 \sum_{j=1}^{N}\left(r_j - \overline{r}\right)^2}} \tag{3.20}$$

where

$$\overline{X}_i = \frac{1}{N}\sum_{j=1}^{N} X_{i,j} \tag{3.21}$$

is the mean value of the time sequence for pixel i. The normalized cross correlation has values in the range of -1 to $+1$. Regions of identical temporal variation have a coefficient of $+1$, with the exception that $\mathbf{x}_i$ or $\mathbf{r}$ are extracted from regions of constant pixel intensity (e.g. background). In this case, the denominator of Eq. (3.20) equals zero. Therefore, the following restrictions have to be

imposed on the computation of the normalized cross correlation:

$$\sum_{j=1}^{N} \left(X_{i,j} - \overline{X}_i\right)^2 \neq 0 \quad \text{and} \quad \sum_{j=1}^{N} \left(r_j - \overline{r}\right)^2 \neq 0 \tag{3.22}$$

Time-intensity curves similar to the reference curve will have high-correlation values and are bright in the similarity map, whereas those with low-correlation values are dark. Therefore, structures in the dynamic image sequence can be segmented from the similarity map based on their temporal changes rather than spatial similarities. It should be noted that cross-correlation does not depend on the absolute magnitude of the time-intensity curves. Regions whose time-intensity curves are differed from the reference curve **r** by an additive or by a multiplicative constant, will have a perfect positive correlation (+1).

By using different reference ROIs, a series of similarity maps containing different segmentation for regions that have similar or different temporal kinetics can be obtained. It was used to investigate local changes and segmentation of rabbit kidney on spatially aligned image sequences obtained from dynamic MR imaging of Gd-DTPA [64]. The similarity mapping technique has also been applied to brain activation studies to extract the activated regions and their temporal dynamics [65]. The same technique has also been used to segment the area of ischemia in the left coronary artery territory, lung tumor, and tentorial meningioma, and localize the focal ischemic region in brain [66].

3.5.3.2 Principal Component Analysis

Principal component analysis (PCA), also called *Karhunen–Loéve transform* or *Hotelling transform*, is probably the most famous method in multivariate data analysis [67]. It was developed independently by Pearson [68] in 1901 and Hotelling [69] in 1933. It has been widely applied in a number of scientific areas such as biology, chemistry, medicine, psychology, and the behavioral sciences. Given a set of multivariate data, PCA explains the variance–covariance structure by linearly transforming the (possibly) correlated variables into a smaller set of uncorrelated (orthogonal) variables called *principal components* (PCs). The first (highest order) component maximally accounts for the variation in the original data and each succeeding component maximally accounts for the remaining variation present in the original data. In other words, higher order components are important as they explain the major variation (also the feature) in

the data, whereas lower order components are unimportant as they mainly contain noise, which can be discarded without causing too much loss of information of the original data. Therefore, dimensionality reduction (or data compression) can be achieved using PCA technique. Separation of tissue types characterized by different features can also be accomplished by careful inspection of the PCs. This is because each PC contains only the representative feature that is specific to that PC and cannot be found elsewhere (theoretically) owing to orthogonality among PCs.

Let the dynamic sequence of images be represented by a matrix $\mathbf{X}$ that has M rows and N columns. Each column represents a time frame of image data and each row represents a pixel vector, i.e., a tissue TAC or a *dixel* [63], which is a time series $\mathbf{x}_i$ as in Eqs. (3.15) and (3.17). Note that there is no explicit assumption on the probability density of the measurements $\mathbf{x}_i$ as long as the first-order and second-order statistics are known or can be estimated from the available measurements. Each of $\mathbf{x}_i$ can be considered as a random process

$$\mathbf{x} = [x_1, x_2, \ldots, x_N]^T \tag{3.23}$$

If the measurements (or random variables) x_j are correlated, their major variations can be accurately approximated with less than N parameters using the PCA. The mean of $\mathbf{x}$ is given by

$$\overline{x} = E\{\mathbf{x}\} \tag{3.24}$$

and the covariance matrix of the same dataset is given by

$$\mathbf{C}_\mathbf{x} = E\{(\mathbf{x} - \overline{x})(\mathbf{x} - \overline{x})^T\} \tag{3.25}$$

which is a $N \times N$ symmetric matrix. The elements of $\mathbf{C}_\mathbf{x}$, denoted by c_{kl}, represent the covariances between the random variables x_k and x_l, whereas the element c_{kk} is the variance of the random variable x_k. If x_k and x_l are uncorrelated, their covariance would be zero, i.e., $c_{kl} = c_{lk} = 0$. The mean and the covariance matrix of a sample of random vectors $\mathbf{x}_i$ can be estimated from its sample mean and sample covariance matrix in a similar manner.

The orthogonal basis of the covariance matrix $\mathbf{C}_\mathbf{x}$ can be calculated by finding its eigenvalues and eigenvectors. It is well known from basic linear algebra that the eigenvectors $\mathbf{e}_k$ and the corresponding eigenvalues λ_k are the solutions of

the equation

$$\mathbf{C}_\mathbf{x}\mathbf{e}_k = \lambda_k \mathbf{e}_k \tag{3.26}$$

for $k = 1, 2, \ldots, N$ and $\lambda_k \neq 0$. There are several numerical methods to solve for λ_k and $\mathbf{e}_k$ in Eq. (3.26). One of the popular approaches is to make use of the symmetrical property of $\mathbf{C}_\mathbf{x}$ and solve for the eigenvalues and eigenvectors by means of Householder reduction followed by QL algorithm with implicit shifts [70, 71]. As $\mathbf{C}_\mathbf{x}$ is a real, symmetric matrix, an equivalent approach is to compute the singular value decomposition (SVD) of the matrix $\mathbf{C}_\mathbf{x}$ directly:

$$\mathbf{C}_\mathbf{x} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T \tag{3.27}$$

where $\mathbf{U}$ is a $N \times N$ column-orthogonal matrix, $\mathbf{V}$ is a $N \times N$ orthogonal matrix that contains the eigenvectors, and $\mathbf{\Lambda}$ is a $N \times N$ diagonal matrix whose squared diagonal elements correspond to the eigenvalues. However, the difference between SVD and the eigen-decomposition should be noted, in particular, the eigen-decomposition of a real matrix might be complex, whereas the SVD of a real matrix is always real.

The ordering of the eigenvectors can be sorted in the order of descending eigenvalues such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \geq 0$. In this way, an ordered orthogonal basis is formed, and the first eigenvector $\mathbf{e}_1$ (the one associated with λ_1) has the direction of the largest variance of the data (the first PC), and the second eigenvector $\mathbf{e}_2$ has the direction of the second largest variance of the data (the second PC), and so on. The PCs are obtained by projecting the multivariate random vectors onto the space spanned by the eigenvectors. Let $\mathbf{\Omega}$ be a matrix that stores the eigenvectors $\mathbf{e}_k$ as row vectors, then the PCs, $\mathbf{y} = [y_1, y_2, \ldots, y_N]^T$, can be calculated as

$$\mathbf{y} = \mathbf{\Omega}(\mathbf{x} - \overline{x}) \tag{3.28}$$

which defines a linear transformation for the random vector $\mathbf{x}$ through the orthogonal basis and $\overline{x}$ is calculated from Eq. (3.24). The kth PC of $\mathbf{x}$ is given by

$$y_k = \mathbf{e}_k^T(\mathbf{x} - \overline{x}) \tag{3.29}$$

which has zero mean. The PCs are also orthogonal (uncorrelated) to one another because

$$E\{y_k y_l\} = E\left\{\left(\mathbf{e}_k^T(\mathbf{x}-\overline{x})\right)\left(\mathbf{e}_l^T(\mathbf{x}-\overline{x})\right)\right\} = \mathbf{e}_k^T\mathbf{C}_\mathbf{x}\,\mathbf{e}_l = 0 \tag{3.30}$$

for $k > l$. The original random vector $\mathbf{x}$ can be reconstructed from $\mathbf{y}$ by

$$\mathbf{x} = \mathbf{\Omega}^T\mathbf{y} + \overline{x} \tag{3.31}$$

where $\mathbf{\Omega}^{-1} = \mathbf{\Omega}^T$ since $\mathbf{\Omega}$ is an orthogonal matrix.

The variances of the PCs can be computed as follows:

$$E\{y_k^2\} = E\left\{\left(\mathbf{e}_k^T(\mathbf{x}-\overline{x})\right)\left(\mathbf{e}_k^T(\mathbf{x}-\overline{x})\right)\right\} = \mathbf{e}_k^T\mathbf{C}_\mathbf{x}\,\mathbf{e}_k = \lambda_k \tag{3.32}$$

which indicates that the variances of the PCs are given by the eigenvalues of $\mathbf{C}_\mathbf{x}$. As the PCs have zero means, a very small eigenvalue (variance) λ_k implies that the value of the corresponding PC is also very small to contribute to the total variances present in the data. Since the eigenvalue sequence $\{\lambda_k\}$ is monotonically decreasing and typically the sequence drops rapidly, it is possible to determine a limit below which the eigenvalues (and PCs) can be discarded without causing significant error in reconstruction of the original dataset using only the retained PCs. Thus, data compression (or dimensionality reduction) can be achieved and this is an important application of PCA. Instead of using all eigenvectors of the covariance matrix $\mathbf{C}_\mathbf{x}$, the random vector $\mathbf{x}$ can be approximated by the highest few basis vectors of the orthogonal basis. Suppose that only the first K rows (eigenvectors) of $\mathbf{\Omega}$ are selected to form a $K \times N$ matrix, $\mathbf{\Omega}_K$, a similar transformation as in Eqs. (3.28) and (3.31) can be derived

$$\tilde{\mathbf{y}} = \mathbf{\Omega}_K(\mathbf{x} - \overline{x}) \tag{3.33}$$

and

$$\hat{\mathbf{x}} = \mathbf{\Omega}_K^T\mathbf{y} + \overline{x} \tag{3.34}$$

where $\tilde{\mathbf{y}}$ represents a truncated PC vector, which contains only the K highest PCs, and $\hat{\mathbf{x}}$ is an approximation of $\mathbf{x}$ with the K highest PCs. It can be shown that the mean square error (MSE) between $\hat{\mathbf{x}}$ and $\mathbf{x}$ is given by

$$E\{\|\hat{\mathbf{x}} - \mathbf{x}\|^2\} = \sum_{k=K+1}^{N} \lambda_k \tag{3.35}$$

The practical issue here is the choice of K beyond which the PCs are insignificant. The gist of the problem lies in how "insignificant" is defined and how much error one could tolerate in using less number of PCs to approximate the original data. Sometimes, a small number of PCs are sufficient to give an accurate approximation to the observed data. A commonly used strategy is to plot the eigenvalues against the number of PCs and detect a cut-off beyond which the eigenvalues become constants. Another approach is to discard the PCs with eigenvalues lower than a specified fraction of the first (largest) eigenvalue. There is no simple answer and one has to trade off between errors and the number of PCs for approximation of the observed data which is the primary concern when PCA is used for data compression.

PCA has been applied to analyze functional images including nuclear medicine [72–77] and dynamic MRI [78, 79] where data visualization, structure and functional classification, localization of diseases, and detection of activation patterns are of primary interests. Moeller and Strother [72] applied PCA to analyze functional activation patterns in brain activation experiments. Strother *et al.* [75] revealed an intra- and intersubject subspace in data and demonstrated that the activation pattern is usually contained in the first PC. A later study conducted by Ardekani *et al.* [76] further demonstrated that the activation pattern may spread across several PCs rather than lie only on the first PC, particularly when the number of subjects increases and/or multicenter data are used. PCA was also applied to aid interpretation of oncologic PET images. Pedersen *et al.* [74] applied PCA to aid analyze of dynamic FDG-PET liver data. Anzai *et al.* [77] investigated the use of PCA in detection of tumors in head and neck, also using dynamic FDG-PET imaging. It was found that the first few highest order component images often contained tumors whereas the last several components were simply noise.

Absolute quantification of dynamic PET or SPECT data requires an invasive procedure where a series of blood samples are taken to form an input function for kinetic modeling (Chapter 2 of *Handbook of Biomedical Image Analysis: Segmentation*, *Volume I*). Sampling blood at the radial artery or from an arterialized vein in a hand is the currently recognized method to obtain the input function. However, arterial blood sampling is invasive and has several potential risks associated with both the patient and the personnel who performed the blood sampling [80]. Drawing ROI around vascular structures (e.g., left ventricle in the myocardium [81] and internal carotid artery in the brain [82]) has been

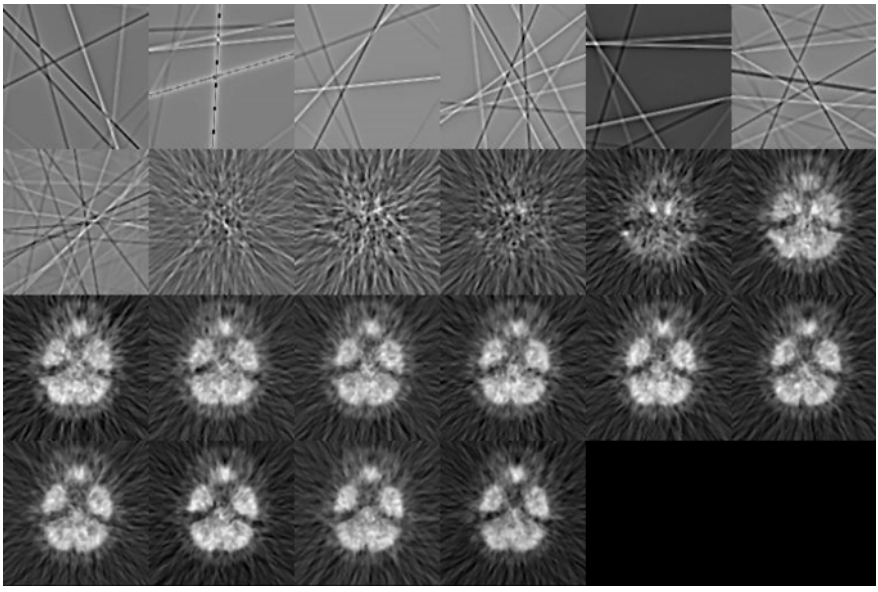

Figure 3.2: A sequence of dynamic neurologic FDG-PET images sampled at the level where the internal carotid arteries are covered. Individual images are scaled to their own maximum.

proposed as a noninvasive method that obviates the need of frequent blood sampling. Delineation of larger vascular structures in the myocardium is relatively straightforward. In contrast, delineation of internal carotid artery in the head and neck is not trivial. A potential application of PCA is the extraction of input function from the dynamic images in which vascular structures are present in the dynamic images. Figure 3.2 show a sequence of dynamic neurologic FDG-PET images sampled at the level in the head where the internal carotid arteries are covered. Figure 3.3 shows the highest 12 PC images. The signal-to-noise ratio (SNR) of the first PC image is very high when comparing with the original image sequence. For PC images beyond the second, they simply represent the remaining variability that the first two PC images cannot account for and they are dominated by noise. The internal carotid arteries can be seen in the second PC image which can be extracted by means of thresholding as mentioned before in sections 3.4.1 and 3.4.4. Figure 3.4 shows a plot of percent contribution to the total variance for individual PCs. As can be seen from the figure, the first and the second PCs contribute about 90% and 2% of the total variance, while the remaining PCs only contribute for less than 0.6% of the total variance individually.

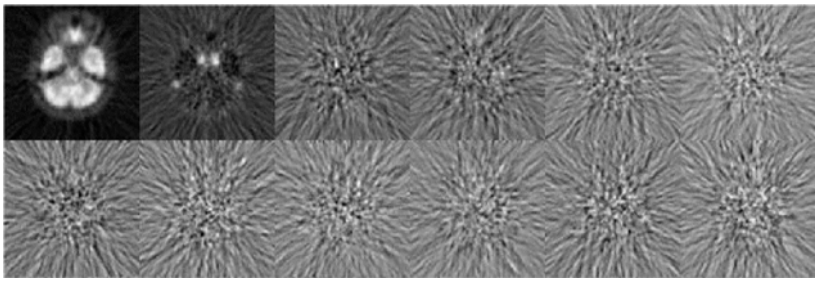

Figure 3.3: From left to right, the figure shows the first six principal component (PC) images (top row), and the 7th to 12th PC images (bottom row) scaled to their own maxima. All but the first two PC images are dominated by noise. The higher order PC images (not shown) look very similar to PC images 3–12.

This means that large amount of information (about 92%) is preserved in only the first two PCs, and the original images can be approximated by making use only the first one or two PCs.

Different from model-led approaches such as compartmental analysis where the physiological parameters in a hypothesized mathematical model are estimated by fitting the model to the data under certain possibly invalid assumptions, PCA is data-driven, implying that it does not rely on a mathematical model.

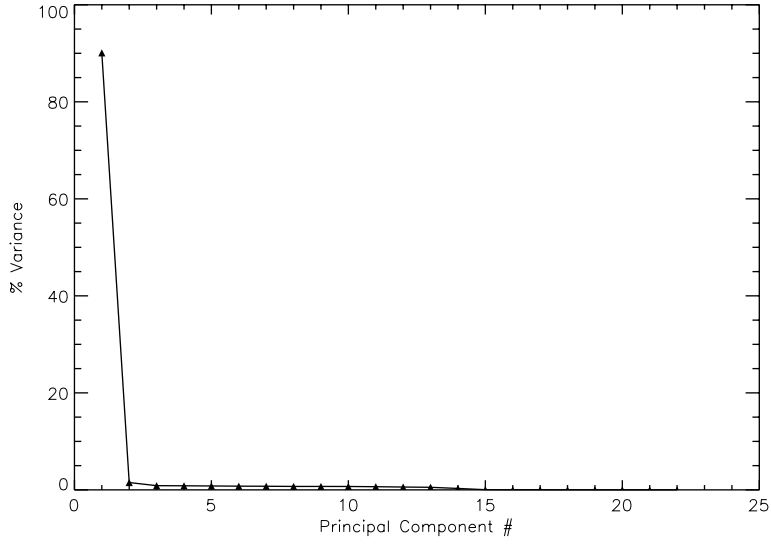

Figure 3.4: The percent variance distribution of the principal component (PC) images.

Instead, it explores the variance–covariance structure of the observed data and finds a set of optimal (uncorrelated) PCs, each of which contains maximal variation present in the measured data. A linear combination of these components can accurately represent the observed data. However, because of lack of model as a constraint, PCA cannot separate signals from statistical noise, which may be an important component if it is highly correlated and dominates the multivariate data. In this case, convincing results of dimensionality reduction or structure exploration may not be achievable as noise is still retained in the higher order components. In addition, the orthogonal components produced by PCA are not necessarily physiological meaningful. Thus, it is difficult to relate the extracted components to the underlying TACs and structures in the multivariate data.

3.5.3.3 Factor Analysis

Factor analysis of dynamic structures (FADS), or *factor analysis* (FA), can be thought of as a generalization of PCA as it produces factors closer to the true underlying tissue response and assumes a statistical model for the observed data. FADS is a semiautomatic technique used for extraction of TACs from a sequence of dynamic images. FADS segments the dynamic sequence of images into a number of structures which can be represented by functions. Each function represents one of the possible underlying physiological kinetics such as blood, tissue, and background in the sequence of dynamic images. Therefore, the whole sequence of images can be represented by a weighted sum of these functions.

Consider a sequence of dynamic images $\mathbf{X}$ of size $M \times N$, with M being the number of pixels in one image and N the number of frames. Each row of $\mathbf{X}$ represents a pixel vector, which is a tissue TAC in PET/SPECT data. Assume that pixel vectors in $\mathbf{X}$ can be represented by a linear combination of factors $\mathbf{F}$, then $\mathbf{X}$ can written as

$$\mathbf{X} = \mathbf{CF} + \boldsymbol{\eta} \tag{3.36}$$

where $\mathbf{C}$ contains factor coefficients for each pixel and it is of size $M \times K$ with K being the number of factors; $\mathbf{F}$ is a $K \times N$ matrix which contains underlying tissue TACs. The additive term $\boldsymbol{\eta}$ in Eq. (3.36) represents measurement noise in $\mathbf{X}$.

Similar to the mathematical analysis detailed before for similarity mapping and PCA, we define $\mathbf{x}_i$ as the ith pixel vector in $\mathbf{X}$, and $\mathbf{f}_k$ being the kth underlying

factor curve (TAC), and c_{ki} being the factor coefficient that represents contribution of the kth factor curve to $\mathbf{x}_i$. Let $\mathbf{Y} = \mathbf{CF}$ and $\mathbf{y}_i$ be a vector which represents the ith row of $\mathbf{Y}$, then

$$\mathbf{y}_i = \sum_{k=1}^{K} c_{ki}\mathbf{f}_k \tag{3.37}$$

and

$$\mathbf{x}_i = \mathbf{y}_i + \boldsymbol{\eta}_i \tag{3.38}$$

where $\boldsymbol{\eta}_i$ represents a vector of noise associated with $\mathbf{x}_i$. Simply speaking, these equations mean that the series of dynamic images $\mathbf{X}$ can be approximated and constructed from some tissue TACs of the underlying structures (represented by a factor model $\mathbf{Y} = \mathbf{CF}$), which are myocardial blood pools, myocardium, liver, and background for cardiac PET/SPECT imaging, for example. The aim of FADS is to project the underlying physiological TACs, $\mathbf{y}_i$ as close as possible to the measured TACs, $\mathbf{x}_i$, so that the MSE between them can be minimized:

$$\Phi(\mathbf{C}, \mathbf{F}) = \sum_{i=1}^{M} \left(\mathbf{x}_i - \sum_{k=1}^{K} c_{ki}\mathbf{f}_k \right)^2 \tag{3.39}$$

Typically, FADS proceeds by first identifying an orthogonal basis for the sequence of dynamic images followed by an oblique rotation. Identification of the orthogonal basis can be accomplished by PCA discussed previously. However, the components identified by PCA are not physiologically meaningful because some components must contain negative values in order to satisfy the orthogonality condition. The purpose of oblique rotation is to impose nonnegativity constraints on the extracted factors (TACs) and the extracted images of factor coefficients [63].

As mentioned in section 3.2, careful ROI selection and delineation are very important for absolute quantification, but manually delineation of ROI is not easy due to high-noise levels present in the dynamic images. Owing to scatter and patient volume effects, the selected ROI may represent "lumped" activities from different adjacent overlapping tissue structures rather than the "pure" temporal behavior of the selected ROI. On the other hand, FADS can separate partially overlapping regions that have different kinetics, and thereby, extraction of TACs corresponding to those overlapping regions is possible.

Although the oblique rotation yield reasonable nonnegative factor curves that are highly correlated with the actual measurements, they are not unique [83] because both factors and factor coefficients are determined simultaneously. It is very easy to see this point by a simple example. Assume that a tissue TAC $\mathbf{x}$ composes of only two factors $\mathbf{f}_1$ and $\mathbf{f}_2$ and c_1 and c_2 being the corresponding factor coefficients. According to Eqs. (3.36) and (3.37), $\mathbf{x}$ can be represented by

$$\mathbf{x} = c_1\mathbf{f}_1 + c_2\mathbf{f}_2 \tag{3.40}$$

which can be written as

$$\mathbf{x} = c_1(\mathbf{f}_1 + \alpha\mathbf{f}_2) + (c_2 - \alpha c_1)\mathbf{f}_2 \tag{3.41}$$

with some constant α. It can be seen that Eqs. (3.40) and (3.41) are equivalent for describing the measured TAC, $\mathbf{x}$, as long as $\mathbf{f}_1 + \alpha\mathbf{f}_2$ and $c_2 - \alpha c_1$ are nonnegative if nonnegativity constraints have to be satisfied. In other words, there is no difference to represent $\mathbf{x}$ using factors $\mathbf{f}_1$ and $\mathbf{f}_2$ and factor coefficients c_1 and c_1, or using factors $\mathbf{f}_1 + \alpha\mathbf{f}_2$ and $\mathbf{f}_2$ and factor coefficients c_1 and $c_2 - \alpha c_1$. Therefore, further constraints such as *a priori* information of the data being analyzed are required [84–87].

FADS has been successfully applied to extract the time course of blood activity in left ventricle from PET images by incorporating additional information about the input function to be extracted [88, 89]. Several attempts have also been made to overcome the problem of nonuniqueness [90, 91]. It was shown that these improved methods produced promising results in a patient planar ^{99m}Tc-MAG$_3$ renal study and dynamic SPECT imaging of ^{99m}Tc-teboroxime in canine models using computer simulations and measurements in experimental studies [90, 91].

3.5.3.4 Cluster Analysis

Cluster analysis has been described briefly in section 3.4.4. One of the major aims of cluster analysis is to partition a large number of objects according to certain criteria into a smaller number of clusters that are mutually exclusive and exhaustive such that the objects within a cluster are similar to each others, while objects in different clusters are dissimilar. Cluster analysis is of potential value in classifying PET data, because the cluster centroids (or centers) are derived

from many objects (tissue TACs) and an improved SNR can be achieved [92]. It has been applied to segment a dynamic [^{11}C]flumazenil PET data [92] and dynamic [^{123}I]iodobenzamide SPECT images [93]. In the following, a clustering algorithm is described. Its application to automatic segmentation of dynamic FDG-PET data for tumor localization and detection is demonstrated in the next section. An illustration showing how to apply the algorithm to generate ROIs automatically for noninvasive extraction of physiological parameters will also be presented.

The segmentation method is based on cluster analysis. Our aim is to classify a number of tissue TACs according to their shape and magnitude into a smaller number of distinct characteristic classes that are mutually exclusive so that the tissue TACs within a cluster are similar to one another but are dissimilar to those drawn from other clusters. The clusters (or clustered ROIs) represent the locations in the images where the tissue TACs have similar kinetics. The kinetic curve associated with a cluster (i.e. cluster centroid) is the average of TACs in the cluster. Suppose that there exists k characteristic curves in the dynamic PET data matrix, $\mathbf{X}$, which has M tissue TACs and N time frames with $k \ll M$ and that any tissue TAC belongs to only one of the k curves. The clustering algorithm then segments the dynamic PET data into k curves automatically based on a weighted least-squares distance measure, $\mathcal{D}$, which is defined as

$$\mathcal{D}\{\mathbf{x}_i, \boldsymbol{\mu}_j\} = \sum_{j=1}^{k} \sum_{i=1}^{M} \|\mathbf{x}_i - \boldsymbol{\mu}_j\|^2_{\mathbf{W}} \tag{3.42}$$

where $\mathbf{x}_i \in \mathbb{R}^N$ is the ith tissue TAC in the data, $\boldsymbol{\mu}_j \in \mathbb{R}^N$ is the centroid of cluster C_j, and $\mathbf{W} \in \mathbb{R}^{N \times N}$ is a square matrix containing the weighting factors on the diagonal and zero for the off-diagonal entries. The weighting factors were used to boost the degree of separation between any TACs that have different uptake patterns but have similar least-squares distances to a given cluster centroid. They were chosen to be proportional to the scanning intervals of the experiment. Although this is not necessarily an optimal weighting, reasonably good clustering results can be achieved.

There is no explicit assumption on the structure of data and the clustering process proceeds automatically in an unsupervised manner. The minimal assumption for the clustering algorithm is that the dynamic PET data can be represented by a finite number of kinetics. As the number of clusters, k, for a given dataset is usually not known *a priori*, k is usually determined by trial and error.

In addition, the initial cluster centroid in each cluster is initialized randomly to ensure that all clusters are nonempty. Each tissue TAC is then allocated to its nearest cluster centroid according to the following criterion:

$$\begin{aligned} & \|\mathbf{x}_l - \boldsymbol{\mu}_i\|^2_{\mathbf{W}} < \|\mathbf{x}_l - \boldsymbol{\mu}_j\|^2_{\mathbf{W}} \\ \Rightarrow \quad & \mathbf{x}_l \in C_i \qquad \forall i, j = 1, 2, \ldots, k, \quad i \neq j \end{aligned} \tag{3.43}$$

where $\mathbf{x}_l \in \mathbb{R}^N$ is the lth tissue TAC in $\mathbf{X}$; $\boldsymbol{\mu}_i \in \mathbb{R}^N$ and $\boldsymbol{\mu}_j \in \mathbb{R}^N$ are the ith and jth cluster centroid, respectively; and C_i represents the ith cluster set. The centroids in the clusters are updated based on Eq. (3.43) so that Eq. (3.42) is minimized. The above allocation and updating processes are repeated for all tissue TACs until there is no reduction in moving a tissue TAC from one cluster to another. On convergence, the cluster centroids are mapped back to the original data space for all voxels. An improved SNR can be achieved because each voxel in the mapped data space is represented by one of the cluster centroids each of which possesses a higher statistical significance than an individual TAC.

Convergence to a global minimum is not always guaranteed because the final solution is not known *a priori* unless certain constraints are imposed on the solution that may not be feasible in practice. In addition, there may be several local minima in the solution space when the number of clusters is large. Restarting the algorithm with different initial cluster centroids is necessary to identify the best possible minimum in the solution space.

The algorithm is similar to the K-means type Euclidean clustering algorithm [40]. However, the K-means type Euclidean clustering algorithm requires that the data are normalized and it does not guarantee that the within-cluster cost is minimized since no testing is performed to check whether there is any cost reduction if an object is moved from one cluster to another.

3.6 Segmentation of Dynamic PET Images

The work presented in this section builds on our earlier research in which we applied the proposed clustering algorithm to tissue classification and segmentation of phantom data and a cohort of dynamic oncologic PET studies [94]. The study was motivated by our on-going work on a noninvasive modeling approach

for quantification of FDG-PET studies where several ROIs of distinct kinetics are required [95, 96]. Manual delineation of ROIs restrain the reproducibility of the proposed modeling technique, and therefore, some other semiautomated and automated methods have been investigated and clustering appears as a promising alternative to automatically segment ROI of distinct kinetics. The results indicated that the kinetic and physiological parameters obtained with cluster analysis are similar to those obtained with manual ROI delineation, as we will see in the later sections.

3.6.1 Experimental Studies

3.6.1.1 Simulated [^{11}C]Thymidine PET Study

To examine the validity of the segmentation scheme, we simulated a dynamic 2-[^{11}C]thymidine (a marker of cell proliferation) PET study. 2-[^{11}C]thymidine was chosen because it is being increasingly used in the research setting to evaluate cancer and treatment response, and it offers theoretical advantages over FDG such as greater specificity in the assessment of malignancy. Also, the kinetics are very similar for most tissues and the data are typically quite noisy. Thus, thymidine data represent a challenging example for testing the clustering algorithm.

Typical 2-[^{11}C]thymidine kinetics for different tissues were derived from eight patients. The data were acquired on an ECAT 931 scanner (CTI/Siemens, Knoxville, TN). The dynamic PET data were acquired over 60 min with a typical sampling schedule (10×30 sec, 5×60 sec, 5×120 sec, 5×180 sec, 5×300 sec) and the tracer TAC in blood was measured with a radial artery catheter following tracer administration. Images were reconstructed using filtered back-projection (FBP) with a Hann filter cut-off at the Nyquist frequency. ROIs were drawn over the PET images to obtain tissue TACs in bone, bone marrow, blood pool, liver, skeletal muscle, spleen, stomach, and tumor. Impulse response functions (IRFs) corresponding to these tissues were determined by spectral analysis of the tissue TACs [97]. The average IRFs for each common tissue type were obtained by averaging the spectral coefficients across the subjects and convolved with a typical arterial input function, resulting in typical TACs for each tissue. The TACs were then assigned to the corresponding tissue types in a single slice of the Zubal phantom [98] which included blood vessels, bone, liver, bone marrow,

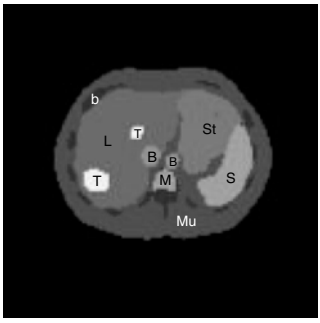

Figure 3.5: A slice of the Zubal phantom. B = blood vessels; b = bone; L = liver; M = marrow; Mu = muscle; S = spleen; St = stomach; T = tumor.

muscle, spleen, stomach, a large and small tumor in the liver (see Fig. 3.5). A dynamic sequence of sinograms was obtained by forward projecting the images into 3.13 mm bins on a 192 $\times$ 256 grid. Attenuation was included in the simulations for the purpose of obtaining the correct scaling of the noise. Poisson noise and blurring were added to simulate realistic sinograms. Noisy dynamic images were then reconstructed using FBP (Hann filter cut-off at the Nyquist frequency). Figure 3.6 shows the metabolite-corrected arterial blood curve and noisy 2-[^{11}C]thymidine kinetics in some representative tissues.

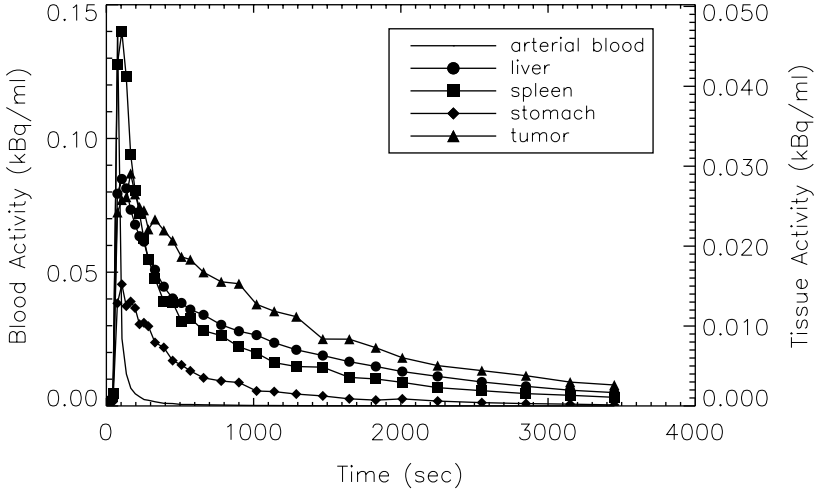

Figure 3.6: Simulated noisy 2-[^{11}C]thymidine kinetics in some representative regions. A metabolite-corrected arterial blood curve, which was used to simulate 2-[^{11}C]thymidine kinetics in different tissues, is also shown.

Figure 3.7: A slice of the Hoffman brain phantom. A tumor in white matter (white spot) and an adjacent hypometabolic region (shaded region) are shown.

3.6.1.2 Simulated FDG-PET Study

Dynamic FDG-PET study was simulated using a slice of the numerical Hoffman brain phantom [99] that modified using a template consisting of five different kinetics (gray matter, white matter, thalamus, tumor in white matter, and an adjacent hypometabolic region in left middle temporal gyrus), as shown in Fig. 3.7. The activities in gray matter and white matter were generated using a five-parameter three-compartment FDG model [100] with a measured arterial input function obtained from a patient (constant infusion of 400 MBq of FDG over 3 min). The kinetics present in the hypometabolic region, thalamus, and tumor were set to 0.7, 1.1, and 2.0 times the activity in gray matter. The kinetics were then assigned to each brain region and a dynamic sequence of sinograms (22 frames, 6×10 sec, 4×30 sec, 1×120 sec, 11×300 sec) was obtained by forward projecting the images into 3.13 mm bins on a 192×256 grid. Poisson noise and blurring were also added to simulate realistic sinograms. Dynamic images were reconstructed using FBP with Hann filter cut-off at the Nyquist frequency. The noisy FDG kinetics are shown in Fig. 3.8 and some of the kinetics are similar to each other due to the added noise and gaussian blurring, although their kinetics are different in the absence of noise and blurring. This is illustrated in the white matter and the hypometabolic region, and the gray matter and thalamus.

3.6.2 Cluster Validation

As mentioned earlier, the optimum number of clusters for a given dataset is usually not known *a priori*. It is advantageous if this number can be determined

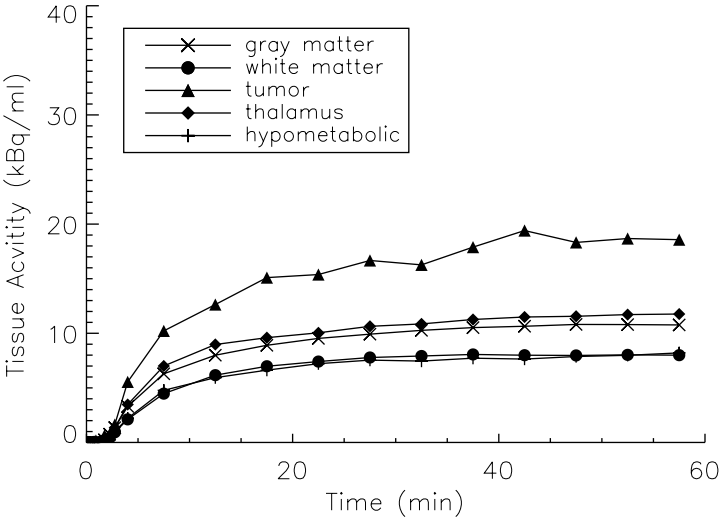

Figure 3.8: Simulated noisy[^{18}F]fluorodeoxyglucose (FDG) kinetics in different regions.

based on the given dataset. In this study, a model-based approach was adopted to cluster validation based on two information-theoretic criteria, namely, Akaike information criterion (AIC) [101] and Schwarz criterion (SC) [102], assuming that the data can be modeled by an appropriate probability distribution function (e.g. Gaussian). Both criteria determine the optimal model order by penalizing the use of a model that has a greater number of clusters. Thus, the number of clusters that yields the lowest value for AIC and/or SC is selected as the optimum. The use of AIC and SC has some advantages compared to other heuristic approaches such as the "bootstrap" resampling technique which requires a large amount of stochastic computation. This model-based approach is relatively flexible in evaluating the goodness-of-fit and a change in the probability model of the data does not require any change in the formulation except the modeling assumptions. It is noted, however, that both criteria may not indicate the same model as the optimum [102].

The validity of clusters is also assessed visually and by thresholding the average mean squared error (MSE) across clusters, which is defined as

$$\text{MSE} = \frac{1}{k} \sum_{j=1}^{k} \sum_{i=1}^{M} \|\mathbf{x}_i - \boldsymbol{\mu}_j\|_{\mathbf{W}}^2. \tag{3.44}$$

Both approaches are subjective but they can provide an insight into the "correct" number of clusters.

3.6.3 Human Studies

The clustering algorithm has been applied to a range of FDG-PET studies and three examples (two patients with brain tumor and one patient with a lung cancer) are presented in this chapter. FDG-PET was chosen to assess the clustering algorithm because it is commonly used in clinical oncologic PET studies. All oncological PET studies were performed at the Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, Sydney, Australia. Ethical permissions were obtained from the Institutional Review Board.

Dynamic neurologic FDG-PET studies were performed on an ECAT 951R whole-body PET tomograph (CTI/Siemens, Knoxville, TN). Throughout the study the patient's eyes were patched and ears were plugged. The patients received 400 MBq of FDG, infused at a constant rate over a 3-min period using an automated injection pump. At least 30 min prior to the study, patient's hands and forearms were placed into hot water baths preheated to 44 °C to promote arterio-venous shunting. Blood samples were taken at approximately 30 sec for the first 6 min, and at approximately 8, 10, 15, 30, and 40 min, and at the end of emission data acquisition. A dynamic sequence of 22 frames was acquired for 60 min following radiotracer administration according to the following schedule: 6×10 sec, 4×30 sec, 1×2 min, 11×5 min. Data were attenuation corrected with a postinjection transmission method [103]. Images were reconstructed on a 128×128 matrix using FBP with a Shepp and Logan filter cut-off at 0.5 of the Nyquist frequency.

The dynamic lung FDG-PET study was commenced after intravenous injection of 487 MBq of FDG. Emission data were acquired on an ECAT 951R whole-body PET tomograph (CTI/Siemens, Knoxville, TN) over 60 min (22 frames, 6×10 sec, 4×30 sec, 1×2 min, and 11×5 min). Twenty one arterial blood samples were taken from the pulmonary artery using a Grandjean catheter to provide an input function for kinetic modeling.

The patient details are as follows:

Patient 1: The FDG-PET scan was done in a female patient, 6 months after resection of a malignant primary brain tumor in the right parieto-occipital

lobe. The scan was done to determine if there was evidence for tumor recurrence. A partly necrotic hypermetabolic lesion was found in the right parieto-occipital lobe that was consistent with tumor recurrence.

Patient 2: A 40-year-old woman had a glioma in the right mesial temporal lobe. The FDG-PET scan was performed at 6 months after tumor resection. A large hypermetabolic lesion was identified in the right mesial temporal lobe that was consistent with tumor recurrence.

Patient 3: A 67-year-old man had an aggressive mesothelioma in the left lung. In the PET images, separate foci of increased FDG uptake were seen in the contralateral lymph nodes as well as in the peripheral left lung.

As they are unnecessary for clustering and the subsequent analysis, low count areas such as the background (where the voxel values should be zero theoretically) and streaks (which are due to reconstruction errors) were excluded by zeroing voxels whose summed activity was below 5% of the mean pixel intensity of the integrated dynamic images. A 3×3 closing followed by a 3×3 erosion operation was then applied to fill any "gap" inside the intracranial/body region to which cluster analysis was applied. Parametric images of the physiological parameter, K, which is defined as the value of $k_1^* k_3^* / (k_2^* + k_3^*)$ [104], were generated by fitting all voxels inside the intracranial/body region using Patlak graphical approach [105]. The resultant parametric images obtained for the raw dynamic images and dynamic images after cluster analysis were assessed visually. Compartmental model fitting using the three-compartment FDG model [104] was also performed on the tissue TACs extracted manually and by cluster analysis to investigate whether there is any disagreement between the parameter estimates.

3.6.4 Results

3.6.4.1 Simulated [^{11}C]Thymidine PET Study

Figure 3.9 shows the segmentation results using different numbers of clusters, k, in the clustering algorithm. The number of clusters is actually varied from 3 to 13 but only some representative samples are shown. In each of the images in Figs. 3.9(a)–3.9(f), different gray levels are used to represent the cluster locations. Figure 3.9 shows that when the number of clusters is small, segmentation

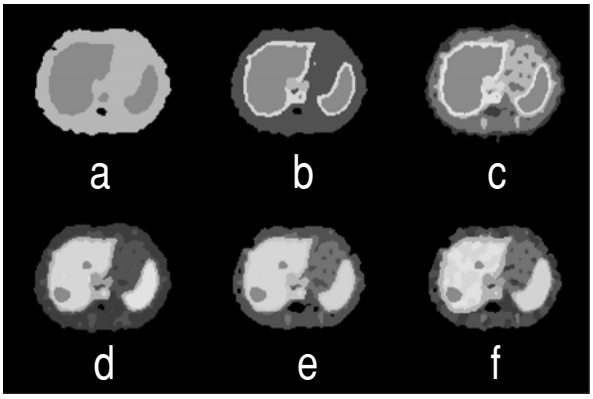

Figure 3.9: Tissue segmentation obtained with different number of clusters. (a) $k = 3$, (b) $k = 5$, (c) $k = 7$, (d) $k = 8$, (e) $k = 9$, and (f) $k = 13$. (Color Slide)

of the data is poor. With $k = 3$, the liver, marrow, and spleen merge to form a cluster and the other regions merge to form a single cluster. With $5 \leq k \leq 7$, the segmentation results improve because the blood vessels and stomach are visualized. However, the hepatic tumors are not seen and the liver and spleen are classified into the same cluster. With $k = 8$, the tumors are visualized and almost all of the regions are correctly identified (Fig. 3.9(d)). Increasing the value of k to 9 gives nearly the same segmentation as in the case of $k = 8$ (Fig. 3.9(e)). Further increasing the value of k, however, may result in poor segmentation because the actual number of tissues present in the data is less than the specified number of clusters. Homogeneous regions are therefore fragmented to satisfy the constraint on the number of clusters (Fig. 3.9(f)). Thus, 8 or 9 clusters appear to provide reasonable segmentation of tissues in the slice and this number agrees with the various kinetics present in the data.

Figure 3.10 plots the average MSE across clusters as a function of k. The average MSE decreases monotonically, as it drops rapidly ($k < 8$) before reaching a plateau ($k \geq 10$). From the trend of the plot, there is no significant reduction in the average MSE with $k > 12$. Furthermore, the decrease in the average MSE is nearly saturated with $k \geq 8$. These results confirm the findings of the images in Fig. 3.9, suggesting 8 or 9 as the optimal number of clusters for this dataset.

Table 3.1 tabulates the results of applying AIC and SC to determine the optimum number of clusters which is the one that gives the minimum value for

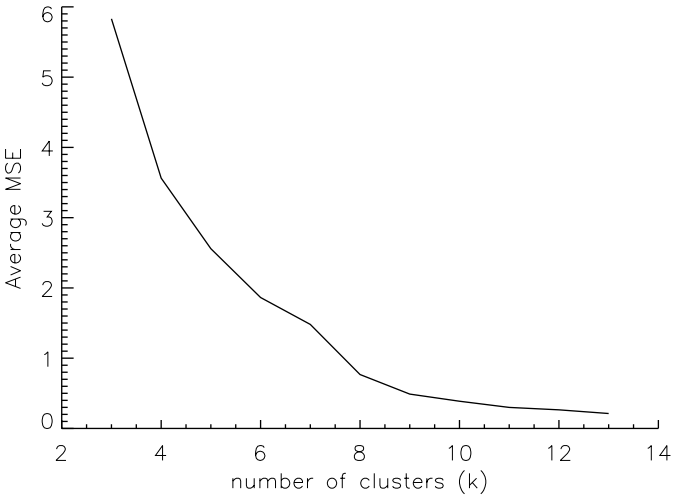

Figure 3.10: Average mean squared error (MSE) as a function of number of clusters.

the criteria. Both criteria indicate that $k = 8$ is an optimal approximation to the underlying number of kinetics. It was found that a good segmentation can be achieved when the number of clusters is the same as that determined by the criteria. Conversely, the segmentation result is poor when the number of clusters is smaller than that suggested by the criteria and there is no significant improvement in segmentation when the number of clusters is larger than that determined by the criteria. The heuristic information given by both criteria also support our visual interpretation of the clustering results, suggesting that the criteria are reasonable approaches to objectively determine the number of clusters.

Table 3.1: Computed values for AIC and SC with different choices of the value of k

	Number of clusters, k										
Criterion	3	4	5	6	7	8	9	10	11	12	13
AIC	99005	95354	93469	90904	88851	**86967**[a]	89769	93038	91994	90840	89807
SC	98654	94888	92887	90206	88038	**86038**[a]	88725	91878	90719	89450	88301

AIC: Akaike information criterion; SC: Schwarz criterion.

[a] Values in bold correspond to the computed minimum of the criterion.

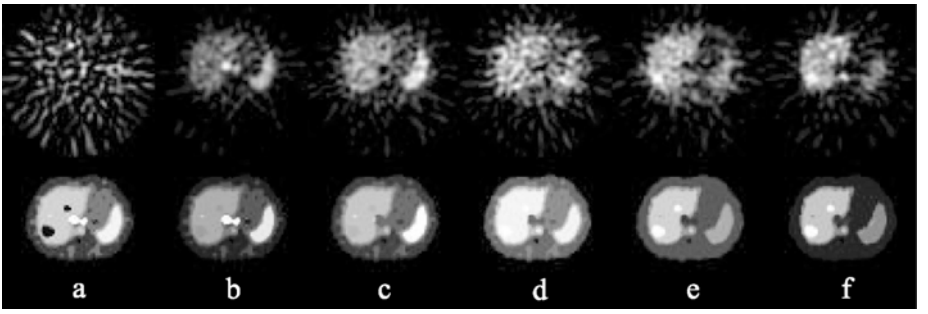

Figure 3.11: Single slice of simulated 2-[^{11}C]thymidine PET study. Top row shows the original reconstructed images at (a) 15 sec, (b) 75 sec, (c) 135 sec, (d) 285 sec, (e) 1020 sec, and (f) 2850 sec postinjection. Bottom row shows same slice at identical time points after cluster analysis. Individual images are scaled to their own maximum.

Application of the clustering algorithm to the simulated PET data is shown in Fig. 3.11. The number of clusters is eight, corresponding to the optimum number of clusters determined by the statistical criteria. The SNR of the images is markedly improved after clustering. In addition, the blood vessels are clearly seen in the frame sampled at 15 and 75 sec after clustering but not in the corresponding frame in the original data. In the original images, it is difficult to identify different tissues which may be due to reconstruction effects and inhomogeneous noise. However, the liver, spleen, muscle, marrow, stomach, and tumors are clearly delineated by the clustering algorithm (bottom row of the figure).

3.6.4.2 Simulated FDG-PET Study

Five cluster images were generated by applying the clustering algorithm to the noisy simulated dynamic images. The number of clusters k was actually varied from 3 to 10 and the optimal k was determined by inspecting the change of average MSE and the visual quality of the cluster images. Figure 3.12 shows the cluster images for $k = 5$ that was found to be the optimum number of clusters for this dataset. It was found that the tumor cannot be located when k was small ($k < 4$). The tumor was located by gradually increasing the number of clusters. However, there was a deteriorated segmentation of all regions when k was large ($k > 7$).

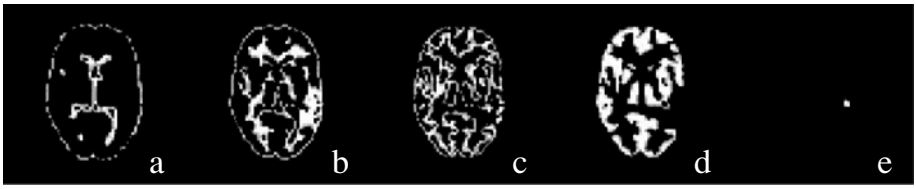

Figure 3.12: Five cluster images obtained from the noisy simulated dynamic FDG-PET data. The images correspond to (a) ventricles and scalp, (b) white matter and left middle temporal gyrus (hypometabolic zone), (c) partial volume between gray matter and white matter, (d) gray matter, deep nuclei, and outer rim of tumor, and (e) tumor.

Although the tumor was small in size, cluster analysis was still able to locate it because of its abnormal temporal characteristics as compared to the other regions. Cluster analysis also performed well in extracting underlying tissue kinetics in gray matter and white matter because of their distinct kinetics. On the contrary, the kinetics in the thalamus and the hypometabolic region were not separated from those in gray and white matter but this was not unexpected since their kinetics were very similar.

Owing to the partial volume effects (PVEs), there were some vague regions whose kinetics were indeterminate (Fig. 3.12(c)) and did not approach gray or white matter. The algorithm was unlikely to assign such kinetics to the cluster corresponding to white matter or to the cluster corresponding to gray matter since the overall segmentation results would be deteriorated. Thus, a cluster was formed to account for the indeterminate kinetics.

3.6.4.3 Human Studies

Segmentation results are shown for dynamic neurologic (Fig. 3.13) and lung (Fig. 3.14) FDG-PET studies. The clusters are represented by differing gray scales and slices were sampled at the level where the lesions were seen on the original reconstructed data. Since there is no *a priori* knowledge about the optimum number of clusters, the value of k was varied in order to determine the optimal segmentation using the AIC and SC as in the phantom study. For Fig. 3.13, eight clusters were found to give the optimal segmentation for these datasets. The locations of the tumors and the rim of increased glucose uptake

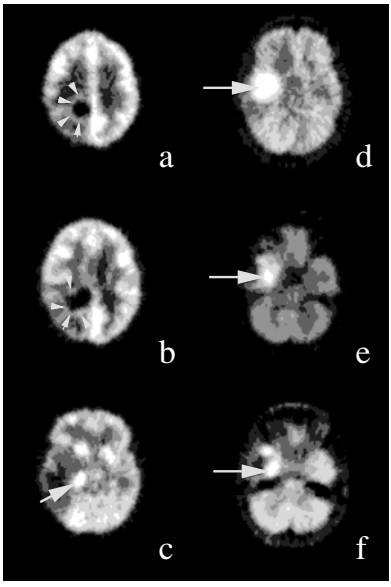

Figure 3.13: Tissue segmentation obtained from Patient 1 at (a) slice 10, (b) slice 13, and (c) slice 21; and Patient 2 at (d) slice 21, (e) slice 24, and (f) slice 26. The number of clusters used is eight. The locations of the solid hypermetabolic portions of the tumors (arrows) and the small rim of increased glucose uptake (arrow heads) identified by cluster analysis are shown.

are identified correctly by the clustering algorithm with the optimal value of clusters.

For Fig. 3.14, the number of clusters was varied from 3 to 13 and only some representative results are shown. Similar to the simulation study, the segmentation results are poor when the number of clusters is small ($k = 3$), while the segmentation is gradually improved by increasing the number of clusters. Based on the AIC and SC, the optimum numbers of clusters for the selected slices (4, 19, and 24) were found to be 8, 8, and 9, respectively. It is not surprising that the optimum number of clusters is different for different slices because of the differing number of anatomical structures contained in the plane and the heterogeneity of tracer uptake in tissues. Nevertheless, the tumor (slice 4), right lung and muscle (slices 4, 19, and 24), blood pool (slices 4, 19, and 24), separate foci of increased FDG uptake (slices 19 and 24), and the injection site (slices 4, 19, and 24) are identifiable with the optimum number of clusters.

Figure 3.15 shows the measured blood TAC at the pulmonary artery and the extracted tissue TACs for the tumor (from slice 4), lung and muscle (from

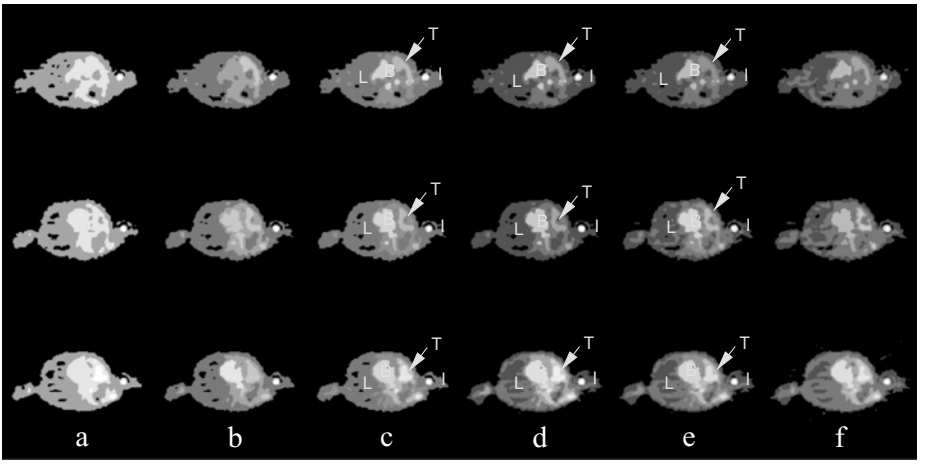

Figure 3.14: Tissue segmentation of the dynamic lung FDG-PET data from Patient 3 in three selected slices: 4 (top row), 19 (middle row), and 24 (bottom row) with different number of clusters. (a) $k = 4$, (b) $k = 7$, (c) $k = 8$, (d) $k = 9$, (e) $k = 10$, and (f) $k = 12$. (I = injection site; B = blood pool; L = lung; T = tumor).

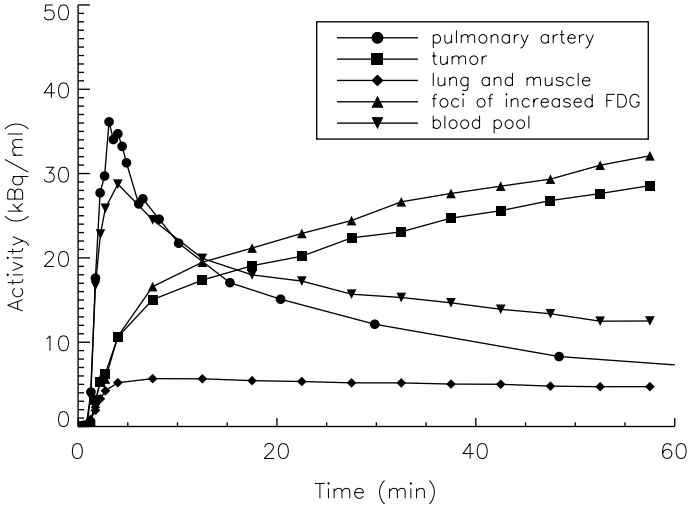

Figure 3.15: Extracted tissue time–activity curve (TACs) corresponding to the tumor, lung, and muscle, foci of increased FDG uptake, and blood pool. The measured blood TAC at the pulmonary artery is also shown.

Table 3.2: Compartmental modeling of the tumor TACs obtained by manually ROI delineation and by cluster analysis

Parameter	Manual delineation	Cluster analysis
k_1^* (ml/min/g)	0.854 ± 17.1	0.921 ± 18.2
k_2^* (min^{-1})	1.987 ± 21.2	2.096 ± 22.5
k_3^* (min^{-1})	0.099 ± 0.9	0.100 ± 0.9
k_4^* (min^{-1})	0.018 ± 1.1	0.017 ± 1.2
K (ml/min/g)	0.041 ± 5.3	0.042 ± 5.3

TAC: Time–activity curve; ROI: region of interest. Values are given as estimate ±% CV.

slice 19), foci of increased FDG uptake (from slice 24), and the blood pool (from slice 19) using the corresponding optimal value of clusters.

The extracted tissue TACs obtained by cluster analysis and manual ROI delineation were fitted to the three-compartment FDG model using nonlinear least squares method and the results obtained for the tumor tissue TAC (Patient 2) are summarized in Table 3.2. There was a close agreement between the parameter estimates for the tissue TACs obtained by different methods in terms of the estimate and the coefficient of variation (CV), which is defined as the ratio of the standard deviation of the parameter estimate to the value of the estimate. Similar results were also found for other regions.

3.7 Extraction of Physiological Parameters and Input Function

Quantification of dynamic PET or SPECT data requires an invasive procedure where a series of blood samples are taken to form an input function for kinetic modeling. One of the potential applications of the clustering algorithm presented earlier is in noninvasive quantitative PET. We have proposed a simultaneous estimation approach to estimate the input function and physiological parameters simultaneously with two or more ROIs and our results with *in vivo* PET data are promising [95]. The method is still limited, however, by the selection of ROIs whose TACs must have distinct kinetics. As the ROIs are drawn manually on the PET images, reproducibility is difficult to achieve. The use of clustering to extract tissue TACs of distinct kinetics has been investigated in three

Table 3.3: Comparison between the estimated input functions obtained using different number of *manually drawn* ROIs and *clustered* ROIs, and the measured input functions

	Number of ROIs			
	2	3	4	5
Manually drawn ROIs				
MSE	0.632	0.365	0.431	0.967
AUC (measured = 24.077)	23.796	23.657	24.188	25.138
Linear regression on AUC ($n = 19$)				
Slope	0.967	0.963	0.984	1.022
Intercept	0.493	0.460	0.609	0.712
r value	0.999	0.999	0.999	0.999
Clustered ROIs				
MSE	0.100	0.096	0.040	0.066
AUC (measured = 24.077)	20.742	23.721	25.430	23.481
Linear regression on AUC ($n = 19$)				
Slope	0.807	0.953	1.067	0.946
Intercept	0.874	0.575	-0.321	0.481
r value	0.993	0.998	0.999	0.999

MSE = Mean square errors between the estimated and the measured input functions; AUC = area under the blood curve; r = coefficient of correlation; ROI = region of interest.

FDG-PET studies. Table 3.3 summarizes the results for the estimation of the input functions by the proposed modeling approach for both manually drawn ROIs and clustered ROIs. The MSE between the estimated and the measured input functions are tabulated. In addition, results of linear regression analysis on the areas under the curves (AUCs) covered by the measured and the estimated input functions are listed for comparison. Regression lines with slopes close to unity and intercepts close to zero were obtained in all cases for manually drawn ROIs and for clustered ROIs.

Figure 3.16 plots the measured input function and the estimated input functions for manually drawn ROIs and clustered ROIs, respectively. The estimated input functions were obtained by simultaneously fitting with three ROIs of distinct kinetics. There was a very good agreement between the estimated input functions and the measured blood curve, in terms of the shape and the peak time estimation at which the peak occurs, despite the overestimation of the peak value. Thus, cluster analysis may be useful as a preprocessing step before our noninvasive modeling technique.

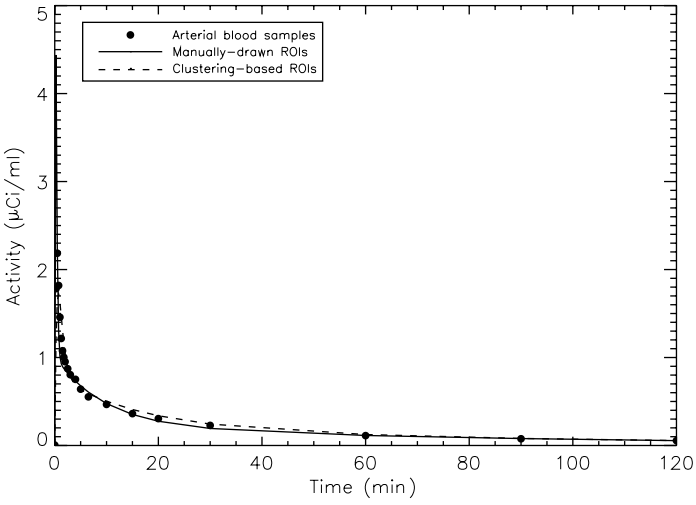

Figure 3.16: Plot of the measured arterial input function, the estimated input function from *manually drawn regions of interest (ROIs)*, and *clustering based ROIs*. The estimated input functions were obtained by simultaneously fitting with three ROIs of distinct kinetics.

Alternatively, clustering can be applied to extract input function directly on the dynamic PET/SPECT data if the vascular structures (e.g. left ventricle [81] and internal carotid artery [82]) are present in the field of view, providing that partial volume and spillover effects are appropriately corrected. Clustering has also been found useful in analyzing PET/SPECT neuroreceptor kinetics in conjunction with simplified techniques for quantification [106]. In particular, identification of regions that are devoid of specific binding is attractive because the kinetics of these regions can be treated as a noninvasive input function to the simplified approach for parametric imaging of binding potentials and relative delivery [107, 108].

3.8 Fast Generation of Parametric Images

Fast generation of parametric images is now possible with current high-speed computer workstations. However, overestimation of parameters and negative parameter estimates, which are not physiologically feasible, occur often when the data are too noisy. Reliable parametric imaging is therefore largely dependent

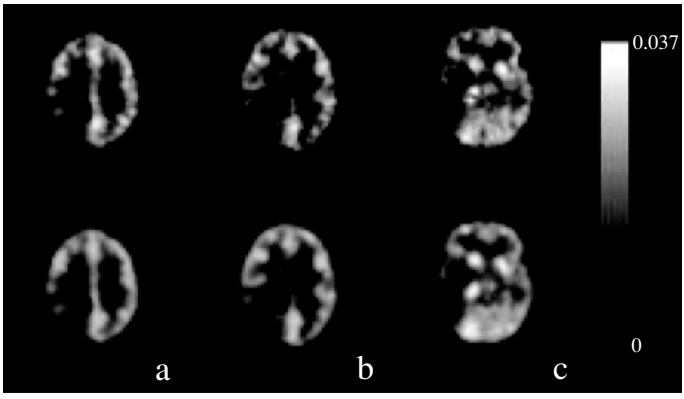

Figure 3.17: Parametric images on a pixel-by-pixel basis of K obtained from Patient 1: (a) slice 10; (b) slice 13; (c) slice 21. Top row shows the images obtained from the raw dynamic images and bottom row shows the images obtained from dynamic images after cluster analysis. The images have been smoothed slightly for better visualization.

on the noise levels inherent in the data which affect, in addition to meaningful parameter estimation, the time required to converge as well as the convergence. Clustering may be useful as a preprocessing step before fast generation of parametric images since only a few characteristic curves which have high statistical significance, need to be fitted as compared to conventional pixel-by-pixel parametric image generation where many thousands of very noisy tissue TACs must be analyzed. The computational advantage and time savings for generation of parametric images (fitting many thousands of kinetic curves versus several curves) are apparent.

Figure 3.17 shows the parametric images of physiological parameters, K, obtained from the neurologic study for Patient 1 in the three selected slices. The top and bottom rows of the images correspond to the results obtained from pixel-by-pixel fitting the TACs in the raw dynamic PET data and data after cluster analysis, respectively. The K images are relatively noisy when compared to the data after cluster analysis because of the high-noise levels of PET data which hampered reliable parametric image generation. However, the visual quality of the K images improves markedly with cluster analysis as a result of the increased SNR of the dynamic images. Low-pass filtering of the original parametric images may improve the SNR but clustering should produce better results because it takes

the tissue TACs with similar temporal characteristics for averaging. Meanwhile, low-pass filtering only makes use of the spatial (adjacent pixels) information for filtering and this will only further degrade the spatial resolution. The feasibility of using the kinetic curves extracted by cluster analysis for noninvasive quantification of physiological parameters and parametric imaging has been investigated and some preliminary data have been reported [109]. Some other recent studies can be found elsewhere [110–115].

3.9 Visualization of Disease Progress

In nuclear medicine, several kinds of organ function can be measured simultaneously with various radiopharmaceuticals under different conditions. This gives us useful information about the stage of disease progress if the relationship between various parameters such as metabolism, blood flow, and hemodynamics can be elucidated. Toyama *et al.* [116] investigated the use of agglomerative hierarchical and K-means clustering methods to study regional vasodilative and vascoconstrictive reactivity and oxygen metabolism before and after revascularization surgery in chronic occlusive cerebrovascular disease. By clustering a four-variable correlation map, whose pixel values on the X, Y, Z, and T axes represent, respectively, the resting cerebral blood flow, the hyperventilatory response, the acetazolamide response, and regional oxygen metabolic rate, anatomically and pathophysiologically different areas can be identified showing the involvement of certain areas with varying degrees of progression between pre- and postsurgical treatment, while functional changes in the revascularized region can be depicted on the clustered brain images. It appears that clustering technique maybe useful for multivariate staging of hemodynamic deficiency in obstructive cerebrovascular disease and it also be suitable for objective representation of multiple PET physiological parameters obtained from ^{15}O-labeled compound studies and also in brain activation studies.

3.10 Characterization of Tissue Kinetics

Kinetic modeling of radiotracer (or radiopharmaceutical) is the core of dynamic PET/SPECT imaging. The aim of modeling is to interpret kinetic data

quantitatively in terms of physiological and pharmacological parameters of a mathematical model, which describes the exchanges (e.g. delivery and uptake) of radiotracer by the tissue. Statistical inferences can then be made regarding the distribution and circulation of tracers within different tissues regions, which are quantitatively represented by the physiological/pharmacological parameters in the model. Successful statistical inference relies heavily on the appropriate use of analysis approaches and *a priori* knowledge of the underlying system as well as the validity of the assumptions being made. What if we know nothing about the underlying system, or little is known about the tracer characteristics and we are unsure if the assumptions (e.g. tissue homogeneity) being made are valid? The use of kinetic modeling could lead to incorrect inferences about the complex physiological or biochemical processes. In this case, data-driven approaches can provide important clues to what is going on inside the underlying system and how the radiotracer behaves inside the tissue, as they interrogate the measured data to characterize the complex processes, with minimal assumptions and independent of any kinetic model.

Evaluation of soft tissue sarcomas (STS) is a challenging clinical problem because these tumors are very heterogeneous, and the treatment of patients with STS is also very complicated. The most essential step in the diagnostic evaluation of STS is tumor biopsy. PET imaging has the ability to differentiate benign from malignant lesions. It can detect intralesional morphologic variation in soft tissue sarcomas, and it is of value in grading tumor, staging, restaging, and prognosis. Fluoromisonidazole (FMISO) has been shown to bind selectively to hypoxic cells *in vitro* and *in vivo* at radiobiologically significant oxygen levels. When labeled with the positron emitter fluorine-18 (^{18}F), its uptake in tissue can be localized and detected quantitatively with high precision by PET. [^{18}F]FMISO uptake has been investigated in various human malignancies [117]. PET imaging with [^{18}F]FMISO, FDG, and ^{15}O-water may provide valuable information complementary to tumor biopsy for better understanding of the biological behavior of STS. As cluster analysis does not rely on tracer assumptions nor kinetic model, it maybe useful in analyzing tissue TACs of STS obtained from [^{18}F]FMISO-PET and FDG-PET, and in looking for any correlations, for instance, tumor volume, hypoxic volume, and vascular endothelial growth factor expression, etc, between different datasets [118].

3.11 Partial Volume Correction in PET

Ideally, after corrections for the physical artifacts (e.g. scatter and attenuation) and calibration, the reconstructed PET image should represent highly accurate radiopharmaceutical distribution in absolute units of radioactivity concentration throughout the field of view of the PET scanner. However, this only holds for organs or structures with dimensions greater than twice the spatial resolution of the scanner, which is characterized by the full width at half maximum height of an image of a line source. When the object or structure being imaged is smaller than this, the apparent activity concentration is diluted. The degree of dilution in activity concentration varies with the size of the structure being imaged and the radioactivity concentration of the imaged structure comparing to its surrounding structures [10]. This phenomenon is known as *partial volume effect* (PVE), which is solely caused by the limited spatial resolution of the PET scanner.

A number of approaches have been proposed to correct or minimize the PVE, including resolution recovery before or during image reconstruction, and incorporation of side information provided from anatomical imaging modalities such as CT and MRI. One of the popular approaches that incorporates MRI segmentation for partial volume correction is the method proposed by Müller-Gärtner *et al.* [54] but the method is only applicable to brain imaging. PET images are first spatially co-registered with MR images obtained from the same subject. The MR images are then segmented into gray matter, white matter, and CSF regions, represented in three separate images. These images are then convolved spatially with a smoothing kernel which is derived from the point spread function of the PET scanner. The convolved white matter, image is then normalized to the counts in a white matter ROI drawn on the PET image so that spillover white matter activity into gray matter regions can be removed. Finally, the resultant image is divided by the smoothed gray matter image so that signals in small structures, which were smoothed severely, can be enhanced.

3.12 Attenuation Correction in PET

Accurate attenuation correction (AC) is essential to emission computed tomography such as PET and SPECT, for both quantitative and qualitative

interpretation of the results (Chapter 2 of *Handbook of Biomedical Image Analysis: Segmentation*, *Volume I*). In PET, for instance, AC factors are most often determined by calculating the pixel-wise ratio of a blank scan acquired before positioning the patient in the gantry of the scanner, and a transmission scan performed with the patient in the gantry. The major drawback of this approach is that statistical noise in the transmission data would propagate to the emission data [119, 120]. Depending on several factors such as body size and composition, transmission scans of 15–30 min are often performed to minimize the propagation of noise to the emission data through AC, at the price of reducing the patient throughput and increasing the errors due to patient motion, causing misalignment between transmission and emission data. Segmented AC methods, which employ image segmentation to partition the transmission images into regions of homogeneous density such as soft tissue, lung, and solid bones whose AC factors are known *a priori* and can be assigned, are particularly useful in cases where propagation of noise in transmission measurements during AC becomes a significant effect. A number of approaches based on the framework of pixel classification techniques and region-based segmentation approaches have been proposed and examined for segmented AC in PET. For example, Huang *et al.* [121] proposed a method where the operator manually defines the outlines of the body and the lung on the attenuation images. Known attenuation coefficients are then assigned to these regions and noiseless AC factors are then obtained by forward projecting the attenuation images with assigned attenuation coefficients. This approach has been further extended by a number of investigators by automating the determination of lung and the body regions using image segmentation techniques. For instance, Xu *et al.* used local thresholding approach to segment attenuation images into air, lung, and soft tissue [122]. Meikle *et al.* [123] used histogram fitting techniques to assign the attenuation values based on an assumed probability distribution for the lung and soft tissue components. Papenfuss *et al.* [124] used expectation-maximization clustering technique in conjunction with thresholding to produce fuzzy segmentation of attenuation images. Likewise, Bettinardi *et al.* [125] proposed an adaptive segmentation technique, also based on fuzzy clustering. This method can automatically determine the number of tissue classes in the attenuation images. The method can generally be applied to any region of the body. At least one of the aforementioned methods is currently in routine use by many PET centers worldwide.

3.13 Application to Analysis of fMRI Data

Functional MRI is a powerful modality for determining neural correlates of cognitive processes. It can be used to monitor changes of physiological parameters such as regional cerebral blood flow, regional cerebral blood volume, and blood oxygenation during cognitive tasks [126]. To extract functional information and detect activated regions using fMRI, the most widely adopted procedures are generally based on statistics theory and are paradigm dependent [65, 127, 128]. Cluster analysis has recently been applied to the fMRI discipline [129–131]. It is anticipated that cluster analysis will have great impact on analysis of fMRI signals for the detection of functional activation sites during cognitive tasks.

3.14 Discussion and Conclusions

In this chapter, a number of segmentation techniques used in, but not specific to functional imaging have been detailed. In particular, tissue segmentation and classification in functional imaging are of primary interests for dynamic imaging, for which cluster analysis is a valuable asset for data analysis because the identified characteristic curves are in the same space as the original data. This certainly has advantages over PCA in terms of interpretation of identified PCs, and over FADS where the factor components are rotated, leading to possibly nonunique factor explanation and interpretation. This chapter focuses on functional segmentation and a clustering technique is presented and discussed in great detail. The proposed technique is an attempt to overcome some of the limitations associated with commonly used manual ROI delineation, which is labor intensive and time-consuming. The clustering technique described is able to provide statistically meaningful clusters because the entire sequence of images are analyzed and different kinetic behaviors and the associated regions are extracted from the dataset, as long as there is a finite number of kinetics in the data. Once the segmentation process is completed, the extracted TACs, i.e. the cluster centroids, are then mapped back to the original data space for all voxels. Thus, an improved SNR can be achieved because each voxel in the mapped data space is represented by one of the cluster centroids each of which possesses a higher statistical significance than an individual TAC in the same spatial

location. Therefore, the extracted TACs obtained by cluster analysis should be more consistent and reproducible.

It is difficult to identify obvious cluster centroids in PET data because they are multidimensional and noisy. Therefore, initial centroids are needed for the proposed algorithm. The initial cluster centroids do not have to be accurate because they are used only as seeds to start the algorithm. However, if the starting centroids are far from the final cluster centroids, more iterations may be required. An incorrect initial selection may occur if a noisy outlier is chosen, resulting in a cluster with a single member. For this case, a lower bound on the final number of members in a cluster should be incorporated to prevent the cluster from being exhausted.

The optimum number of clusters for cluster analysis is usually not known *a priori*. The number of clusters, k, is also dependent on a number of factors mentioned previously. In addition, different choices for the values of k may result in different partitions of data. In this study, we limited the range for the values of k and applied the clustering algorithm to the simulated and real data. Nevertheless, it is reasonable to assume that the limited number of clusters used in this study is feasible, given that there is a finite number of kinetics present in the data. With the use of information-theoretic approaches to cluster validation, one can objectively determine the optimum number of clusters for the given dataset. However, caution should be taken when using the criteria as they are model dependent. The optimum number of clusters suggested by the criteria may not make sense if the specified probability distribution function for the observed data is not appropriate. There are a number of statistical criteria for the determination of the optimal number of clusters in addition to those used in this study and we are currently exploring various approaches, including the minimum description length to the cluster validation problem.

A limitation of the proposed algorithm is that it cannot differentiate anatomical structures having similar kinetics but are unconnected spatially. It is expected that future work will consider the addition of other information such as the geometry and the coordinates of the structure concerned. Another related issue is tissue heterogeneity [132] although this effect is usually ignored. In the work described, we did not attempt to solve this problem for cluster analysis. However, some heuristic interpretations could be made. In anatomy, most of the anatomical structures are discrete and well separated, they should easily be segmented by the proposed algorithm. Because of biological variations, a tissue

type may have inherent heterogeneity in it. A typical example is tumor, which are naturally heterogeneous. Activity concentration in a small tissue structure can be underestimated or overestimated due to partial volume effects, which cause the structure being imaged to mix with adjacent structures of possibly markedly different kinetics within the image volume, resulting in a mixed kinetics of the structures involved. As a finite number of clusters is assumed to be present in the raw PET data, the clustering algorithm will automatically look for the cluster centers that best represent the dataset without any *a priori* knowledge about the data and without violating the specified number of clusters. Therefore, certain regions which are indeterminate but their kinetics are similar, may be grouped together due to the constraint on the number of clusters, resulting in the formation of vague clusters. Further studies are required to investigate tissue heterogeneity in cluster analysis.

In earlier work, O'Sullivan [110] used cluster analysis as an intermediate step to extract "homogeneous" TACs from data containing a heterogeneous mix of kinetics resulting from spillover and partial volume effects for parametric mapping. This method is very similar to FADS but still there is a main difference between them: FADS extracts physiological factors (TACs) that could (theoretically) be found in the original data, whereas the set of "homogeneous" TACs does not necessarily correspond to the underlying physiology. In this current work, cluster analysis is used to extract kinetic data with different temporal characteristics as well as for parametric mapping. This is important for data analysis because data with different temporal behavior are better characterized by the extracted features seen in a spatial map. A spatial map is simpler to interpret when compared to the original multidimensional data. However, similar to O'Sullivan's approach [110], our method is data driven and is independent of the properties of tracer that may be required by other methods [111]. Thus, the clustering approach can be applied to a wide range of tracer studies.

It is anticipated that cluster analysis has a great deal of potential in PET data analysis for various neurodegenerative conditions (e.g. dementias) or diseases such as multiple system atrophy, Lewy Body disease, and Parkinsons disease where numerical values for glucose metabolism and the patterns of glucose hypometabolism may aid in the diagnosis and the assessment of disease progression. Localization of seizure foci in patients with refractory extratemporal epilepsy is also important as it is a difficult management problem for surgical epilepsy programs for this patient group. Functional segmentation may be

a useful tool in this regard. Investigation of the applicability of cluster analysis to whole-body PET for lesion localization and assessment of treatment response in a variety of oncological conditions will also be a fruitful research direction. Combining information provided by structural images for segmentation of functional image data will certainly become one of the key research areas in a new, hybrid PET/CT imaging technique, which will likely replace the existing PET alone facilities and will become a new standard of cancer imaging in the near future. With this hybrid imaging technique, precisely coregistered anatomical (CT) and functional (PET) images can be acquired in a single scanning session, and accurate localization of lesions could be achieved with the use of CT images as they provide very clear boundary delineation and anatomical information.

The above list of segmentation methods and applications are by no means complete. In fact, segmentation is one of the most difficult problems in medical image analysis but it is very useful in many medical imaging applications. Tremendous efforts have been devoted to cope with different segmentation problems. Continuing advances in exploitation and development of new conceptual approaches for effective extraction of all information and features contained in different types of multidimensional images are of increasingly importance in this regard. The following quote by the late philosopher, Sir Karl Popper, is worth noting when we think about new ideas and using analysis tools [133]:

> . . . at any moment we are prisoners caught in the framework of our theories; our expectations; our language. But . . . if we try, we can break out of our framework at any time. Admittedly, we shall find ourselves again in a framework, but it will be a better and roomier one; and we can at any moment break out of it again.

There is no magic method that suits all problems. One has to realize the strengths and limitations of the technique, and understand what kind of information the technique provides, and careful definition of the goals of segmentation is essential. It is also important to remember that new ideas and techniques may bring us something valuable that we are eager to see but something may be overlooked or missed out in the mean time, because we are bounded by the framework of the ideas or techniques, just like a prisoner, as Popper said. What we can only hope is that the new idea or the new technique, i.e. the prison, will be a better and roomier one where we can break out of it again at any time if there is a need!

Acknowledgment

This work was partially supported by the Hong Kong Polytechnic University under Grant G-YX13. Some of the results presented in this chapter were obtained in the period 1999 to 2002 during which the author was sustained financially by the National Health and Medical Research Council (NHMRC) of Australia.

Questions

1. *What is the major purpose of image segmentation? Why is it so significant in medical image analysis?*
2. *Identify the major classes of techniques for image segmentation.*
3. *List the advantages and disadvantage of using edge detection techniques for image segmentation.*
4. *What are the advantages and disadvantages of using manual region of interest (ROI) delineation with respect to using a template?*
5. *What are the disadvantages of using similarity mapping for segmentation?*
6. *What are the common and differences between principal component analysis (PCA) and factor analysis of dynamic structures (FADS)?*
7. *What are the major advantages of cluster analysis over other multivariate analysis approaches such as PCA and FA?*
8. *What are the major advantages of using clustering for characterizing tissue kinetics?*

Bibliography

[1] Rosenfeld, A. and Kak, A. C., Digital Image Processing, Academic Press, New York, 1982.

[2] Bajcsy, R. and Kovacic, S., Multiresolution elastic matching, Comp. Vision Graph. Image Proc., Vol. 46, pp. 1–21, 1989.

[3] Lim, K. O. and Pfefferbaum, A., Segmentation of MR brain images into cerebrospinal fluid spaces, white, and gray matter, J. Comput. Assist. Tomogr., Vol. 13, pp. 588–593, 1989.

[4] Brzakovic, D., Luo, X. M., and Brzakovic, P., An approach to automated detection of tumors in mammograms, IEEE Trans. Med. Imaging, Vol. 9, pp. 233–241, 1990.

[5] Liang, Z., MacFall, J. R., and Harrington, D. P., Parameter estimation and tissue segmentation from multispectral MR images, IEEE Trans. Med. Imaging, Vol. 13, pp. 441–449, 1994.

[6] Ardekani, B. A., Braun, M., Hutton, B. F., Kanno, I., and Iida, H., A fully automatic multimodality image registration algorithm, J. Comput. Assist. Tomogr., Vol. 19, pp. 615–623, 1995.

[7] Bankman, I. N., Nizialek, T., Simon, I., Gatewood, O. B., Weinberg, I. N., and Brody, W. R., Segmentation algorithms for detecting microcalcifications in mammograms, IEEE Trans. Inform. Technol. Biomed., Vol. 1, pp. 141–149, 1997.

[8] Small, G. W., Stern, C. E., Mandelkern, M. A., Fairbanks, L. A., Min, C. A., and Guze, B. H., Reliability of drawing regions of interest for positron emission tomographic data, Psych. Res., Vol. 45, pp. 177–185, 1992.

[9] White, D. R., Houston, A. S., Sampson, W. F., and Wilkins, G. P., Intra- and interoperator variations in region-of-interest drawing and their effect on the measurement of glomerular filtration rates, Clin. Nucl. Med., Vol. 24, pp. 177–181, 1999.

[10] Hoffman, E. J., Huang, S. C., and Phelps, M. E., Quantitation in positron emission computed tomography, 1: Effect of object size, J. Comput. Assist. Tomogr., Vol. 3, pp. 299–308, 1979.

[11] Mazziotta, J. C., Phelps, M. E., Plummer, D., and Kuhl, D. E., Quantitation in positron emission compted tomography, 5: Physical-anatomical effects, J. Cereb. Blood Flow Metab., Vol. 5, pp. 734–743, 1981.

[12] Hutchins, G. D., Caraher, J. M., and Raylman, R. R., A region of interest strategy for minimizing resolution distortions in quantitative myocardial PET studies, J. Nucl. Med., Vol. 33, pp. 1243–1250, 1992.

[13] Welch, A., Smith, A. M., and Gullberg, G. T., An investigation of the effect of finite system resolution and photon noise on the bias and precision of dynamic cardiac SPECT parameters, Med. Phys., Vol. 22, pp. 1829–1836, 1995.

[14] Bezdek, J., Hall, L., and Clarke, L., Review of MR image segmentation techniques using pattern recognition, Med. Phys., Vol. 20, pp. 1033–1048, 1993.

[15] Mazziotta, J. C. and Koslow, S. H., Assessment of goals and obstacles in data acquisition and analysis from emission tomography: Report of a series of international workshops, J. Cereb. Blood Flow Metab., Vol. 7(Suppl. 1), pp. S1–S31, 1987.

[16] Mazziotta, J. C., Pelizzari, C. A., Chen, G. T., Bookstein, F. L., and Valentino, D., Region of interest issues: The relationship between structure and function in the brain, J. Cereb. Blood Flow Metab., Vol. 11, pp. A51–A56, 1991.

[17] Fu, K. S. and Mui, J. K., A survey on image segmentation, Pattern Recogn., Vol. 13, pp. 3–16, 1981.

[18] Haralick, R. M. and Shapiro, L. G., Survey: Image segmentation techniques, Comput. Vision Graphics Image Proc., Vol. 29, pp. 100–132, 1985.

[19] Pal, N. R. and Pal, S. K., A review on image segmentation techniques, Pattern Recogn., Vol. 26, pp. 1227–1249, 1993.

[20] Gonzalez, R. C. and Woods, R. E., Digital Image Processing, Addison-Wesley, Reading, MA, 1993.

[21] Castleman, K. R., Digital Image Processing, Prentice Hall, Upper Saddle River, NJ, 1996.

[22] Kittler, J., Illingworth, J., and Foglein, J., Threshold based on a simple image statistics, Comp. Vision Graph. Image Proc., Vol. 30, pp. 125–147, 1985.

[23] Chow, C. K. and Kaneko, T., Automatic boundary detection of the left ventricle from cineangiograms, Comput. Biomed. Res., Vol. 5, pp. 388–410, 1972.

[24] Marr, D. and Hildreth, E., Theory of edge detection, Proc. Roy. Soc. London, Vol. 27, pp. 187–217, 1980.

[25] Sun, Y., Lucariello, R. J., and Chiaramida, S. A., Directional low-pass filtering for improved accuracy and reproducibility of stenosis quantification in coronary arteriograms, IEEE Trans. Med. Imaging, Vol. 14, pp. 242–248, 1995.

[26] Faber, T. L., Akers, M. S., Peshock, R. M., and Corbett, J. R., Three-dimensional motion and perfusion quantification in gated single-photon emission computed tomograms, J. Nucl. Med., Vol. 32, pp. 2311–2317, 1991.

[27] Hough, P. V. C., A method and means for recognizing complex patterns, US Patent 3069654, 1962.

[28] Deans, S. R., The Radon Transform and Some of Its Applications, Wiley, New York, 1983.

[29] Radon, J., Über die bestimmung von funktionen durchihre integralwärte längs gewisser männigfaltigkeiten, Bertichte Säechsiche Akad. Wissenschaften (Leipzig), Math. Phys. Klass, Vol. 69, pp. 262–277, 1917.

[30] Kalviainen, H., Hirvonen, P., Xu, L., and Oja, E., Probabilistic and non-probabilistic Hough transforms: Overview and comparisons, Image Vision Comput., Vol. 13, pp. 239–252, 1995.

[31] Kassim, A., Tan, T., and Tan, K., A comparative study of efficient generalized Hough transforms techniques, Image Vision Comput., Vol. 17, pp. 737–748, 1999.

[32] Martelli, A., Edge detection using heuristic search methods, Comp. Graph. Image Proc., Vol. 1, pp. 169–182, 1972.

[33] Nilsson, N. J., Principles of Artificial Intelligence, Springer-Verlag, Berlin, 1982.

[34] Geiger, D., Gupta, A., Costa, A., and Vlontzos, J., Dynamic programming for detecting, tracking, and matching deformable contours, IEEE Trans. Patt. Anal. Mach. Intell., Vol. 17, pp. 294–302, 1995.

[35] Barret, W. A. and Mortensen, E. N., Interactive live-wire boundary detection, Med. Image Analy., Vol. 1, pp. 331–341, 1996.

[36] Zucker, S., Region growing: Childhood and adolescence, Comp. Graph. Image Proc., Vol. 5, pp. 382–399, 1976.

[37] Hebert, T. J., Moore, W. H., Dhekne, R. D., and Ford, P. V., Design of an automated algorithm for labeling the cardiac blood pool in gated SPECT images of radiolabeled red blood cells, IEEE Trans. Nucl. Sci., Vol. 43, pp. 2299–2305, 1996.

[38] Kim, J., Feng, D. D., Cai, T. W., and Eberl, S., Automatic 3D temporal kinetics segmentation of dynamic emission tomography image using adaptive region growing cluster analysis, In: Proceedings of 2002 IEEE Medical Imaging Conference, Vol. 3, IEEE, Norfolk, VA, pp. 1580–1583, 2002.

[39] Hartigan, J. A., Clustering Algorithms, Wiley, New York, 1975.

[40] Cooper, L., M-dimensional location models: Application to cluster analysis, J. Reg. Sci., Vol. 13, pp. 41–54, 1973.

[41] Bezdek, J. C., Ehrlich, R., and Full, W., FCM: The fuzzy c-means clustering algorithm, Comp. Geosci., Vol. 10, pp. 191–203, 1984.

[42] Ball, G. H. and Hall, D. J., A clustering technique for summarizing multi-variate data, Behav. Sci., Vol. 12, pp. 153–155, 1967.

[43] Anderberg, M. R., Cluster Analysis for Applications, Academic Press, New York, 1973.

[44] McLachlan, G. J. and Krishnan, T., The EM Algorithm and Extensions, Wiley, New York, 1997.

[45] Kass, M., Witkin, A., and Terzopoulos, D., Snakes: Active contour models, Int. J. Comput. Vis., Vol. 1, pp. 321–331, 1987.

[46] Terzopoulos, D. and Fleischer, K., Deformable models, Visual Comput., Vol. 4, pp. 306–331, 1988.

[47] Fischler, M. A. and Elschlager, R. A., The representation and matching of pictorial structures, IEEE Trans. Comput., Vol. 22, pp. 67–92, 1973.

[48] Widrow, B., The "rubber-mask" technique, Pattern Recogn., Vol. 5, pp. 175–211, 1973.

[49] McInerney, T. and Terzopoulos, D., Deformable models in medical image analysis: A survey, Med. Image Analy., Vol. 1, pp. 91–108, 1996.

[50] Mykkänen, J. M., Tohka, J., and Ruotsalainen, U., Automated delineation of brain structures with snakes in PET, In: Physiological Imaging of the Brain with PET, Gjedde, A., Hansen, S. B., Knudsen, G., and Paulson, O. B., eds., Academic Press, San Diego, pp. 39–43, 2001.

[51] Chiao, P. C., Rogers, W. L., Fessler, J. A., Clinthorne, N. H., and Hero, A. O., Motion-based estimation with boundary side information or boundary regularization, IEEE Trans. Med. Imaging, Vol. 13, pp. 227–234, 1994.

[52] Chiao, P. C., Rogers, W. L., Clinthorne, N. H., Fessler, J. A., and Hero, A. O., Model-based estimation for dynamic cardiac studies using ECT, IEEE Trans. Med. Imaging, Vol. 13, pp. 217–226, 1994.

[53] Meltzer, C. C., Leal, J. P., Mayberg, H. S., Wagner, H. N., and Frost, J. J., Correction of PET data for partial volume effects in human cerebral cortex by MR imaging, J. Comput. Assist. Tomogr., Vol. 14, pp. 561–570, 1990.

[54] Müller-Gärtner, H. W., Links, J. M., Price, J. L., Bryan, R. N., McVeigh, E., Leal, J. P., Davatzikos, C., and Frost, J. J., Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects, J. Cereb. Blood Flow Metab., Vol. 12, pp. 571–583, 1992.

[55] Fox, P. T., Perlmutter, J. S., and Raichle, M. E., A stereotatic method of anatomical localization for positron emission tomography, J. Comput. Assist. Tomogr., Vol. 9, pp. 141–153, 1985.

[56] Talairach, J., Tournoux, P., and Rayport, M., Co-planar Stereotaxic Atlas of the Human Brain, Thieme, Inc., New York, 1988.

[57] Thompson, P. and Toga, A., A surface-based technique for warping three-dimensional images of the brain, IEEE Trans. Med. Imaging, Vol. 15, pp. 402–417, 1996.

[58] Bremner, J. D., Bronen, R. A., De Erasquin, G., Vermetten, E., Staib, L. H., Ng, C. K., Soufer, R., Charney, D. S., and Innis, R. B., Development and reliability of a method for using magnetic resonance imaging for the definition of regions of interest for positron emission tomography, Clin. Pos. Imag., Vol. 1, pp. 145–159, 1998.

[59] Maintz, J. B. A. and Viergever, M. A., A survey of medical image registration, Med. Imag. Analy., Vol. 2, pp. 1–37, 1998.

[60] Pelizzari, C. A., Chen, G. T. Y., Spelbring, D. R., Weichselbaum, R. R., and Chen, C. T., Accurate three-dimensional registration of CT, PET and/or MR images of the brain, J. Comput. Assist. Tomogr., Vol. 13, pp. 20–26, 1989.

[61] Woods, R. P., Mazziotta, J. C., and Cherry, S. R., MRI-PET registration with automated algorithm, J. Comput. Assisted Tomogr., Vol. 17, pp. 536–546, 1993.

[62] Rogowska, J., Similarity methods for dynamic image analysis, In: Proceedings of International AMSE Conference on Signals and Systems, Vol. 2, Warsaw, Poland, 15–17 July 1991, pp. 113–124.

[63] Barber, D. C., The use of principal components in the quantitative analysis of gamma camera dynamic studies, Phys. Med. Biol., Vol. 25, pp. 283–292, 1980.

[64] Rogowska, J. and Wolf, G. L., Temporal correlation images derived from sequential MR scans, J. Comput. Assist. Tomogr., Vol. 16, pp. 784–788, 1992.

[65] Bandettini, P. A., Jesmanowicz, A., Wong, E. C., and Hyed, J. S., Processing strategies for time-course datasets in functional MRI of the human brain, Magn. Res. Med., Vol. 30, pp. 161–173, 1993.

[66] Rogowska, J., Preston, K., Hunter, G. J., Hamberg, L. M., Kwong, K. K., Salonen, O., and Wolf, G. L., Applications of similarity mapping in dynamic MRI, IEEE Trans. Med. Imaging, Vol. 14, pp. 480–486, 1995.

[67] Jolliffe, I., Principal Component Analysis, Springer, New York, 1986.

[68] Pearson, K., On lines and planes of closest fit to systems of points in space, Phil. Mag., Vol. 6, pp. 559–572, 1901.

[69] Hotelling, H., Analysis of a complex of statistical variables into principal components, J. Edu. Psycho., Vol. 24, pp. 417–441, 1933.

[70] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., Numerical Recipes in C. The Art of Scientific Computing, Cambridge University Press, New York, 1992.

[71] Golub, G. H. and Van Loan, C. F., Matrix Computations, 3rd edn., John Hopkins University Press, Baltimore, 1996.

[72] Moeller, J. R. and Strother, S. C., A regional covariance approach to the analysis of functional patterns in positron emission tomographic data, J. Cereb. Blood Flow Metab., Vol. 11, pp. A121–A135, 1991.

[73] Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S., Functional connectivity: The principal component analysis of large (PET) data sets, J. Cereb. Blood Flow Metab., Vol. 13, pp. 5–14, 1993.

[74] Pedersen, F., Bergström, M., and Långström, B., Principal component analysis of dynamic positron emission tomography images, Eur. J. Nucl. Med., Vol. 21, pp. 1285–1292, 1994.

[75] Strother, S. C., Anderson, J. R., Schaper, K. A., Sidtis, J. S., and Rottenberg, D. A., Linear models of orthogonal subspaces and networks from functional activation PET studies of the human brain, In: Information Processing in Medical Imaging, Bizais, Y., Barillot, C., and Di Paola, R., eds., Kluwer, Dordrecht, The Netherlands, pp. 299–310, 1995.

[76] Ardekani, B. A., Strother, S. C., Anderson, J. R., Law, I., Paulson, O. B., Kanno, I., and Rottenberg, D. A., On the detection of activation patterns using principal components analysis, In: Quantitative Functional Brain Imaging with Positron Emission Tomography, Carson, R. E., Daube-Witherspoon, M. E., and Herscovitch, P., eds., Academic Press, San Diego, pp. 253–257, 1998.

[77] Anzai, Y., Minoshima, S., Wolf, G. T., and Wahl, R. L., Head and neck cancer: Detection of recurrence with three-dimensional principal components analysis at dynamic FDG PET, Radiology, Vol. 212, pp. 285–290, 1999.

[78] Andersen, A. H., Gash, D. M., and Avison, M. J., Principal component analysis of the dynamic response measured by fMRI: A generalized linear systems framework, Mag. Res. Imag., Vol. 17, pp. 795–815, 1999.

[79] Baumgartner, R., Ryner, L., Richter, W., Summers, R., Jarmasz, M., and Somorjai, R., Comparison of two exploratory data analysis methods for fMRI: Fuzzy clustering vs. principal component analysis, Mag. Res. Imag., Vol. 18, pp. 89–94, 2000.

[80] Correia, J., A bloody future for clinical PET? [editorial], J. Nucl. Med., Vol. 33, pp. 620–622, 1992.

[81] Iida, H., Rhodes, C. G., De Silva, R., Araujo, L. I., Bloomfield, P. M., Lammertsma, A. A., and Jones, T., Use of the left ventricular time-activity curve as a non-invasive input function in dynamic Oxygen-15-Water positron emission tomography, J. Nucl. Med., Vol. 33, pp. 1669–1677, 1992.

[82] Chen, K., Bandy, D., Reiman, E., Huang, S. C., Lawson, M., Feng, D., Yun, L. S., and Palant, A., Noninvasive quantification of the cerebral metabolic rate for glucose using positron emission tomography, ^{18}F-fluoro-2-deoxyglucose, the Patlak method, and an image-derived input function, J. Cereb. Blood Flow Metab., Vol. 18, pp. 716–723, 1998.

[83] Houston, A. S., The effect of apex-finding errors on factor images obtained from factor analysis and oblique transformation, Phys. Med. Biol., Vol. 29, pp. 1109–116, 1984.

[84] Nirjan, K. S. and Barber, D. C., Factor analysis of dynamic function studies using *a priori* physiological information, Phys. Med. Biol., Vol. 31, pp. 1107–1117, 1986.

[85] Šámal, M., Kárný, M., Sùrová, H., and Dienstbier, Z., Rotation to simple structure in factor analysis of dynamic radionuclide studies, Phys. Med. Biol., Vol. 32, pp. 371–382, 1987.

[86] Buvat, I., Benali, H., Frouin, F., Bazin, J. P., and Di Paola, R., Target apex-seeking in factor analysis on medical sequences, Phys. Med. Biol., Vol. 38, pp. 123–128, 1993.

[87] Sitek, A., Di Bella, E. V. R., and Gullberg, G. T., Factor analysis with *a priori* knowledge—Application in dynamic cardiac SPECT, Phys. Med. Biol., Vol. 45, pp. 2619–2638, 2000.

[88] Wu, H. M., Hoh, C. K., Buxton, D. B., Schelbert, H. R., Choi, Y., Hawkins, R. A., Phelps, M. E., and Huang, S. C., Factor analysis for extraction of blood time-activity curves in dynamic FDG-PET studies, J. Nucl. Med., Vol. 36, pp. 1714–1722, 1995.

[89] Wu, H. M., Huang, S. C., Allada, V., Wolfenden, P. J., Schelbert, H. R., Phelps, M. E., and Hoh, C. K., Derivation of input function from FDG-PET studies in small hearts, J. Nucl. Med., Vol. 37, pp. 1717–1722, 1996.

[90] Sitek, A., Di Bella, E. V. R., and Gullberg, G. T., Factor analysis of dynamic structures in dynamic SPECT imaging using maximum entropy, IEEE Trans. Nucl. Sci., Vol. 46, pp. 2227–2232, 1999.

[91] Sitek, A., Gullberg, G. T., and Huesman, R. H., Correction for ambiguous solutions in factor analysis using a penalized least squares objective, IEEE Trans. Med. Imaging, Vol. 21, pp. 2166–225, 2002.

[92] Ashburner, J., Haslam, J., Taylor, C., Cunningham, V. J., and Jones, T., A cluster analysis approach for the characterization of dynamic PET data, In: Quantification of Brain Function using PET, Myers, R., Cunningham, V., Bailey, D., and Jones, T., eds., Academic Press, San Diego, pp. 301–306, 1996.

[93] Acton, P. D., Pilowsky, L. S., Costa, D. C., and Ell, P. J., Multivariate cluster analysis of dynamic iodine-123 iodobenzamide SPET dopamine D_2 receptor images in schizophrenia, Eur. J. Nucl. Med., Vol. 24, pp. 111–118, 1997.

[94] Wong, K. P., Feng, D., Meikle, S. R., and Fulham, M. J., Segmentation of dynamic PET images using cluster analysis, IEEE Trans. Nucl. Sci., Vol. 49, pp. 200–207, 2002.

[95] Wong, K. P., Feng, D., Meikle, S. R., and Fulham, M. J., Simultaneous estimation of physiological parameters and the input function—*In vivo* PET data, IEEE Trans. Inform. Technol. Biomed., Vol. 5, pp. 67–76, 2001.

[96] Wong, K. P., Meikle, S. R., Feng, D., and Fulham, M. J., Estimation of input function and kinetic parameters using simulated annealing: Application in a flow model, IEEE Trans. Nucl. Sci., Vol. 49, pp. 707–713, 2002.

[97] Cunningham, V. J. and Jones, T., Spectral analysis of dynamic PET studies, J. Cereb. Blood Flow Metab., Vol. 13, pp. 15–23, 1993.

[98] Zubal, I. G., Harrell, C. R., Smith, E. O., Rattner, Z., Gindi, G., and Hoffer, P. B., Computerized three-dimensional segmented human anatomy, Med. Phys., Vol. 21, pp. 299–302, 1994.

[99] Hoffman, E. J., Cutler, P. D., Digby, W. M., and Mazziotta, J. C., 3-D phantom to simulate cerebral blood flow and metabolic images for PET, IEEE Trans. Nucl. Sci., Vol. 37, pp. 616–620, 1990.

[100] Hawkins, R. A., Phelps, M. E., and Huang, S. C., Effects of temporal sampling, glucose metabolic rates, and disruptions of the blood-brain barrier on the FDG model with and without a vascular compartment: Studies in human brain tumors with PET, J. Cereb. Blood Flow Metab., Vol. 6, pp. 170–183, 1986.

[101] Akaike, H., A new look at the statistical model identification, IEEE Trans. Automatic Control, Vol. AC-19, pp. 716–723, 1974.

[102] Schwarz, G., Estimating the dimension of a model, Ann. Stat., Vol. 6, pp. 461–464, 1978.

[103] Hooper, P. K., Meikle, S. R., Eberl, S., and Fulham, M. J., Validation of post injection transmission measurements for attenuation correction in neurologic FDG PET studies, J. Nucl. Med., Vol. 37, pp. 128–136, 1996.

[104] Huang, S. C., Phelps, M. E., Hoffman, E. J., Sideris, K., Selin, C., and Kuhl, D. E., Noninvasive determination of local cerebral metabolic rate of glucose in man, Am. J. Physiol., Vol. 238, pp. E69–E82, 1980.

[105] Patlak, C. S., Blasberg, R. G., and Fenstermacher, J., Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data, J. Cereb. Blood Flow Metab., Vol. 3, pp. 1–7, 1983.

[106] Gunn, R. N., Lammertsma, A. A., and Cunningham, V. J., Parametric imaging of ligand-receptor interactions using a reference tissue model and cluster analysis, In: Quantitative Functional Brain Imaging with Positron Emission Tomography, Carson, R. E., Daube-Witherspoon, M. E., and Herscovitch, P., eds., Academic Press, San Diego, pp. 401–406, 1998.

[107] Lammertsma, A. A. and Hume, S. P., Simplified reference tissue model for PET receptor studies, NeuroImage, Vol. 4, pp. 153–158, 1996.

[108] Gunn, R. N., Lammertsma, A. A., Hume, S. P., and Cunningham, V. J., Parametric imaging of ligand-receptor binding in PET using a simplified reference region model, NeuroImage, Vol. 6, pp. 279–287, 1997.

[109] Wong, K. P., Feng, D., Meikle, S. R., and Fulham, M. J., Non-invasive determination of the input function in PET by a Monte Carlo approach and cluster analysis, J. Nucl. Med., Vol. 42, No. 5(Suppl.), p. 183P, 2001.

[110] O'Sullivan, F., Imaging radiotracer model parameters in PET: A mixture analysis approach, IEEE Trans. Med. Imaging, Vol. 12, pp. 399–412, 1993.

[111] Kimura, Y., Hsu, H., Toyama, H., Senda, M., and Alpert, N. M., Improved signal-to-noise ratio in parametric images by cluster analysis, NeuroImage, Vol. 9, pp. 554–561, 1999.

[112] Bentourkia, M., A flexible image segmentation prior to parametric estimation, Comput. Med. Imaging Graphics, Vol. 25, pp. 501–506, 2001.

[113] Kimura, Y., Senda, M., and Alpert, N. M., Fast formation of statistically reliable FDG parametric images based on clustering and principal components, Phys. Med. Biol., Vol. 47, pp. 455–468, 2002.

[114] Zhou, Y., Huang, S. C., Bergsneider, M., and Wong, D. F., Improved parametric image generation using spatial-temporal analysis of dynamic PET studies, NeuroImage, Vol. 15, pp. 697–707, 2002.

[115] Bal, H., DiBella, E. V. R., and Gullberg, G. T., Parametric image formation using clustering for dynamic cardiac SPECT, IEEE Trans. Nucl. Sci., Vol. 50, pp. 1584–1589, 2003.

[116] Toyama, H., Takazawa, K., Nariai, T., Uemura, K., and Senda, M., Visualization of correlated hemodynamic and metabolic functions in cerebrovascular disease by a cluster analysis with PET study, In: Physiological Imaging of the Brain with PET, Gjedde, A., Hansen, S. B., Knudsen, G. M., and Paulson, O. B., eds., Academic Press, San Diego, pp. 301–304, 2001.

[117] Koh, W. J., Rasey, J. S., Evans, M. L., Grierson, J. R., Lewellen, T. K., Graham, M. M., Krohn, K. A., and Griffin, T. W., Imaging of hypoxia in human tumors with [F-18]fluoromisonidazole, Int. J. Radiat. Oncol. Biol. Phys., Vol. 22, pp. 199–212, 1992.

[118] Marsden, P. K., Personal communication, 2003.

[119] Huang, S. C., Hoffman, E. J., Phelps, M. E., and Kuhl, D. E., Quantitation in positron emission computed tomography, 2: Effects of inaccurate attenuation correction, J. Comput. Assist. Tomogr., Vol. 3, pp. 804–814, 1979.

[120] Dahlbom, M. and Hoffman, E. J., Problems in signal-to-noise ratio for attenuation correction in high-resolution PET, IEEE Trans. Nucl. Sci., Vol. 34, pp. 288–293, 1987.

[121] Huang, S. C., Carson, R. E., Phelps, M. E., Hoffman, E. J., Schelbert, H. R., and Kuhl, D. E., A boundary method for attenuation correction in positron computed tomography, J. Nucl. Med., Vol. 22, pp. 627–637, 1981.

[122] Xu, M., Luk, W. K., Cutler, P. D., and Digby, W. M., Local threshold for segmented attenuation correction of PET imaging of the thorax, IEEE Trans. Nucl. Sci., Vol. 41, pp. 1532–1537, 1994.

[123] Meikle, S. R., Dahlbom, M., and Cherry, S. R., Attenuation correction using count-limited transmission data in positron emission tomography, J. Nucl. Med., Vol. 34, pp. 143–144, 1993.

[124] Papenfuss, A. T., O'Keefe, G. J., and Scott, A. M., Segmented attenuation correction in whole body PET using neighbourhood EM clustering, In: 2000 IEEE Medical Imaging conference, IEEE Publication, Lyon, France, 2000.

[125] Bettinardi, V., Pagani, E., Gilardi, M. C., Landoni, C., Riddell, C., Rizzo, G., Castiglioni, I., Belluzzo, D., Lucignani, G., Schubert, S., and Fiazio, F., An automatic classification technique for attenuation correction in positron emission tomography, Eur. J. Nucl. Med., Vol. 26, pp. 447–458, 1999.

[126] Ogawa, S., Lee, T. M., Kay, A. R., and Tank, D. W., Brain magnetic resonance imaging with contrast dependent on blood oxygenation, Proc. Natl. Acad. Sci. USA, Vol. 87, pp. 9868–9872, 1990.

[127] Bullmore, E. and Brammer, B., Statistical methods of estimation and inference for functional MR image analysis, Magn. Reson. Med., Vol. 35, pp. 261–277, 1996.

[128] Lange, N., Statistical approaches to human brain mapping by functional magnetic resonance imaging, Stat. Med., Vol. 15, pp. 389–428, 1996.

[129] Moser, E., Diemling, M., and Baumgartner, R., Fuzzy clustering of gradient-echo functional MRI in the human visual cortex. Part II: Quantification, J. Magn. Reson. Imaging, Vol. 7, pp. 1102–1108, 1997.

[130] Goutte, C., Toft, P., Rostrup, E., Nielsen, F. Å., and Hansen, L. K., On clustering fMRI time series, NeuroImage, Vol. 9, pp. 298–310, 1999.

[131] Fadili, M. J., Ruan, S., Bloyet, D., and Mazoyer, B., A multistep unsupervised fuzzy clustering analysis of fMRI time series, Hum. Brain Mapping, Vol. 10, pp. 160–178, 2000.

[132] Schmidt, K., Lucignani, G., Moresco, R. M., Rizzo, G., Gilardi, M. C., Messa, C., Colombo, F., Fazio, F., and Sokoloff, L., Errors introduced by tissue heterogeneity in estimation of local cerebral glucose utilization with current kinetic models of the [^{18}F]fluorodeoxyglucose method, J. Cereb. Blood Flow Metab., Vol. 12, pp. 823–834, 1992.

[133] Popper, K. R., Normal science and its dangers, In: Criticism and the Growth of Knowledge, Lakatos, I. and Musgrave, A., eds., Cambridge University Press, Cambridge, pp. 51–58, 1970.