Segmentation of VOI From Multidimensional Dynamic PET Images by Integrating Spatial and Temporal Features

Jinman Kim, *Member, IEEE*, Weidong Cai, *Member, IEEE*, Dagan Feng, *Fellow, IEEE*, and Stefan Eberl, *Member, IEEE*

***Abstract*—Segmentation of multidimensional dynamic positron emission tomography (PET) images into volumes of interest (VOIs) exhibiting similar temporal behavior and spatial features is a challenging task due to inherently poor signal-to-noise ratio and spatial resolution. In this study, we propose VOI segmentation of dynamic PET images by utilizing both the three-dimensional (3-D) spatial and temporal domain information in a hybrid technique that integrates two independent segmentation techniques of cluster analysis and region growing. The proposed technique starts with a cluster analysis that partitions the image based on temporal similarities. The resulting temporal partitions, together with the 3-D spatial information are utilized in the region growing segmentation. The technique was evaluated with dynamic 2-[^{18}F] fluoro-2-deoxy-D-glucose PET simulations and clinical studies of the human brain and compared with the k-means and fuzzy c-means cluster analysis segmentation methods. The quantitative evaluation with simulated images demonstrated that the proposed technique can segment the dynamic PET images into VOIs of different kinetic structures and outperforms the cluster analysis approaches with notable improvements in the smoothness of the segmented VOIs with fewer disconnected or spurious segmentation clusters. In clinical studies, the hybrid technique was only superior to the other techniques in segmenting the white matter. In the gray matter segmentation, the other technique tended to perform slightly better than the hybrid technique, but the differences did not reach significance. The hybrid technique generally formed smoother VOIs with better separation of the background. Overall, the proposed technique demonstrated potential usefulness in the diagnosis and evaluation of dynamic PET neurological imaging studies.**

***Index Terms*—Image segmentation, multivolume rendering, positron emission tomography (PET), quantitative evaluation.**

I. Introduction

THE functional imaging modality of positron emission tomography (PET) allows *in vivo* studies of physiological processes [1]. To estimate the physiological function, dynamic PET images are acquired in four dimensions (three spatial dimensions and one temporal dimension), by quantifying the radiotracer distribution at predetermined sampling times. From these data, a physiological tissue time–activity curve (TTAC) can be extracted for each voxel and the physiological parameter value for that voxel can then be calculated by the application of a tracer kinetic model to the TTAC. Although physiological parameter estimation can be performed from the kinetic modeling of individual voxels, estimation based on volume of interest (VOI) segmentation comprising functionally and structurally related voxels can be of considerable benefit in a variety of neuroimaging applications [2]–[4]. For instance, the image classification based on the characteristics of the TTAC in dynamic PET images is often facilitated by using regional parameters from the segmentation definition [3]. In another study, the automatic segmentation of the blood vessel time–activity curves can eliminate the need for arterial blood sampling [4]. Furthermore, VOI segmentation enables three-dimensional (3-D) visualization of the multidimensional PET data, and therefore, provides the ability to navigate the VOI in three spatial directions [5].

Segmentation of VOI from dynamic PET images can be performed manually by a physician. Such manual delineation, however, is subjective, may not be reproducible, and is time consuming to perform. An alternate approach to manual delineation is the registration of the PET data with an anatomical imaging modality such as magnetic resonance imaging (MRI) [6], [7], allowing the MRI segmentation results to be used as a reference. Accurate and correct VOI information may be obtained from the MRI segmentation reference. However, this approach limits the segmentation to those PET images, which have corresponding MRI images. The semiautomatic or automatic segmentation of dynamic PET images into VOIs without the use of other imaging modalities is a challenging problem, primarily due to the inherently poor spatial resolution and signal-to-noise ratio (SNR) in the PET images. Conventional segmentation techniques, which rely only on spatial domain data and local statistical properties, may yield inaccurate segmentation, as they may have difficulties in accurately differentiating between adjacent tissue structures with different functional characteristics. It is likely that the segmentation can be improved by utilizing the additional physiological kinetic features from the temporal domain.

Dynamic PET images have been segmented using cluster analysis techniques for various applications, including radiotracer modeling [8], segmentation of regions of interest (ROIs)

Manuscript received January 13, 2005; revised April 19, 2005, October 4, 2005, and January 4, 2006. This work was supported in part by ARC and RGC grants.

J. Kim and D. Feng are with the Biomedical and Multimedia Information Technology Group, School of Information Technologies, University of Sydney, Sydney, NSW 2006, Australia, and also with the Center for Multimedia Signal Processing, Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Kowloon, Hong Kong (e-mail: jinman@it.usyd.edu.au; feng@it.usyd.edu.au).

W. Cai is with the Biomedical and Multimedia Information Technology Group, School of Information Technologies, University of Sydney, Sydney, NSW 2006, Australia (e-mail: tomc@it.usyd.edu.au).

S. Eberl is with the Biomedical and Multimedia Information Technology Group, School of Information Technologies, University of Sydney, Sydney, NSW 2006, Australia, and also with the Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, Sydney, NSW 2050, Australia (e-mail: Stefan@it.usyd.edu.au).

Digital Object Identifier 10.1109/TITB.2006.874192

[9], extraction of blood time–activity curves [4], and the generation of parametric images from huge data sets [10]. In cluster analysis of the dynamic PET images, the voxel-based TTACs are classified into a predefined number of clusters where the similarity of a given TTAC to the cluster mean TTAC is measured by the Euclidean distance between its data points and that of the cluster mean TTAC. In [11], an alternate similarity measure was presented based on the shape of the TTAC rather than the amplitude of the data points.

These techniques only utilize the temporal information of individual voxel-based TTACs, disregarding the information available in the spatial connectivity of voxels from a particular structure. It is a reasonable assumption that the voxels in the same structure are contiguous and neighboring voxels are more likely to belong to a given structure than distant voxels [12], [13]. Recently, a technique to combine the spatial correlation between voxels with the TTAC kinetic features was described using the Markov random field (MRF) model in expectation maximization (EM) iterations to segment dynamic PET images [12]. The study showed that the use of neighboring connectivity could improve the segmentation of tissue structures. Although such a technique could be applicable to using 3-D spatial correlation, this technique, as with others, was inherently implemented for a single slice of dynamic PET images, with two-dimensional (2-D) spatial correlation, thus, disregarding the 3-D spatial connectivity information available in the "slice" dimension. The study in [13] demonstrated the use of 3-D spatial connectivity, however, only for static PET images without utilizing the temporal information.

When utilizing the 3-D spatial connectivity for image segmentation, a region-growing (RG) approach, which postulates that the 3-D neighboring voxels are likely to belong to the same structure, has been shown to be an effective segmentation technique [14], [15]. RG starts from a set of "seed" voxels and a region is grown around each seed by appending the seed's immediate neighbors, which satisfy the similarity tolerance criterion. The RG technique is, however, largely dependent on the selections of seeds and tolerance criterion for successful segmentation. On the other hand, high-level knowledge of the image features can be exploited to select appropriate parameters [16]–[18]. Methods applied to the selection of parameters for RG techniques include intensity thresholding [15] and histogram analysis [16]. In [17], the seeds for the seeded RG segmentation were initially selected from the results of color-edge extraction segmentation. Such hybrid segmentation has been shown to provide more accurate segmentation results [17], [18]. The selection of RG parameters from the specific image information can facilitate segmentation; however, these parameter selection techniques are image dependent, and only applicable to the spatial domain. In dynamic PET images, both the spatial and temporal domain information can be exploited for the selection of RG parameters.

In this study, we describe and validate a VOI segmentation of dynamic PET images by utilizing both the 3-D spatial and temporal domain information, in a four-dimensional (4-D) hybrid technique that integrates the two independent segmentation techniques of cluster analysis and RG. The technique starts with a functional cluster analysis, which partitions the image based on the temporal similarities. The resulting temporal partitions, together with the 3-D spatial information are then utilized in the RG segmentation. Cluster analysis and RG complement each other by utilizing both the temporal kinetic features and the 3-D spatial connectivity. We demonstrate our hybrid segmentation technique by applying it to the 4-D dynamic 2[^{18}F] fluoro-2-deoxy-D-glucose (FDG) PET simulations and clinical studies of the human brain and comparing it with the conventional segmentation techniques of k-means cluster analysis (KCA) and fuzzy c-means (FCM) cluster analysis.

II. Method

Our proposed 4-D dynamic PET image segmentation technique consists of four major stages as shown in the flowchart in Fig. 1: 1) background removal and feature extraction; 2) seed selection from cluster analysis; 3) RG using the KCA seeds; and 4) VOI refinements.

A. Background Removal and Feature Extraction

Prior to segmentation, low-background counts and the streaks from the reconstruction of PET images outside the brain are removed (set to zero) by thresholding. The threshold value is set at 15% of the maximum counts in each slice of the summed temporal frames. Artifacts from the thresholding, such as isolated voxels and gaps are then removed and/or filled by $3 \times 3 \times 3$ morphological opening and subsequent closing filters [19]. The mask generated from this background removal process is then applied to the individual temporal frames.

The TTACs are extracted for each of the N voxels, which were not removed by the background masking to form kinetic feature vectors $\mathbf{f}(t)$ with time points $t(t = 1, 2, \ldots, T)$, where T is the total number of time points. An optimal image sampling schedule (OISS) technique [20], [21] is applied to the extracted feature vectors. In the design of the OISS, an objective function based on the Fisher information matrix [21] is used to discriminate between different experimental protocols and sampling schedules. Only five temporal frames are needed in the proposed segmentation of dynamic PET images based on the OISS with a five-parameter FDG model. The application of OISS considerably reduces the dimension of the feature vectors while increasing the SNR of the individual frames, thus, potentially improving the measure of similarity among the feature vectors.

B. Potential Seed and Tolerance Criterion Parameter Selection Based on KCA

The KCA segmentation based on [22] is initially applied to assign each of the N feature vectors into one of a set number of distinct cluster groups. Voxels are assigned to the cluster group with the smallest Euclidean distance D between the cluster centroid $\bar{\mathbf{f}}_{c_j}(t)$ of jth cluster group c_j and the feature vector $\mathbf{f}_i(t)$ defined by

$$D\left(\mathbf{f}_i(t), \bar{\mathbf{f}}_{c_j}(t)\right) = \left(\sum_{t=1}^{T} s(t)\left(\mathbf{f}_i(t) - \bar{\mathbf{f}}_{c_j}(t)\right)^2\right)^{1/2} \quad (1)$$

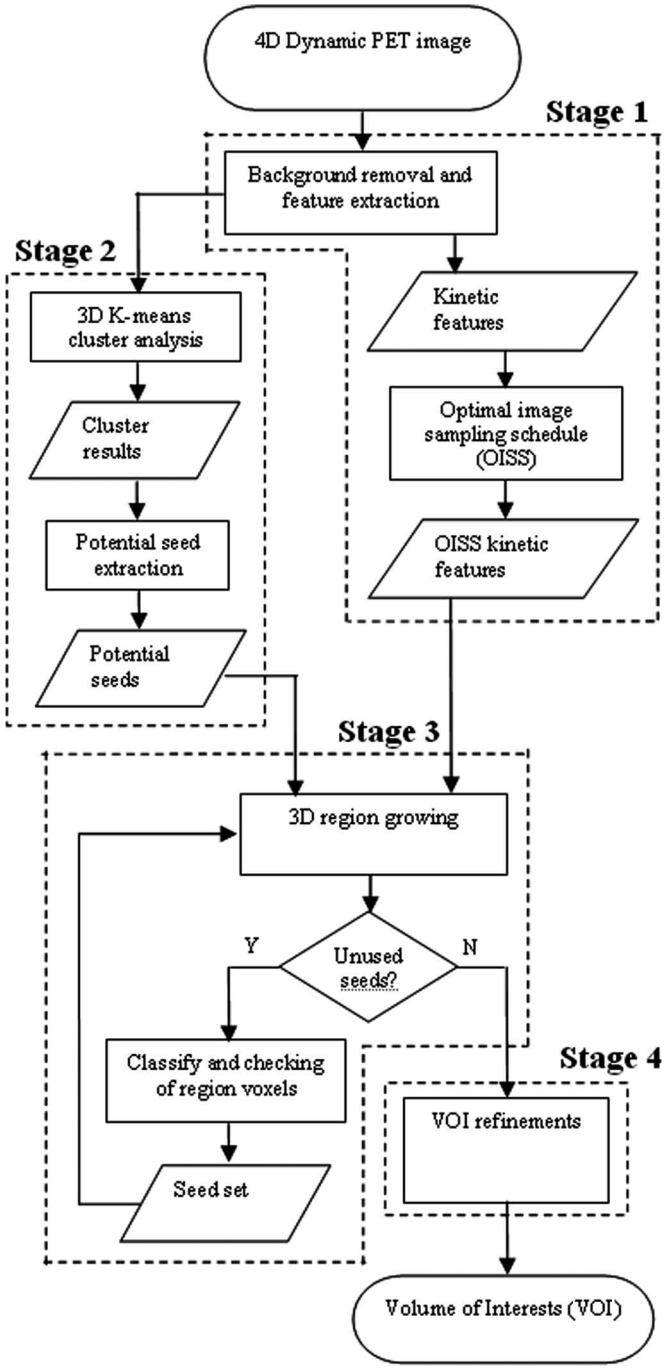


Fig. 1. Flowchart of the proposed hybrid technique that consists of four major stages. After background removal and feature extraction (stage 1), the KCA (stage 2) is applied to automatically extract the potential seeds and the tolerance parameters based on the statistics of the segmentation result for use in the RG technique (stage 3). Finally, the segmented regions are refined (stage 4).

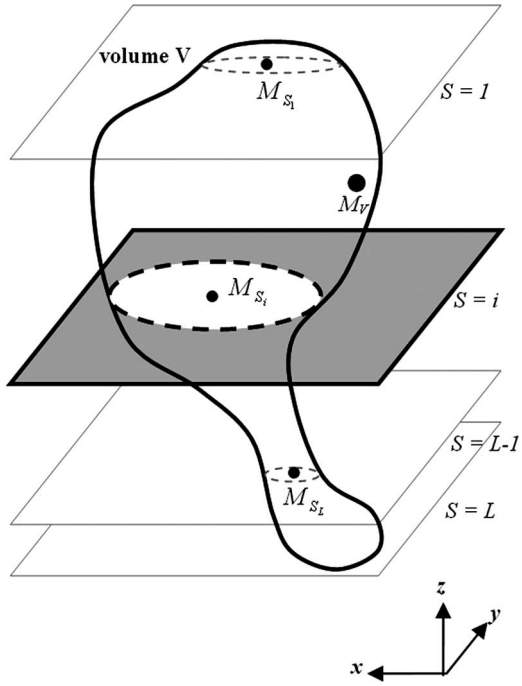


Fig. 2. For each segmented volume V spanning multiple slices ($S = 1, 2, \ldots L$), potential seeds are extracted from each slice using the 2-D centroid moment of the region $M_{S_1}, \ldots, M_{S_L}$. The volume centroid M_v may not correctly represent the entire volume for use in RG segmentation.

where $s(t)$ is the scale factor of time point $t(t = 1, 2, \ldots, T)$ equal to the duration of the tth frame divided by the total dynamic acquisition time. The scale factor gives more weight to the later, longer frames containing more reliable data. Allocation of the N voxels to the appropriate clusters is iterated until the number of voxels changing from one cluster to another does not vary. All the centroids of the clusters for the KCA segmentation are randomly initialized to be exclusive of each other and to include only nonbackground voxels. To avoid the false selection of a noisy outlier as a centroid, which may result in a cluster with a single member, the average of the neighboring voxels to the random seed is used as the centroid.

From the KCA segmentation results, for each 3-D cluster, closed groups of mutually exclusive voxels are extracted as volumes, and the potential seeds are extracted from these volumes for use in the RG stage. Isolated voxels that fail to form a group are discarded. Due to the volume being irregularly shaped, the 3-D centroid *moment* of the volume M_v may fall outside the volume from which it was extracted and is, thus, not suitable as a seed. In addition, a centroid, which resides within its volume, may be located at, or near the boundary of the volume and, thus, does not represent the entire volume correctly when applied to RG. Therefore, a 2-D centroid moment is extracted from every 2-D region in the axial planes that the volume occupies as illustrated in Fig. 2. For a centroid moment that also falls outside a 2-D region, the nearest voxel to the 2-D centroid moment that resides within the 2-D region is selected as the centroid. The centroid of the 2-D region, referred to as M_{S_i} of *slice* $S_i = S_1, S_2, \ldots, S_L$, is calculated as the spatial location of the potential seed. The average of feature vectors in the region $\bar{\mathbf{f}}_{S_i}(t)$ is then calculated and set as the potential seed's feature vector.

For each region where the potential seed is extracted, standard deviation (SD) of the Euclidean distance measures given in (1) between all the feature vectors residing in the region and their region centroid is computed. The $\pm 2 \times \mathrm{SD}$ is then assigned to the potential seed to approximate the tolerance parameter τ,

which corresponds to 95.4% of the feature vectors residing in the region and is, thus, likely to exclude only noisy and most dissimilar feature vectors. The τ is used in the RG stage to determine whether a feature vector is similar enough to the seed feature vector to be potentially allocated to the RG around the seed.

The extracted potential seeds from all the volumes are then ordered to minimize the inherent seed dependence limitation of the RG technique [23], according to two equally weighted criteria: 1) the number of feature vectors belonging to the seed's region and 2) the spatial distance between the centroid of the 2-D region M_{S_i} and the centroid of the volume M_v from where the 2-D region was extracted, as shown in Fig. 2. The seed extracted from the region with the largest number of feature vectors and the lowest distance to the volume centroid is given the highest order.

C. 4-D RG Based on Kinetic Features

Using the potential seeds, a region grows into VOIs from each seed by appending at most 26 immediate neighboring feature vectors that satisfy two conditions. First, Euclidean distance D given by (1) between a neighboring feature vector $\mathbf{f}_n(t)$ and the seed's feature vector $\bar{\mathbf{f}}_S(t)$ must be within the tolerance criterion parameter τ of the seed, i.e., $D(\mathbf{f}_n(t), \bar{\mathbf{f}}_s(t)) \leq \tau$. Second, α immediate neighboring feature vectors of $\mathbf{f}_n(t)$ must also be within the tolerance criterion τ of the same seed. The α value acts as a connectivity controller where the smaller the value, the more feature vectors satisfy the condition, however, the weaker the connectivity. To tolerate atypical voxels due to the noise in the dynamic PET images, α is empirically set to 22, which approximates 85% of the total neighbors. For the first and last image slices (axial slices), which do not have two adjacent slices, α is set to 14.

With every addition of a feature vector, $\bar{\mathbf{f}}_S(t)$ is adapted progressively by averaging the feature vectors. With the addition of feature vectors, the spatial location of the seed is not affected. The appended feature vector is then ranked against all of the already appended feature vectors to the seed according to $D(\mathbf{f}_n(t), \bar{\mathbf{f}}_s(t)) \times w$, where D is the Euclidean distance in (1) and w is a weight factor calculated as the inverse of the spatial distance between the centers of $\mathbf{f}_n(t)$ and $\bar{\mathbf{f}}_S(t)$. The weight factor is able to take into account the non-isotropic voxel dimensions frequently encountered in PET data. In the continuation of the RG, the immediate neighbors to the appended feature vectors are considered for the RG to the seed. The appended feature vector, which has the highest ranking, is used first; thus, the continuation of the RG utilizes the appended feature vector, which is most similar to the seed. The RG is continued until no new feature vectors are appended to the seed region, and this procedure is repeated using the next seed if it has not already been classified to a region until all the seeds have been depleted.

Although the potential seeds are ordered, the seed ordering may still influence the RG result. In the proposed RG technique, the RG for each seed proceeds without restrictions on the feature vectors, which have already been classified into another region (previous seed regions). Hence, feature vectors may be classified into more than one region from multiple seeds, causing an overlap of regions. The feature vectors from the overlapped regions are reclassified, so that such a feature vector is assigned to a region that has the highest neighboring connectivity. The neighboring connectivity is the count of the immediate neighbors that an overlapped region has with the feature vector. If the neighboring connectivity between the overlapped regions is equal, the overlapped feature vector is assigned to the region that has the highest kinetic similarity based on (1) between the feature vector and the centroids of the overlapped regions. Such an approach relies less on the importance of the seed ordering and the threshold selection, which may potentially improve the RG segmentation.

D. Physiological VOI Refinements

After segmenting, the dynamic PET images into VOIs, noisy outliers, or weakly connected feature vectors are classified to a VOI to which the majority of its 26 neighbors belong.

III. Simulation Evaluations

A. 4-D Dynamic Brain Phantom Simulation Study

Computer simulations were performed to evaluate the performance and reliability of the proposed method. The simulations were based on the anatomical Zubal brain phantom [24], which was reduced to the two tissue types of white matter (WM) and gray matter (GM). The slices were scaled to a $128 \times 128 \times 100$ volume with voxel dimension of $2.2 \times 2.2 \times 1.4$ mm^3. A five-parameter FDG model, which takes the effects of cerebral blood volume (CBV) into account, was used to define the physiological TTAC values to construct the dynamic temporal frame sequences representing the WM and GM structures. The model parameters were derived from the human brain studies in [25], and an input plasma tissue activity was obtained from our clinical studies. Tissue sample values were generated over a set of intervals using the conventional sampling schedule (CSS) sequences consisting of 10×0.2, 2×0.5, 2×1, 1×1.5, 1×3.5, 2×5, 1×10, and 3×30 min scans, based on the schedule of the clinical studies. To simulate realistic PET measurements, Gaussian blur was added with a full-width half-maximum (FWHM) value of 8 mm and dynamic sequences of sinogram projections were obtained by forward projection. The composite dynamic projection data were scaled by the acquisition time of the frames and then scaled to appropriate voxel counts to allow realistic noise to be added. Based on voxel counts observed in clinical studies, the data were scaled such that the maximum voxel count in the 30-min frames at peak uptake is 500 counts. After scaling, the low counts in the projections of early frames with counts less than 1 but greater than 0.2 were rounded up to a count of 1. Poisson noise was then added to the scaled projection data, which were then converted into counts per minute. Finally, the projection data were reconstructed using filtered backprojection with a Hann filter and a cutoff frequency of 0.75 of the Nyquist frequency. To investigate the effect of different noise levels, data sets with eight different levels of noise were constructed by scaling the data to maximum voxel counts at peak uptake ranging from 100 to 800 counts, in steps of 100 counts, respectively, for noise levels 1–8. OISS was applied directly to the projection data from the above CSS frames. The counts from the projection were

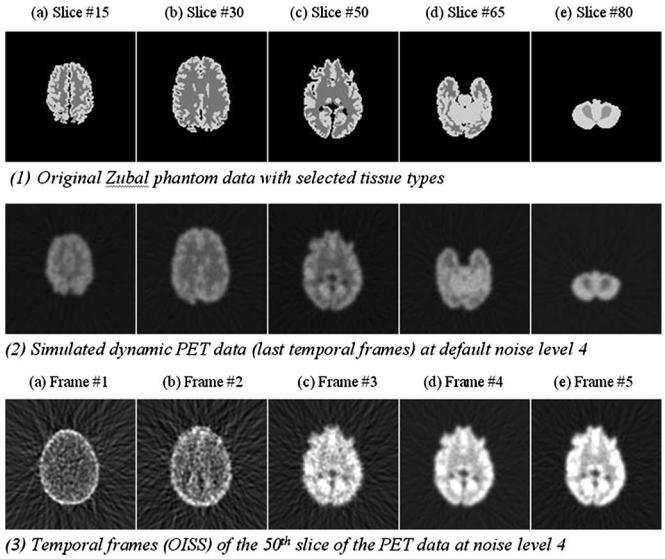


Fig. 3. Selected slices out of 100 slices from the default noise level 4 simulated images. Row (1) is original Zubal phantom with the tissue types assigned as WM (dark shade) and GM (light shade), (2) is the last temporal frame from the dynamic PET-simulated image, and (3) is the five temporal frame sequence of the 50th slice using the OISS. All images are adjusted to their local maximum intensity for display.

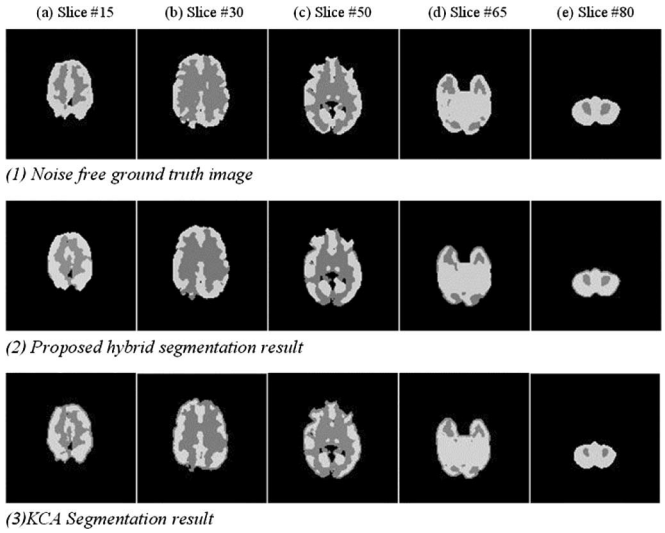


Fig. 4. Selected slices from the segmentation of the dynamic simulated PET images. Light shade of gray represents the GM and the dark shade represents the WM. Row (1) is the ground truth, (2) is the proposed hybrid segmentation result, and (3) is the KCA segmentation result.

resampled according to the scheduling time derived from [20] consisting of 1×0.731, 1×2.483, 1×11.250, 1×60.933, and 1×44.617 min scans prior to reconstruction.

Noise-free "ground truth" images were also constructed by adding partial volume effects using the same Gaussian blur function as in the noisy simulations. The voxels mixed at the boundary in the smoothed images were reclassified to the tissue matter with a contribution greater than 50%. For boundaries separating tissue from zero count regions, such as ventricles, the boundary was defined at the 40% level.

B. Segmentation Evaluation

The segmentation results were evaluated with overlapping volume measures of Dice similarity coefficient (DSC) [26] applied directly to the 3-D segmentation results in relation to the ground truth. The DSC measures the spatial overlap between the estimated A_{Est} and true A_{True} segmented volumes given by

$$\mathrm{DSC} = 2\frac{|A_{\mathrm{Est}} \cap A_{\mathrm{True}}|}{|A_{\mathrm{Est}}| + |A_{\mathrm{True}}|} \tag{2}$$

where $|A_{\mathrm{Est}}|$ is the cardinality of the volume A_{Est}. The DSC is in the range of [0, 1], where 1 represents two overlapping areas of identical size and location.

C. Results

Fig. 3 shows the examples of the simulated images. The first row illustrates the selected slices of the original noise-free image with tissue structures classified into WM and GM, the second row displays the simulated PET images at noise level 4 (last temporal frame), and the third row demonstrates the temporal frame sequence of a single slice (50th) from simulated PET images at noise level 4, which corresponds to the noise in typical clinical studies.

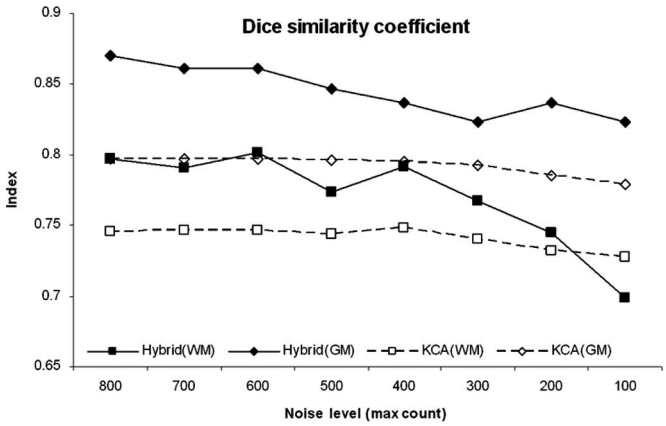


Fig. 5. Plots' DSC evaluation measures for the hybrid (Hybrid) and k-means (KCA) segmentation techniques for the simulations at the eight noise levels for the WM and GM.

In the segmentation of the simulated dynamic PET images, the background in the images was removed by using the head mask derived from the ground truth, which contains the boundaries of the tissue structures. Two clusters were applied, corresponding to WM and GM in the simulated images. No other segmentation parameters were adjusted. In Fig. 4, the segmentation results of the proposed hybrid technique (second row) and the KCA technique (third row) are presented. The KCA segmentation results used to compare with the hybrid segmentation are from the initial segmentation stage of the hybrid segmentation technique. The visual differences between the hybrid and KCA segmentation results illustrated in the second and third rows, respectively, in comparison to the ground truth (first row), are minor, with both techniques being able to separate the WM and GM correctly. The segmentation improvements of the proposed hybrid technique are more evident in our quantitative results using DSC index plotted in Fig. 5. In total, 80 slices as a volume

were applied in the quantitative evaluation. At low noise levels, the proposed hybrid technique performs better than KCA with indices up to 9% higher. The performance of the hybrid technique degrades slightly with increasing noise, while the KCA results are similar for all the noise levels, leading to a less pronounced difference in the performance between the two methods at high noise levels. These results suggest that the use of a 3-D neighboring connectivity in the spatial domain, in addition to the temporal features, does improve on the segmentation of WM and GM, particularly at moderate to low noise levels.

IV. Clinical Evaluation

We evaluated our proposed segmentation technique with clinical dynamic brain PET images from twelve patient studies and the results are compared with the KCA results from the initial stage of the proposed segmentation technique. In addition, the segmentation of dynamic PET images based on FCM cluster analysis [27] is also compared. In the FCM cluster analysis approach to the dynamic PET segmentation, a fuzzy logic algorithm assigns a probabilistic membership function (weighting) to every voxel representing the likelihood that the voxel is a member of a particular cluster. Upon convergence of fuzzy cluster analysis, voxels are assigned to the cluster for which it has the highest fuzzy membership.

Due to the lack of ground truth available for clinical studies, dynamic PET segmentation results are typically compared to the manual delineation of VOI structures by physicians [9], [10] or evaluated with a registered anatomical imaging modality. In this study, we use an objective quantitative evaluation method for the clinical PET images without the need for manual delineation of VOI structures or registration with an anatomical modality. In our evaluation, the segmentation results are transformed into a standard brain atlas. The transformed results are then compared to the corresponding segmentation of the MRI reference image. Although the proposed ground truth does not entirely reflect the segmentation results to be expected from PET images, the quantitative evaluation measures are able to provide comparative analysis of the segmentation technique qualities.

A. Clinical 4-D Dynamic FDG PET Study

The proposed segmentation technique is evaluated using clinical FDG brain PET images (SIEMENS ECAT 951R). The number of image slices is 31, with each slice acquired in 22 temporal frames using the CSS scanning schedule of 6×10 s frames, 4×0.5 min frames, 1×2.0 min frames, and 11×5.0 min frames. The dynamic images are decay corrected to the time of the injection, attenuation corrected, and then reconstructed using filtered backprojection with a Shepp and Logan filter cutoff at 0.5 of the Nyquist frequency. The reconstructed images are 128×128 with voxel dimensions of $1.841 \times 1.841 \times 3.375$ mm. For OISS with the five-parameter FDG model, the images are reframed into 1×41.0, 1×136.0, 1×567.0, 1×1711.0, and 1×1145.0 s scans.

B. Segmentation Evaluation

1) Transformation of the Segmented PET Images Into a Standard Brain Atlas: The anatomical standardization procedure of the 3-D stereotactic surface projection (3-D SSP) transformation from the NEUROSTAT package [28]–[30] is applied to warp the static functional brain PET images into a standard brain atlas. The algorithm deforms the image into a standard stereotactic atlas by linear scaling of the image to correct individual brain sizes and nonlinear warping to minimize regional anatomical variations. Only the last temporal frame from the dynamic PET images is used in the anatomical standardization, creating library files of the standardization procedure transformation parameters. These library files are then applied to the corresponding segmentation results to warp the segmented images into the same standardized image frame of reference. Voxels in the warped segmentation images having contributions from more than one tissue structures are assigned to the most related structure. In addition, slices with incomplete information are removed, since the brain atlas covers a region larger than our clinical studies acquired in a single-bed position. For evaluation purposes, the anatomically warped segmentation results consisting of regions from eight different clusters are manually classified into only two structures of WM and GM, using the region-wide (or cluster-wide) average TTAC values and the knowledge of the human brain anatomy and the physiological functions of the brain tissue structures. Based on this classification, the cerebrospinal fluid (CSF) structures are combined with the WM.

2) Construction of the Ground Truth: The corresponding MRI reference image from the NEUROSTAT library is in the same anatomical frame of reference as the warped segmentation result images. The reference image has a high spatial resolution, removed background, and excellent discrimination of tissue structures. This MRI reference image has been segmented into WM and GM using the FCM segmentation technique, to correspond to the dynamic PET image segmentation. The FCM technique was selected to provide an alternative segmentation technique in the construction of the ground truth. The segmented MRI reference image is then used as the ground truth for comparison to the warped segmentation results from the proposed hybrid and KCA techniques. The FCM segmentation of the clinical images were also comparatively evaluated. The MRI was segmented in 3-D and the result is stripped of sharp edges, and voxels, which are mixed at the boundary, are assigned as in the simulation construction (Section III-A). Finally, the images are filtered for isolated or small regions that are either merged into another region or removed.

3) Evaluation Measures: The DSC measure given in (3) is applied to the clinical studies using the MRI segmentation as the ground truth.

C. Results

The number of clusters used with the proposed hybrid segmentation technique was eight. Fig. 6 shows selected slices (out of 31) from the clinical brain dynamic PET images of patient study #s01 with a cerebral tumor and the results obtained by

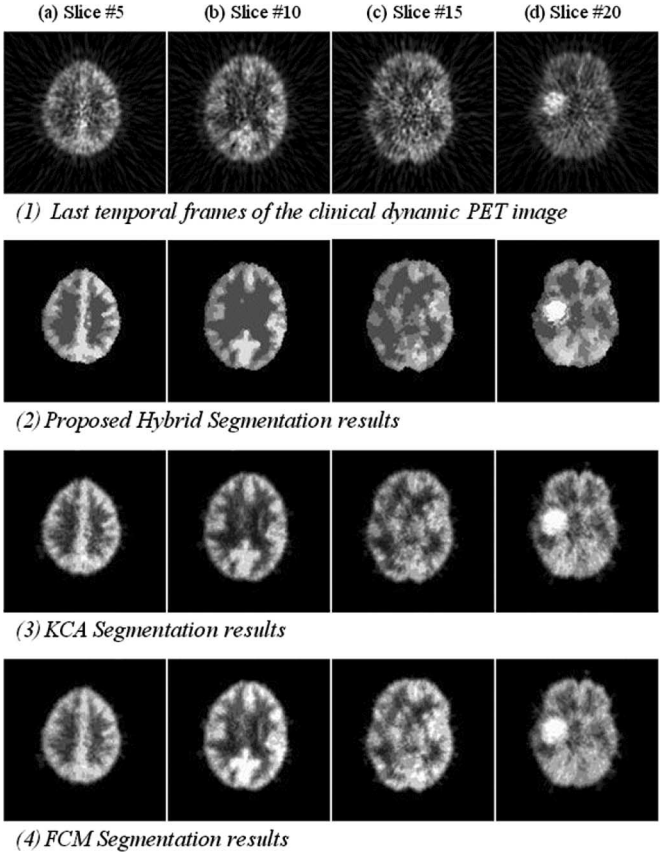


Fig. 6. Row (1) shows selected slices out of 31 slices from the original clinical dynamic PET image of patient study #s01, (2) is the corresponding segmentation result from the proposed hybrid technique, and (3) is the KCA segmentation results, followed by the FCM results in (4). Different shades of gray are applied to individual cluster groups and to the VOIs.

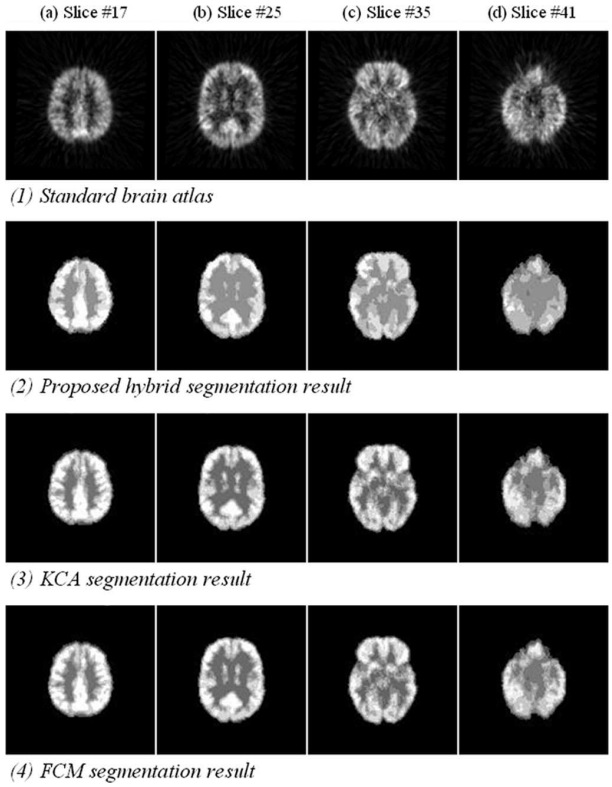


Fig. 7. Row (1) shows selected slices out of 60 slices from the warped patient study #s02 to the NEUROSTAT brain atlas. The anatomically warped segmentation results by the proposed hybrid, KCA, and FCM are presented in rows (2)–(4), respectively.

the three different segmentation techniques. Individual cluster groups are assigned a different shade of gray in ascending order based on the average of the feature vectors in the cluster group. The same shade of gray is assigned to all the VOIs in the hybrid segmentation results having its seeds extracted from the same cluster group. The KCA results presented are from the initial stage of the proposed hybrid technique, and for the FCM cluster analysis, the same background removal, similarity measure of weighted feature vectors, and eight clusters were applied to be consistent with the other segmentation results. From visual inspection, all three segmentation techniques were able to separate the different tissue structures, with clear identification of the tumor and other tissues with high glucose uptake. Different structures such as the cortex are also clearly segmented. However, the proposed hybrid technique appears to have better separation of the tissue structures from the background and stronger connectivity of voxels resulting in cleaner and smoother VOIs. It is interesting to note that in cluster analyses, not all clusters form into a group of functionally related voxels representing a tissue structure. Instead, there are clusters consisting only of boundary voxels between tissue structures scattered among the entire segmentation result. Such boundary voxels are not evident in the proposed segmentation technique, and therefore, more voxels are classified into related VOIs.

The quantitative evaluation of the clinical PET image segmentation was applied to 11 patient studies (#s02 to #s12) considered to have normal brain anatomy. Anatomically warped clinical images of patient study #s02 are presented in the first row of Fig. 7, followed by the hybrid, KCA, and FCM segmentation results (ordered gray levels for cluster groups and VOIs). The warped results demonstrate that the segmented structures are maintained and are correctly transformed for all segmentation results. Similar to the findings in Fig. 6, the VOIs of the proposed hybrid segmentation results are more strongly connected, avoiding loosely connected and scattered voxels evident with the cluster analysis techniques.

The visual findings are confirmed by the DSC evaluation measures of the proposed hybrid technique (Hybrid) against the conventional KCA and FCM for 11 clinical patient studies presented in Table I. The FCM segmentation of the NEUROSTAT MRI brain reference image provided the ground truth. Based on these results, the average indices of the proposed hybrid technique are consistently much higher in the WM; however, for the GM, the indices are slightly lower than those obtained with the two cluster analysis techniques. The higher indices in the WM are attributed to the better separation of the background, and the segmentation of WM structures, which appeared to resemble the ground truth more closely. However, improved WM and decreased GM suggests that the hybrid segmentation has over-segmented the WM and under-segmented the GM

TABLE 1
DICE SIMILARITY COEFFICIENT (1.0 IS PERFECT) EVALUATION MEASURES FOR SEGMENTATION PERFORMANCE

Patient study #	White matter			Grey matter		
	Hybrid	KCA	FCM	Hybrid	KCA	FCM
s02	0.550	0.390	0.407	0.688	0.698	0.698
s03	0.441	0.392	0.398	0.721	0.691	0.682
s04	0.557	0.449	0.446	0.620	0.653	0.649
s05	0.499	0.453	0.453	0.668	0.690	0.689
s06	0.514	0.398	0.400	0.590	0.659	0.658
s07	0.525	0.448	0.455	0.665	0.623	0.605
s08	0.378	0.346	0.351	0.635	0.648	0.646
s09	0.382	0.408	0.400	0.649	0.640	0.646
s10	0.477	0.379	0.394	0.668	0.678	0.669
s11	0.413	0.294	0.315	0.646	0.664	0.662
s12	0.363	0.425	0.378	0.674	0.679	0.684
Average	0.464	0.398	0.400	0.657	0.666	0.662
Standard deviation	0.072	0.048	0.042	0.035	0.024	0.026
Standard error	0.022	0.014	0.013	0.011	0.007	0.008

when compared to the cluster analysis methods. The SD and the standard error results suggest that all three techniques are consistent among the 11 clinical studies with the hybrid having slightly larger variations. The two-tailed paired Student's t-test was used to determine the statistically significant differences of the DSC indices between the hybrid, the KCA, and FCM techniques. The two-tailed p-value of the WM DSC indices was 0.009 for KCA/hybrid and 0.002 for FCM/hybrid pairs. For the GM, the values were 0.335 for KCA/hybrid and 0.593 for FCM/hybrid pairs. A p-value below 0.05 was considered as statistically significant. These results show that the hybrid technique is statistically reliable and significantly better in segmenting the WM; however, for GM, the difference was found not to be significant.

The DSC result was also compared to the KCA segmentation of the MRI reference as an additional ground truth, which further demonstrates the comparative performance of the hybrid technique to the cluster analysis techniques. The findings were similar to the results obtained from the FCM ground truth with all the indices being slightly higher (maximum 3% difference) in the DSC while the SD and standard error were about the same. Although the KCA ground truth resulted in higher indices, the comparative differences in the evaluation indices between the hybrid, KCA, and FCM techniques were not significant when compared to the results obtained using the FCM ground truth. The KCA ground truth evaluation further supports the results obtained using the FCM ground truth. Overall, our results demonstrate that the proposed hybrid technique provides improved segmentation of the WM, but there is no significant difference in the segmentation quality of the GM compared to the other techniques.

V. DISCUSSION

The proposed hybrid method showed a clear advantage over the k-means method in the simulation studies, particularly, at low to moderate noise levels typically encountered in clinical studies. At high noise levels, the advantage of the hybrid method was less pronounced, but nonetheless evident. The simulation studies provide the advantage that the ground truth is known, and thus, the performance of the segmentation methods can be objectively measured. However, simulations may not necessarily reflect all physiological and imaging variability and effects, such as differences in tracer uptake kinetics between patients, patient motion during the study, etc. In addition, the simulations were limited to simulating and segmenting only GM and WM areas in the brain. Thus, the techniques were applied to clinical studies to confirm that the findings of the simulation studies are also applicable to clinical studies. Visually, all three techniques were successful in segmenting the clinical studies, being able to differentiate between WM, GM, tumor, and other structures. However, the hybrid technique provided better connectivity within a given structure and resulted in fewer disconnected or spurious segmentation clusters compared with the other techniques.

Objective evaluation of segmentation results in clinical studies is challenging, since the ground truth is not known. Manual delineation of structures is subjective and may bias the results in favor of a particular segmentation technique. As in these clinical studies, coregistered anatomical imaging, such as MRI, is not necessarily available. Instead, we took the novel approach of warping the imaging to a standard atlas and then using the segmentation MRI data available as part of the standard atlas as ground truth. This approach has a number of limitations. It relies on the accurate transformation of the images into the standard atlas coordinate system. While the accuracy of the transformation has not been assessed in this study, the technique has been extensively validated [30], [31]. It is obviously not the true ground truth and this is reflected by the worse quantitative segmentation evaluation measures compared to the simulations. In its current form, it is applicable only to normal brain studies and, thus, is not applicable to patients with tumors. Nevertheless, given the agreement between these objective measures with the visual evaluation of the segmentation results, this approach proved useful in objectively comparing the segmentation performance of the proposed hybrid technique with the techniques with only cluster analysis.

The centroids of the clusters for the initial KCA segmentation were randomly initialized. This could potentially lead to different segmentation results if the segmentation algorithm is applied multiple times to the same data since the initial, randomly selected seeds will be different for each application. To test for this effect and to evaluate the reproducibility of the hybrid technique, the hybrid technique was applied 10 times to the same simulation data, but each time with a different set of randomly selected initial cluster centroids in the KCA. These segmentations, when compared to each other and their ground truth using the DSC measure (Section III-B), gave identical results. Thus, the final segmentation results were not affected by the initial random seeds. However, the number of iterations required for the KCA to terminate varied for these segmentations, with random cluster centroids that were further from the final cluster centroids requiring more iterations.

The segmentation parameters used in the hybrid segmentation were empirically derived to produce optimal segmentation results. The background-removal threshold was estimated based on the counts in the dynamic PET images. The threshold value is not critical and moderate changes in the threshold value had little effect on the results. The optimal number of clusters was selected based on the composition of the tissue structures and the experimental results. Due to the limited variation in patient studies (brain studies only), the same number of clusters, values of the tolerance criterion and connectivity controller were used for all the studies, derived empirically by varying these parameters and evaluating the segmentation results. As these parameters were derived to be optimal for our clinical study sets, the performance of the proposed hybrid technique applied to other dynamic PET images with different imaging properties may not be optimal without parameter tuning. Future studies will investigate the effect of these parameters, and the application of the hybrid technique to other dynamic PET images using different tracers and acquisition protocols.

VI. Conclusion

We have developed a VOI segmentation technique for the 4-D dynamic PET images of the human brain. The proposed hybrid technique uses both the spatial and temporal domain information available in the dynamic PET images by integrating cluster analysis and RG techniques. Overall, the proposed segmentation technique showed improvements in segmenting the VOI structures in the simulation studies and in the segmentation of the WM of the clinical studies, but did not show significant differences in the segmentation of the GM, compared to conventional cluster analysis approaches.

References

[1] D. Feng, D. Ho, H. Lida, and K. Chen, "Techniques for functional imaging," in invited chapter, in *Medical Imaging Systems Techniques and Applications: General Anatomy*, C. T. Leondes, Ed. Amsterdam, The Netherlands: Gordan and Breach, 1997, pp. 85–145.

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

[3] J. T. Yap, V. J. Cunningham, T. Jones, M. Cooper, C.-T. Chen, and P. Price, "Classification of dynamic PET images using a priori kinetic factors," in *Quantitative Functional Brain Imaging With PET*. New York: Academic, 1998, pp. 107–115.

[4] M. Liptrot, K. H. Adams, L. Martiny, L. H. Pinborg, M. N. Lonsdale, N. V. Olsen, S. Holm, C. Svarer, and G. M. Knudsen, "Cluster analysis in kinetic modeling of the brain: A noninvasive alternative to arterial sampling," *Neuroimage*, vol. 21, pp. 483–493, 2004.

[5] I. Bankman, *Handbook of Medical Imaging: Processing and Analysis Management*. New York: Academic, 2000.

[6] K.-P. Lin, S.-C. Huang, L. R. Baxter, and M. E. Phelps, "A general technique for interstudy registration of multifunction and multimodality images," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 6, pt. 1, pp. 2850–2855, Dec. 1994.

[7] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Imag. Anal.*, vol. 2, pp. 1–36, 1998.

[8] F. O'Sullivan, "Imaging radiotracer model parameters in PET: A mixture analysis approach," *IEEE Trans. Med. Imag.*, vol. 12, no. 3, pp. 399–412, Sep. 1993.

[9] K. P. Wong, D. Feng, S. R. Meikle, and M. J. Fulham, "Segmentation of dynamic PET images using cluster analysis," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 1, pt. 1, pp. 200–207, Feb. 2002.

[10] H. Guo, R. Eanaut, K. Chen, and E. Rieman, "Clustering huge data sets for parametric PET imaging," *Biosystems*, vol. 71, pp. 81–92, 2003.

[11] J. G. Brankov, N. P. Galastsanos, Y. Yongyi, and M. N. Wernick, "Segmentation of dynamic PET or fMRI images based on a similarity metric," *IEEE Trans. Nucl. Sci.*, vol. 50, no. 5, pt. 2, pp. 1410–1414, Oct. 2003.

[12] J. L. Chen, S. R. Gunn, M. S. Nixon, and R. N. Gunn, "Markov random field models for segmentation of PET images," in *Proc. IEEE Int. Conf. Inf. Process. Med. Imag.*, 2001, pp. 468–474.

[13] J. Mykkänen, J. Tohka, and U. Ruotsalainen, "Delineation of brain structures from positron emission tomography images with deformable models," in *Proc. Med. Informatics*, 2003, pp. 33–38.

[14] Y. Lu, T. Jiang, and Y. Zang, "Region growing method for the analysis of functional MRI data," *Neuroimage*, vol. 20, pp. 455–465, 2003.

[15] R. K. Justice and E. M. Stokely, "3-D segmentation of MR brain images using seeded region growing," in *Proc. IEEE Int. Conf. Med. Biol. Soc.*, Oct. 31–Nov. 3, 1996, vol. 3, pp. 1083–1084.

[16] Y. Tuduki, K. Murase, M. Izumida, H. Miki, K. Kikuchi, K. Murakami, and J. Ikezoe, "Automated seeded region growing algorithm for extraction of cerebral blood vessels from magnetic resonance angiographic data," in *Proc. IEEE Int. Conf. Electr. Biomed. Eng.,* Jul. 23–28, 2000, vol. 3, pp. 1756–1759.

[17] J. Fan, D. Y. Yau, A. K. Elmagarmid, and W. G. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1454–1466, Oct. 2001.

[18] T. Pavlidis and Y.-T. Liow, "Integrating region growing and edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 3, pp. 225–233, Mar. 1990.

[19] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[20] X. Li, D. Feng, and K. Chen, "Optimal image sampling schedule: A new effective way to reduce dynamic image storage and functional image processing time," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 710–719, Oct. 1996.

[21] C. Cobelli, A. Ruggeri, J. J. DiStefano III, and E. M. Landaw, "Optimal design of multi output sampling schedule—software and applications to endocrine—metabolic and pharmacokinetic models," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 4, pp. 249–256, Apr. 1985.

[22] B. Everitt, *Cluster Analysis*. New York: Halsted, 1974.

[23] A. Mehnert and P. Jackway, "An improved seeded region growing algorithm," *Pattern Recognit. Lett.*, vol. 18, pp. 1065–1071, 1997.

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

[25] R. A. Hawkins, M. E. Phelps, and S. C. Haung, "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.

[26] A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, and A. C. Palmer, "Morphometric analysis of white matter lesions in MR images: Methods and validation," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 716–724, Dec. 1994.

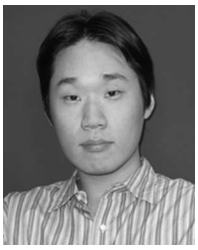
[27] J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithm*. Norwell, MA: Kluwer, 1981.

[28] S. Minoshima, K. L. Berger, K. S. Lee, and M. A. Mintun, "An automated method for rotational correction and centering of three-dimensional functional brain images," *J. Nucl. Med.*, vol. 33, pp. 1579–1585, 1992.

[29] S. Minoshima, R. A. Koeppe, M. A. Mintun, K. L. Berger, S. F. Taylor, K. A. Frey, and D. E. Kuhl, "Automated detection of the intercommissural line for stereotactic localization of functional brain images," *J. Nucl. Med.*, vol. 34, pp. 322–329, 1993.

[30] S. Minoshima, R. A. Koeppe, K. A. Frey, and D. E. Kuhl, "Anatomic standardization: Linear scaling and nonlinear warping of functional brain images," *J. Nucl. Med.*, vol. 35, pp. 1528–1537, 1994.

[31] K. Ishii, F. Willoch, S. Minoshima, A. Drzezga, E. P. Ficaro, D. J. Cross, D. E. Kuhl, and M. Schwaiger, "Statistical brain mapping of 18F-FDG PET in Alzheimer's disease: Validation of anatomic standardization for atrophied brains," *J. Nucl. Med.*, vol. 42, pp. 548–557, 2001.



Jinman Kim (M'06) received the B.S. degree (honors) in computer science and technology in 2001 from the University of Sydney, Sydney, Australia, where he is currently working toward the Ph.D. degree in information technologies.

He is also with the Center for Multimedia Signal Processing, Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Kowloon, Hong Kong. His research interests include the development of multidimensional image segmentation, image enhancement, information visualization, content-based image retrieval, and computer-aided diagnosis.

Weidong Cai (S'99–M'01) received the B.S. degree in computer science from HuaQiao University, Quanzhou, China, in 1989, and the Ph.D. degree in computer science from the University of Sydney, Sydney, Australia, in 2001.

Prior to his doctoral study, he worked in industry for five years. After graduation, he was a Postdoctoral Research Associate at the Centre for Multimedia Signal Processing (CMSP), Hong Kong Polytechnic University. He was then a Lecturer in 2001 and is currently a Senior Lecturer in the School of Information Technologies, University of Sydney. His research interests include computer graphics, image processing and analysis, data compression and retrieval, and multimedia database and computer modeling with biomedical applications.

Dagan Feng (S'88–M'88–SM'94–F'03) received the M.E. degree in electrical engineering and computing science from Shanghai JiaoTong University, Shanghai, China, in 1982, and the M.Sc. degree in biocybernetics and the Ph.D. degree in computer science from the University of California, Los Angeles, in 1985 and 1988, respectively.

After briefly working as an Assistant Professor at the University of California, Riverside, he joined the University of Sydney, Sydney, Australia, at the end of 1988, where he was a Lecturer, Senior Lecturer, Reader, Professor, Head of Department of Computer Science, Head of School of Information Technologies, and is currently the Associate Dean of the Faculty of Science. He is also the Honorary Research Consultant, Royal Prince Alfred Hospital, Sydney, the largest hospital in Australia; the Chair-Professor of Information Technology, Hong Kong Polytechnic University, Hong Kong; the Advisory Professor, Shanghai JiaoTong University; and a Guest Professor with Northwestern Polytechnic University, Xian, China, with Northeastern University, Shenyang, China, and with Tsinghua University, Beijing, China. He is the Founder and Director of the Biomedical and Multimedia Information Technology Research Group. He has published over 400 scholarly research papers, pioneered several new research directions, and made a number of landmark contributions in his field with significant scientific impact and social benefit. His research area is biomedical and multimedia information technology.

Dr. Feng is a Fellow of the Australia Computer Society, the Australian Academy of Technological Sciences and Engineering, the Hong Kong Institution of Engineers, and the Institution of Electrical Engineers, U.K. He is also the Special Area Editor of the IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE and is the current Chairman of IFAC-TC-BIOMED. He is the recipient of the Crump Prize for Excellence in Medical Engineering (USA).

Stefan Eberl (M'97) received the B.E. degree (honors) in electrical engineering from the New South Wales Institute of Technology, Sydney, Australia, in 1982, and the M.Sc. degree in physics and the Ph.D. degree in biomedical engineering from the University of New South Wales, Sydney, in 1997 and 2001, respectively.

He is currently a Principal Hospital Scientist in the Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, Sydney, and is an Adjunct Associate Professor in the School of Information Technologies, University of Sydney, Sydney. His research interests include physiological parameter estimation from functional imaging, image registration, and optimizing use of the combination of functional/anatomic data.