

A New Way for Multidimensional Medical Data Management: Volume of Interest (VOI)-Based Retrieval of Medical Images With Visual and Functional Features

Jinman Kim, *Member, IEEE*, Weidong Cai, *Member, IEEE*, Dagan Feng, *Fellow, IEEE*, and Hao Wu, *Student Member, IEEE*

Abstract—The advances in digital medical imaging and storage in integrated databases are resulting in growing demands for efficient image retrieval and management. Content-based image retrieval (CBIR) refers to the retrieval of images from a database, using the visual features derived from the information in the image, and has become an attractive approach to managing large medical image archives. In conventional CBIR systems for medical images, images are often segmented into regions which are used to derive two-dimensional visual features for region-based queries. Although such approach has the advantage of including only relevant regions in the formulation of a query, medical images that are inherently multidimensional can potentially benefit from the multidimensional feature extraction which could open up new opportunities in visual feature extraction and retrieval. In this study, we present a volume of interest (VOI) based content-based retrieval of four-dimensional (three spatial and one temporal) dynamic PET images. By segmenting the images into VOIs consisting of functionally similar voxels (e.g., a tumor structure), multidimensional visual and functional features were extracted and used as region-based query features. A prototype VOI-based functional image retrieval system (VOI-FIRS) has been designed to demonstrate the proposed multidimensional feature extraction and retrieval. Experimental results show that the proposed system allows for the retrieval of related images that constitute similar visual and functional VOI features, and can find potential applications in medical data management, such as to aid in education, diagnosis, and statistical analysis.

Index Terms—Functional imaging, image segmentation, multidimensional features, region-based image retrieval.

I. INTRODUCTION

THE research and development of digital medical imaging, together with the rapid advancements in picture archiving and communication systems (PACS), have led to an increased demand for efficient medical-imaging data storage and retrieval techniques [1], [2]. The volume of medical data online in PACS often exceeds several terabytes, due to the rapid

growth in both the usage and the range of digital medical diagnostic modalities [3], [4]. It is well acknowledged that medical image databases, which enable access to a patient's historical data, including multidimensional medical images from previous examinations and the opportunity for statistical and comparative image analyses, are key components in preventive medicine and future diagnosis [4], [5]. However, these medical images are primarily indexed by text keywords that limit the image features to textual descriptions and retrievals to text-based queries. Although text-based retrieval is capable of supporting a high degree of image-content semantics, it is likely that text-based retrieval is unable to sufficiently describe the visual features of the images [6], [7]. The content-based image retrieval (CBIR) of medical images according to its domain-specific image features is an important alternative and complement to traditional text-based retrieval using keywords [3], [6], [8]–[16]. In recent years, various CBIR systems have been introduced for medical images. In medical CBIR systems, visual features commonly applied to broad images are unable to fully describe the characteristics and diagnostic meanings of these images [3], [8]. Furthermore, due to the differences within the medical-imaging modalities, the CBIR systems are designed and implemented to domain-specific medical-imaging modalities. Consequently, to extract features for a particular imaging modality, image-processing algorithms that are domain specific to the imaging applications must also be developed.

A common approach of feature extraction is to segment the images into regions (objects) based on a certain similarity criterion. Regions from the segmentation results can then be used in region-based queries for CBIR, enabling the user to include only the relevant regions when formulating a query [9], [10]. Chu *et al.* [11] described a knowledge-based image retrieval of computed tomography (CT) and magnetic resonance imaging (MRI) images where brain lesions were automatically segmented and represented within a knowledge-based semantic model, providing a mechanism for accessing and processing spatial, evolutionary, and temporal queries. Cai *et al.* [12] proposed a CBIR system for functional dynamic positron-emission tomography (PET) images of the human brain, where the segmented clusters of tissue–time activity curves from the temporal domain were used to measure the functional similarity for retrieval. In another study [13], the delineations of the region of interests were manually performed on the “key-frame” from

Manuscript received May 11, 2005; revised November 12, 2005. This work was supported in part by ARC and RGC grants.

J. Kim, W. Cai, and H. Wu are with the Biomedical and Multimedia Information Technology (BMIT) Group, School of Information Technologies, The University of Sydney, Sydney, NSW 2006, Australia.

D. Feng is with the Biomedical and Multimedia Information Technology (BMIT) Group, School of Information Technologies, The University of Sydney, Australia, and also with the Center for Multimedia Signal Processing (CMSP), Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong.

Digital Object Identifier 10.1109/TITB.2006.872045

the frame-stack of high-resolution CT (lung) images and were used as the features to represent the entire image. Other medical CBIR systems using segmentation to represent the regions have also been described, such as the retrieval of tumor shapes [14], clusters of microcalcification from mammography [15], and the shapes of regions in spine X-ray images [16]. The majority of these systems were based on medical-imaging modalities that are multidimensional [11]–[15]; however, the feature extraction and retrieval were all performed using two-dimensional slices, thus, not taking full advantage of the information available in the three-dimensional (3-D) spatial domain. In contrast, Guimond *et al.* [17] introduced the user-selected volume of interest (VOI) for the retrieval of pathological brain MRI images. The basic approach in this study was a density-based registration algorithm using global or local affine followed by free-form transformations for the best match between the user-defined VOI and images in the database. However, this approach relies on the VOI definition to be a subsection of the image, rather than the segmented regions, and therefore, the retrieval similarity measure was not region based.

In this study, we propose a 3-D VOI-based feature extraction and retrieval of multidimensional dynamic [^{18}F]2-*fluoro-deoxy-glucose* (FDG) PET images. As dynamic PET images consist of both 3-D spatial and one-dimensional (1-D) temporal domains, functional and visual features may be utilized that are domain specific and not available in other medical images. These images were segmented into VOIs consisting of voxels that have similar kinetic behaviors and were used to represent functional structures in the images. Visual and functional features were extracted from these VOIs, which allow the users to formulate region-based queries based on the extracted features for content-based retrieval, i.e., search for previous diagnosis and treatment results for cases with a cerebral tumor near the cerebellum of the brain. A prototype VOI-based functional image-retrieval system (VOI-FIRS) has been developed to demonstrate the multidimensional feature extraction and retrieval of dynamic PET images.

II. AUTOMATED VOI SEGMENTATION OF DYNAMIC PET IMAGES USING FUZZY C-MEANS CLUSTER ANALYSIS

In the segmentation of dynamic PET images, cluster analysis based on kinetic behavior has previously been found to be effective in classifying kinetic patterns [12], [18], [19]. In this study, the voxels in the dynamic PET images were automatically segmented using four-dimensional fuzzy c-means cluster analysis [20] into cluster groups based on voxel's kinetic behavior. A detailed description of the theory and implementation of the fuzzy c-means cluster analysis of dynamic PET images may be found in [21]. To simplify the cluster analysis, all frames were summed and low-count background areas in the summed image were removed (set to zero) by applying a threshold. Isolated voxels and gaps from thresholding were then eliminated by a $3 \times 3 \times 3$ morphological opening filter followed by a closing filter [22]. Tissue time activity curves (TTACs) were extracted for each of the N nonzero voxels in the summed image to form the kinetic feature vector comprising the voxel values at time t ($t = 1, 2, \dots, T$) where T is the total number

of time points. On the basis of the optimal image-sampling schedule technique [23]–[25], the TTACs were then reduced to the five temporal frames needed to identify the five parameters of the FDG model. The application of optimal image-sampling schedule considerably reduced the dimension of the feature vector while increasing the signal-to-noise ratio of the individual frames, thus, potentially improving the cluster analysis relying on the similarity measure of the feature vectors [25].

The fuzzy c-means cluster analysis was applied to assign each of the N feature vectors to a set number C of distinct cluster groups. For each cluster, centroids were assigned as the feature vectors of distinct, randomly selected nonbackground voxels. The value of each centroid voxel was replaced with the average of the $3 \times 3 \times 3$ surrounding voxels to avoid false selection of a noisy outlier that could have resulted in a cluster with a single member. The cluster analysis minimizes the objective function J , according to

$$J = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^P D_{ij}^2 \quad (1)$$

where P ($1 \leq P \leq \infty$) is a weighting exponent on each fuzzy membership, which determines the amount of fuzziness of the resulting classification, and u_{ij} is the membership degree of the i th feature vector in the cluster j . The similarity measure between the i th feature vector $\mathbf{f}_i(t)$ and the cluster centroid $\bar{\mathbf{f}}_{c_j}(t)$ of the j th cluster group c_j was calculated using the Euclidean distance D_{ij} given by

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

where $s(t)$ is a scale factor of time point t ($t = 1, 2, \dots, T$) equal to the duration of the t th frame divided by the total dynamic acquisition time. The scale factor gives more weight to the later, longer frames that contain more reliable data. The minimization of J was achieved by iteratively updating the u_{ij}

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left[\frac{D_{ij}}{D_{ik}} \right]^{\frac{2}{P-1}}} \quad (3)$$

and the cluster centroids $\bar{\mathbf{f}}_{c_j}(t)$

$$\bar{\mathbf{f}}_{c_j}(t) = \frac{\sum_{i=1}^N u_{ij}^P \mathbf{f}_i(t)}{\sum_{i=1}^N u_{ij}^P}. \quad (4)$$

Thus, a probabilistic weighting was assigned to every voxel i representing it to be likely a member of each cluster j . For any voxel i , the sum of the assigned membership degrees was 1.0, i.e., $\sum_{j=1}^C u_{ij} = 1.0$. The procedure was terminated when the convergence inequality

$$\max_{i,j} \{ |u_{ij}^{m+1} - u_{ij}^m| \} < \varepsilon \quad (5)$$

was satisfied, where m was the iteration step and $0 < \varepsilon < 1$. Upon convergence, a cluster map was created by assigning to each voxel a value equal to the cluster number for which it had the highest degree of fuzzy membership.

The optimal number of clusters was determined as follows. The fuzzy validation measure given in [26] was evaluated for

integer values of C in the range $(L - 3 < C < L + 3)$, where L was the number of tissue types expected to be present. A smaller value of the validation measure indicated a well-defined cluster scheme with more compact and more separate clusters. Additionally, mean-square error across clusters, calculated as the objective function J in (1) divided by C after termination of the cluster analysis, was treated as an indicator of the similarity of the voxels within the cluster group. Values of the parameter C , P , and ε were empirically determined (see Section VI).

From the cluster-analysis result, region-growing segmentation [22] was applied to the voxels in each cluster to construct the VOIs. The effect of this process was to group the voxels that were spatially connected (neighboring voxels), and separate the different structures that may have been classified into a cluster due to the similarity of voxel's kinetic behavior. After region growing, an erosion filter was applied to remove weakly connected voxels in the VOIs. Multiple VOIs may be formed from each cluster and for VOIs comprising less than 100 voxels were considered insignificant and discarded.

III. VOI FEATURE EXTRACTION

A. VOI Functional and Physiological Features

The mean TTAC values for each identified VOI were indexed into the feature database as numeric values. The physiological parameters in the VOI, such as the local cerebral metabolic rate of glucose [27], corresponding to the glucose consumption and energy requirement of functional components, were also estimated with weighted nonlinear least square algorithm [28] and indexed as numeric values.

B. VOI Visual Features

Our approach to VOI location was based on the transformation of the segmented images to the standard atlas. We utilize the anatomical standardization procedure of the 3-D stereotactic surface projection transformation in the NEUROSTAT package [29]–[31], which 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. The transformation was applied to the dynamic PET images, creating library files of transformation procedures. These library files were then applied to the corresponding segmentation results to warp the segmented images into the same standardized image frame of reference. After the transformation of the segmented images, the centroid moment for each VOI was calculated and indexed into the database. Another visual feature extracted was the volumes of the VOIs, measured in number of voxels.

C. Textual Attribute Features

In dynamic PET image acquisition, textural information regarding the image study is often stored in the image header. The header information usually consists of patient information, such as patient ID, name, sex, age, height, weight, and study number, and attributes related to the patient diagnosis, including physician name, physical examination data, medical history,

and pathological results. The information from the image header was parsed into individual textual attributes and indexed into the feature database.

IV. SIMILARITY MEASURES OF DYNAMIC PET FEATURES

A. Query by VOI Functional and Physiological Features

The TTAC feature vectors of the VOI can be searched in two ways. Firstly, the user can sketch a TTAC curve and adjust the values of individual time-activity points manually. Secondly, the user can select from previously user-defined TTAC curves saved in the database. The sketched or selected curve was then compared to the TTACs stored in the feature index database using the Euclidean distance in (2). Other physiological features can be searched by numeric differences.

B. Query by VOI Visual Features

The volume feature can be retrieved by measuring the absolute differences between the two volumes. In the similarity measure of the VOI location, the Zubal phantom [32], which consists of labeled anatomical structures, was used as the reference image. The labeled phantom was transformed into the standard co-ordinate system as with dynamic PET images (Section III-B). The similarity of the VOI location was measured using the 3-D spatial distance (in voxels) between the user-defined point from the Zubal atlas and the centroids of the VOIs indexed in the database.

C. Query by Textual Attribute Features

Textual features are measured for similarity by traditional keyword matching using the structured query language (SQL) [33]. Multiple attributes can be selected and combined using logical relationships such as “and,” “or,” “greater than,” etc.

D. Combination Feature Query

The query features of TTAC, volume, and location, described in Sections IV-A and IV-B, can be combined into a single query by assigning different weights to these features. This allows the user to formulate a query using multiple features and to prioritize specific features. For every exclusive VOI retrieved from all M queries, the combined query rank R_{CQ} was calculated based on the summation of the ranks from the individual queries according to

$$R_{CQ}(\text{VOI}_j) = \sum_i^M \frac{R(\text{VOI}_j, Q_i)}{R_T} w_i \quad (6)$$

where VOI_j is the j th exclusive VOI, R_T is the maximum number of returned VOI results common among all M queries (Q_1, Q_2, \dots, Q_N) , $R(\text{VOI}_j, Q_i)$ is the rank of the j th VOI from the i th query Q_i , and w_i (sum to 100%) is the user-defined weight assigned to Q_i . The returned VOIs from the combined queries were then ranked based on descending R_{CQ} measures. Textual patient attributes can also be combined without the weighting function.

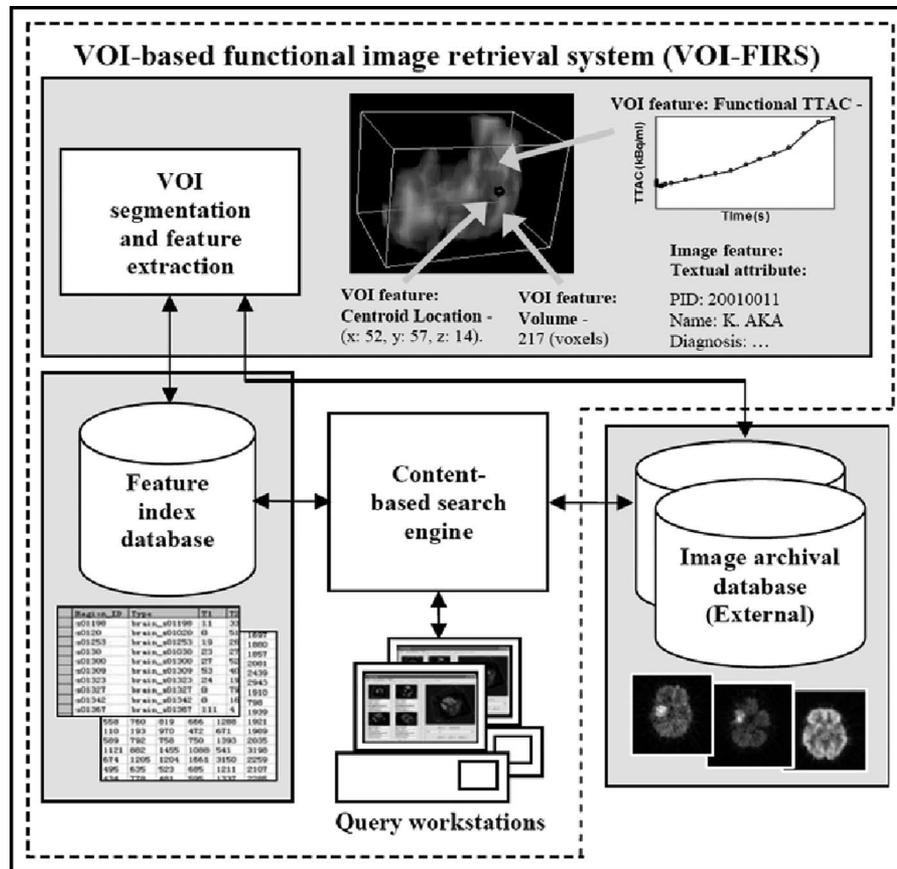


Fig. 1. System block diagram of the proposed VOI-FIRS. After VOI segmentation of dynamic PET images stored in image archival database, visual and functional features were extracted and then indexed in the feature index database (feature extraction exemplified using a sample VOI). The workstations are used to query the content-based search engine, which retrieves the images stored in the external image archival database.

V. DESIGN AND DEVELOPMENT OF VOI-FIRS

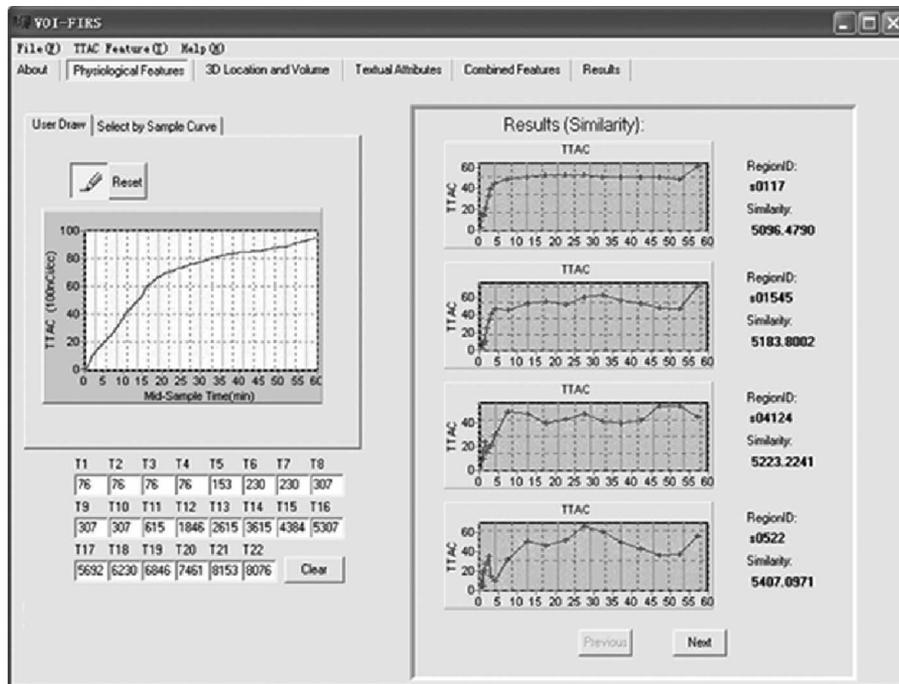
Fig. 1 illustrates the overall system block diagram of the VOI-FIRS for the content-based retrieval of dynamic PET images. The system is composed of a segmentation and feature extraction module and a content-based search engine. The dynamic PET images from the archival database (external storage) were initially segmented into VOIs. The features were then extracted from the VOIs and indexed as visual and functional features in the feature index database. The feature index database can be queried using the graphical user interface from the workstations. On the basis of the user query, the feature index database was searched and the images from the external image archive were retrieved.

The VOI-FIRS has been developed on a PC (Pentium IV Mobile, 2.0 GHz, 512 RAM, 32M graphics) with Microsoft Windows XP platform using Microsoft SQL server 2000, Borland C++ builder, and the visualization tool kit [34]. With this system configuration, the VOI-FIRS was responsive to the user interaction and in the retrieval process. Fig. 2 shows the graphical user interface of the query components in VOI-FIRS. There are four individual query entry windows corresponding to the four query options. With the “Query by functional and physiologic features” illustrated in Fig. 2(a), the user can manually draw the TTAC feature curve within the labeled grid, or se-

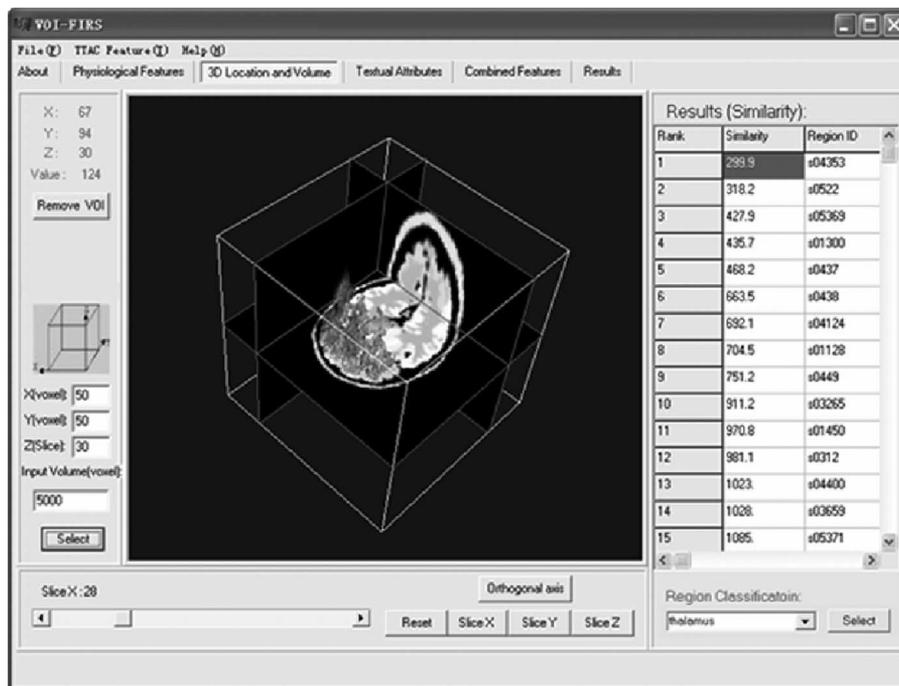
lect from a list of predefined sample TTACs (not shown). Once the selection has been made, the TTAC curve can be manually adjusted for individual TTAC points. As the TTAC curve is concentrated in the early temporal frames, the drawn curve can be zoomed for closer inspection. The “Query by VOI visual features” is illustrated in Fig. 2(b), where the user can select the location from the transformed Zubal phantom used as the standard brain atlas (center). The user can navigate in the 3-D viewing space (rotation, scaling, and translation) and change the viewing planes of saggital, coronal, and transaxial slices, providing both the conventional and 3-D orthogonal views. The location feature can be selected either by selecting a point on the Zubal phantom (center), setting voxel co-ordinates numerically (left center), or on the labeled Zubal structure (bottom right). The volume of the VOI can also be set numerically, which works independently to the location feature. The “Query by textual attribute features” and “Query by combined features” query windows (not shown) allow the users to input key words and weights assigned to features, respectively.

VI. EXPERIMENTAL RESULTS

To validate the proposed VOI-based feature extraction and retrieval, a feature index database was constructed consisting of more than 300 unique and independent VOIs from 13 dynamic



(a)



(b)

Fig. 2. Graphical user interface of the VOI-FIRS. (a) “Query by functional and physiological features” interface which shows the user-drawn TTAC curve (left) and the retrieved VOIs with their TTAC curves and similarity indices (right). (b) “Query by visual features” interface, exemplified with the rendered surface corresponding to the user-selected “prefrontal lobes” in the orthogonal Zubal slices.

FDG PET (SIEMENS ECAT 951R scanner) of clinical human brain studies. These studies included two tumor cases, three normal cases, and eight other neurological cases. The numbers of image slices for the dynamic PET studies was 31, with each slice acquired in 22 temporal frames comprising 6×10 s scans, and 4×0.5 , 1×2.0 , and 11×5.0 min scans. The dynamic im-

ages were decay corrected to the time of injection, attenuation corrected and then reconstructed using filtered back-projection. The reconstructed images were 128×128 with voxel dimensions of $1.841 \times 1.841 \times 3.375$ mm.

The result of applying fuzzy c-means cluster analysis to a patient study is shown in Fig. 3 with different gray levels used

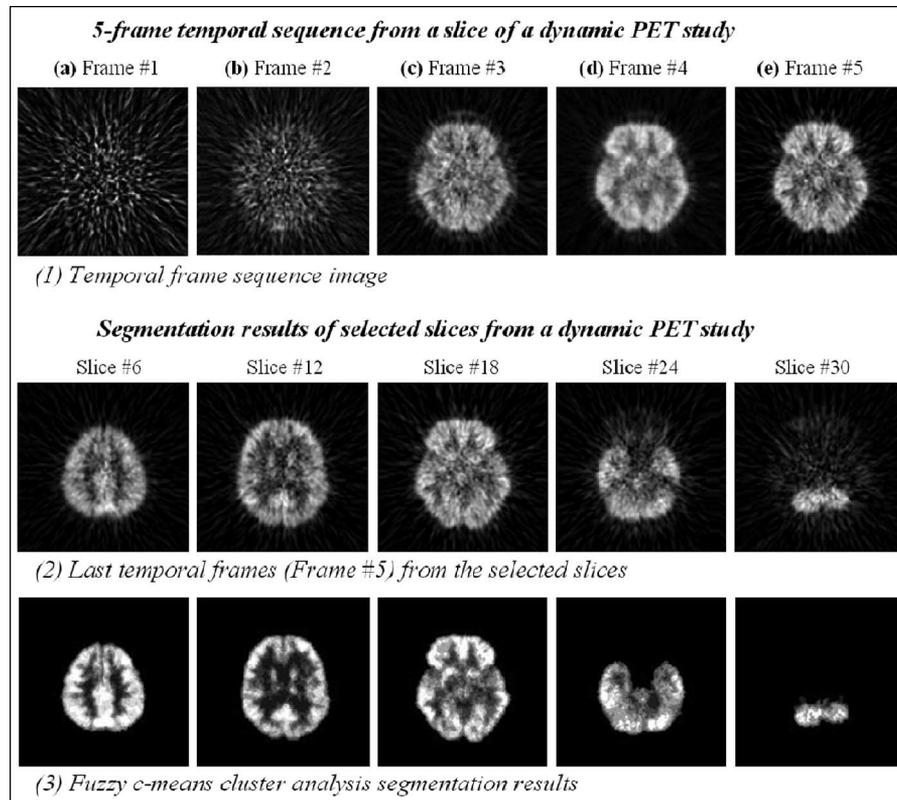


Fig. 3. (1) Dynamic five-frame temporal sequences using the optimal image-sampling schedule, respectively of (a)–(e) from the 18th slice (trans-axial) of a patient study; (2) selected slices out of 31, respectively of (a)–(e) from the patient study (last temporal frames); and (3) corresponding cluster analysis results. All images have been adjusted to their local maximum intensity ranges for display.

TABLE I
FUZZY VALIDATION AND THE MEAN-SQUARE ERROR MEASURES IN THE FCM CLUSTER ANALYSIS OF A DYNAMIC PET IMAGE

Validation Measure	Number of clusters (C)						
	4	5	6	7	8	9	10
Fuzzy validation ($\times 10$)	0.33	0.36	0.38	0.43	0.67	0.76	0.63
Relative differences (%)	n/a	6.21	6.04	10.34	54.88	22.53	-30.98
Mean square error	4.86	3.17	2.21	1.66	1.32	1.07	0.86
Relative differences (%)	n/a	-42.26	-23.81	-13.85	-8.60	-6.02	-5.46

to distinguish the identified clusters. Multiple VOIs were constructed from each cluster. From visual inspection, the cluster analysis appeared to have correctly separated the prominent tissue structures, such as the cortex and the thalamus. Table I shows the FCM cluster analysis validation as a function of cluster number C in the segmentation of a dynamic PET image in Fig. 3. In the fuzzy validation, there was a large decrease in performance occurring at $C = 8$ (largest relative difference). Although a small index of fuzzy validation indicates a cluster scheme with well-defined clusters, when C was small ($C < 6$), it resulted in an undesirable effect of merging functionally different clusters together. Thus, to maximize the classification of individual functional types, the optimal cluster number was taken to be $C = 7$, which corresponds to the index of fuzzy validation prior to the large increase. The mean-square error indices decreased monotonically with increasing C , with rapid

drop in performance before slowing down at $C = 7$. Based on tradeoff between these measures, the optimal value of C was calculated to be 7. The threshold value for the background removal (15%) was not found to be critical and moderate changes ($\pm 5\%$) had little effect on the results. Empirically derived cluster analysis parameter values of $P = 2.0$ and $\varepsilon = 0.1$ gave acceptable results. Variation of these parameters, $1 < P < 3$ and $0.05 < \varepsilon < 0.2$, appeared to have insignificant effect on the fuzzy validation and the mean-square error measures with 2%–3% differences in both validation measures from the variation of these parameters.

The graphical user interface of the VOI-FIRS for retrieved results is shown in Fig. 4. Each thumbnail result [Fig. 4(a)–(d)] can be individually navigated in orthogonal planes. The similarity indices from the retrieved VOIs are presented below the images. Any of the retrieved images can be enlarged and

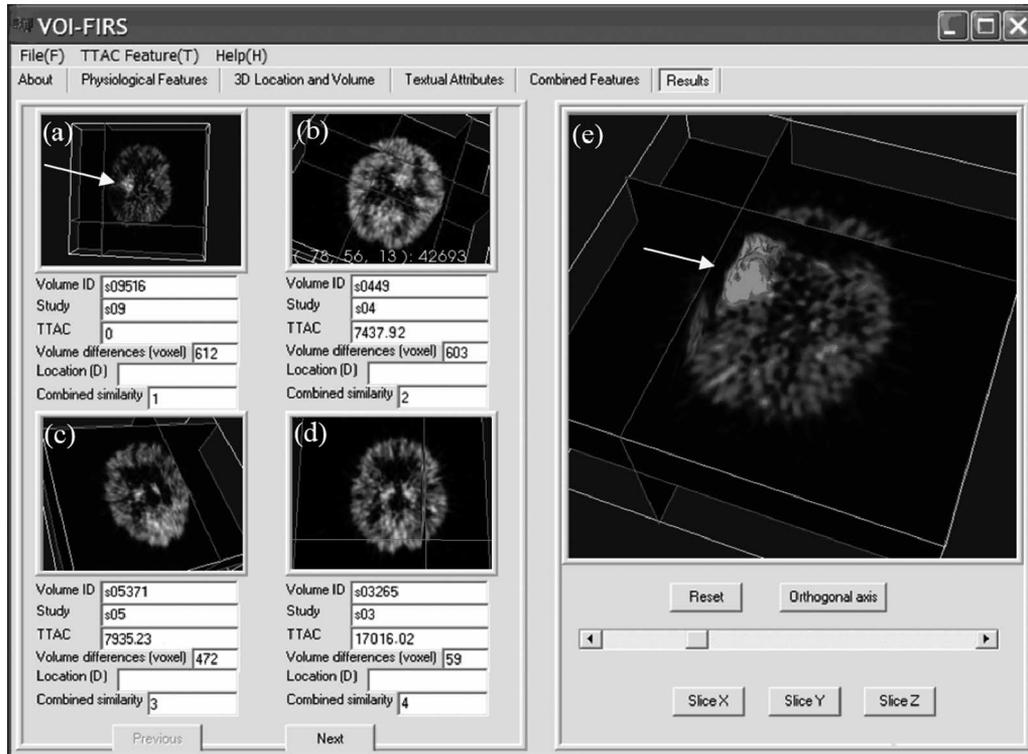


Fig. 4. The graphical user interface of the query results in VOI-FIRS. Thumbnails in the left side of the interface represent the ranked results respectively of (a)–(d). Right side is the enlarged view of the selected thumbnail result. In the enlarged view, the retrieved VOI (tumor VOI) can be surface rendered.

the VOI surface rendered [Fig. 4(e)]. The segmentation result from which the features were extracted can also be retrieved from the database for inspection of the segmented VOIs. In this example, the combination query was formulated to search for images having similar functional behavior and size apparent in a malignant brain tumor cases. The query was formulated to include a TTAC curve as a functional feature derived from an existing tumor VOI in the database, and 2000 voxels as the volume feature, with equal weights applied to the two features. The highest ranked result shown in Fig. 4(a) and enlarged in Fig. 4(e) with the VOI surface rendered is of a patient study with a prominent tumor (indicated by an arrow). The result shows that the query successfully identified VOIs with high kinetic behavior and user-defined volume, which is in agreement with the similarity indices for the query features shown below the images. As expected, the similarity index of the TTAC feature for the 1st-ranked VOI was 0, since the TTAC feature used in the query was derived from this VOI.

The formulated query used in Fig. 4 was modified to include an additional location feature to observe the behavior of the combined query retrieval. The location feature was defined as the center of the brain with voxel coordinates of 64, 64, 17 (x -, y -, z -axes) and equal weights were then distributed to the three features. The 12 highest combined ranked VOIs are tabulated in Table II. In the “Rank” columns, the numbers in bold is the ranks and the numbers inside the brackets are the weighted ranks. In all similarity measures, smaller indices represent better similarity. On the basis of this result, the addition of the location feature and subsequent redistribution of the feature weights had the effect

of retrieving VOIs, which were relevant to the query in all three features, with changes in combined ranks of all retrieved VOIs apart from the 1st combined ranked VOI.

The effect from the distribution of weights applied to the features is further exemplified to provide insights into the stability of the VOI-FIRS. Based on the result in Table II, the differences in the combined weighted ranks (column 1) among the VOIs can be used to indicate the change in VOI ranks from the redistribution of feature weights. As the VOIs are ranked according to the sum of the weighted ranks from individual features, it is likely that a VOI with large differences in weighted ranks to other VOIs (such as the first three VOIs to all the other VOIs), will not be affected by subtle changes in the distribution of feature weights. Indeed, by adjusting the weight to 40% for TTAC, and 30% for the volume and location in the same query, all the ranks of the VOIs were changed apart from the first three VOIs. The largest change in the combined rank occurred for VOI s04-37 with increase in rank from 12th to 7th. The raise in the rank was attributed to the fact that the weight of the high-ranked functional feature (3rd) was increased whereas the low-ranked location feature (56th) was decreased. The individual feature ranks calculated during the combined query can also be used to retrieve VOIs based on only one particular feature, for example, using only the functional feature from the query above, the retrieved VOIs will result in s09 (combined rank of 1), s06 (combined rank of 20), and s04 (combined rank of 3).

Another example of image retrieval is illustrated in Fig. 5. The sample TTAC, which approximates a pattern found in gray matter of PET images [Fig. 5(a)], and the location of

TABLE II
SIMILARITY MEASURES OF THE COMBINED QUERY RESULT USING THREE FEATURES WITH EQUAL WEIGHTS APPLIED

Combined Rank (weighted)	VOI Number (ID)	TTAC Similarity Measure		Volume Similarity Measure		Location Similarity Measure	
		Rank (weighted)	Index ($\times 1000$)	Rank (weighted)	Index	Rank (weighted)	Index
1 (7.67)	s09-516	1 (0.33)	0.00	9 (3.00)	612	13 (4.33)	8.37
2 (9.33)	s09-147	11 (3.67)	13.47	11 (3.67)	779	6 (2.00)	7.48
3 (13.67)	s03-265	16 (5.33)	17.34	2 (0.67)	59	23 (7.67)	10.49
4 (19.33)	s06-433	49 (16.33)	42.29	1 (0.33)	33	8 (2.67)	8.06
5 (20.00)	s04-38	5 (1.67)	7.39	15 (5.00)	922	40 (13.33)	16.31
6 (22.33)	s05-371	6 (2.00)	7.44	7 (2.33)	472	54 (18.00)	22.56
7 (23.33)	s01-447	20 (6.67)	19.73	14 (4.67)	914	36 (12.00)	15.62
8 (24.00)	s06-446	32 (10.67)	30.08	36 (12.00)	12332	4 (1.33)	7.35
=8 (24.00)	s06-523	40 (13.33)	36.62	31 (10.33)	10955	1 (0.33)	6.08
10 (24.67)	s04-49	7 (2.33)	7.94	8 (2.67)	603	59 (19.67)	30.53
=10 (24.67)	s06-414	14 (4.67)	17.02	55 (18.33)	21604	5 (1.67)	7.35
12 (25.00)	s04-37	3 (1.00)	6.53	16 (5.33)	1496	56 (18.67)	26.25

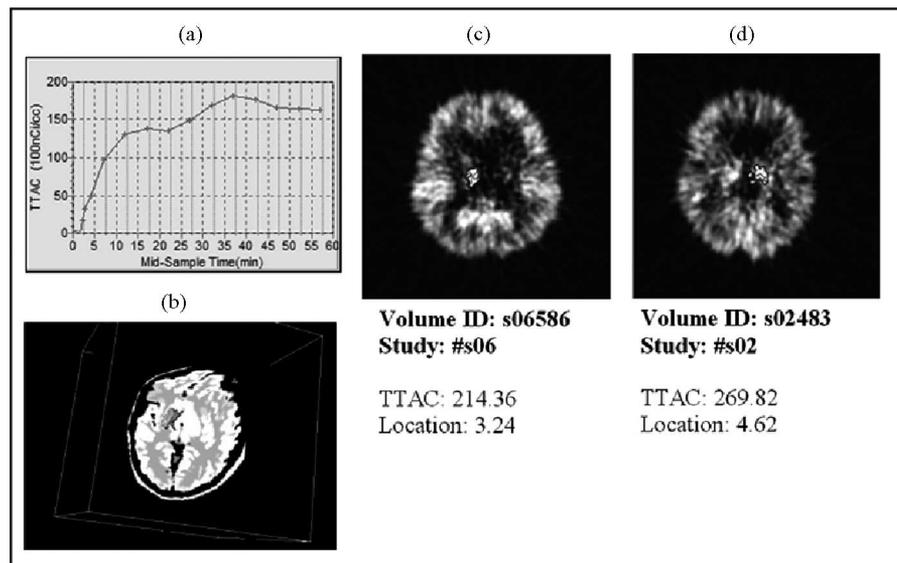


Fig. 5. (a) TTAC curve. (b) Selection of a location (right thalamus) in the standard atlas (Zubal). (c) Retrieved result with (a) and (b) as the query features. (d) Retrieved result with (a) and a different location feature (left thalamus) as the query feature.

the “right thalamus” selected from the labeled structures in the Zubal phantom [Fig. 5(b)] were set as the query features. Weighting was set to 50% for the functional feature and 50% for the location feature. The highest ranked retrieved VOI is shown in Fig. 5(c), where the query identified a VOI representing the right thalamus. Fig. 5(d) is the highest ranked result from changing the location feature to “left thalamus.” Utilizing the spatial properties of the PET images in the location feature, VOIs based on user-defined structures were identified, which may have not been possible from the functional feature alone.

VII. DISCUSSION AND FUTURE WORK

The VOI-FIRS demonstrated the use of multidimensional feature extraction and retrieval, enabling the use of 3-D visual and 1-D functional features within dynamic PET images for

content-based retrieval. Existing CBIR systems that utilize features, such as color and texture, may not be the optimal features when applied to dynamic PET images. Color is not captured in grayscale, and due to the statistical noise and tracer redistribution in dynamic PET images, texture is also unlikely to be a relevant feature. However, other visual features such as VOI shape and grayscale gradient may introduce additional searching capabilities, which could improve the retrieval performance. To represent the image features accurately, the segmentation of dynamic PET images is an important component for overall success of the content-based retrieval system. In this study, the application of fuzzy c-means cluster analysis was well suited in the segmentation of dynamic PET images for the construction of VOIs. However, this cluster analysis was based on the voxel’s

kinetic behavior (TTAC) alone and did not consider the potentially useful spatial information in the voxels. We are currently investigating the segmentation of dynamic PET images using both the spatial and temporal information, which may prove useful in the feature extraction and retrieval for CBIR.

The database of the VOI-FIRS was constructed using only 13 dynamic PET studies. Although only small database samples were indexed, all the images were of the same human body structure and acquired from the same scanner with identical imaging procedures (reconstruction algorithm, voxel dimension, spatial resolution, quantification, etc.). Therefore, we believe that our results were able to provide an insight into the capabilities of CBIR system in enhancing the searching capabilities of conventional text-based retrieval systems. For instance, by combining the text-based feature of "patients age <45 and >30" with the query set in Fig. 4, the retrieved result will only identify those patients with a VOI representing a tumor in the specified age group. Further studies will investigate the possible integration and evaluation of VOI-FIRS within a real hospital environment, which may enable evaluation of VOI-FIRS with large and varying dynamic PET study databases.

VIII. CONCLUSION

When performing content-based retrieval of multidimensional dynamic PET images, it is important to take into account the domain-specific requirements. Our VOI-FIRS has the advantage that visual and functional features extracted from the 4-D dynamic PET images were utilized in a region-based query for CBIR. This approach could facilitate identification of similar patient studies with particular pathologic process and potentially increase our understanding of underlying disease states and improve specificity in diagnosing diseases.

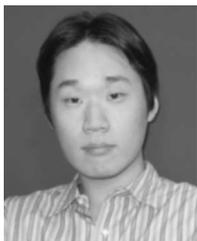
ACKNOWLEDGMENT

The authors are grateful for the generous suggestions by the referees and the staff at the Royal Prince Alfred hospital for supplying the clinical images.

REFERENCES

- [1] D. Feng, W. C. Siu, and H. J. Zhang., *Multimedia Information Retrieval and Management—Technological Fundamentals and Applications*, NY: Springer-Verlag, 2003.
- [2] H. K. Huang, *PACS: Basic Principle and Applications*, NY: Wiley-Liss, 1999.
- [3] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval system in medical applications—Clinical benefits and future directions," *Int. J. Med. Informatics*, vol. 73, pp. 1–23, 2004.
- [4] J. M. Bueno, F. Chino, A. J. M. Traina, and C. Traina, Jr., "How to add content-based image retrieval capability in a PACS," in *Proc. IEEE Computer Based Med. Syst.*, Jun. 4–7, 2002, pp. 321–326.
- [5] I. N. Bankman in *Handbook of Medical Imaging—Processing and Analysis*. San Diego, CA: Academic, 2000.
- [6] L. Zheng, A. W. Wetzel, J. Gilbertson, and M. J. Becich, "Design and analysis of a content-based pathology image retrieval system," *IEEE Trans. Inf. Technol. Biomed.*, vol. 7, no. 4, pp. 249–255, Dec. 2003.
- [7] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.
- [8] H. D. Tagare, C. C. Jaffe, and J. Duncan, "Medical image databases: A content-based retrieval approach," *J. Amer. Med. Inf. Assoc.*, vol. 4, pp. 184–198, 1997.
- [9] K. Vu, K. A. Hua, and W. Tavanapong, "Image retrieval based on regions of interest," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 4, pp. 1045–1049, Jul.–Aug. 2003.
- [10] B. G. Prasad, K. K. Biswas, and S. K. Gupta, "Region-based image retrieval using integrated color, shape, and location index," *Comput. Vis. Image Understand.*, vol. 94, pp. 193–233, 2004.
- [11] W. W. Chu, C. C. Hsu, A. F. Cardenas, and R. K. Taira, "Knowledge-based image retrieval with spatial and temporal constructs," *IEEE Trans. Knowl. Data Eng.*, vol. 10, no. 6, pp. 872–888, Nov.–Dec. 1998.
- [12] W. Cai, D. Feng, and R. Fulton, "Content-based retrieval of dynamic PET functional images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 4, no. 2, pp. 152–158, Jun. 2000.
- [13] C. Shyu, C. Brodley, A. Kak, A. Kosaka, A. Aisen, and L. Broderick, "ASSERT: A physician-in-loop content-based image retrieval system for HRCT image databases," *Comput. Vis. Image Understand.*, vol. 75, pp. 111–132, 1999.
- [14] P. Korn, N. Sidropoulos, C. Faloutsos, E. Siegel, and Z. Protopapas, "Fast and effective retrieval of medical tumor shapes," *IEEE Trans. Knowl. Data Eng.*, vol. 10, no. 6, pp. 889–904, Nov.–Dec. 1998.
- [15] I. El-Naqa, Y. Yang, N. P. Galatsanos, R. M. Nishikawa, and M. N. Wernick, "A similarity learning approach to content-based image retrieval: Application to digital mammography," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1233–1244, Oct. 2004.
- [16] S. Antani, D. J. Lee, R. Long, and G. R. Thoma, "Evaluation of shape similarity measurement methods for spine X-ray images," *J. Vis. Commun. Image Represent.*, vol. 15, pp. 285–302, 2004.
- [17] A. Guimond and G. Subsol, "Automatic MRI database exploration and applications," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 11, pp. 1345–1365, 1997.
- [18] 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, pp. 200–207, Feb. 2002.
- [19] H. Guo, R. Eanaut, K. Chen, and E. Rieman, "Clustering huge data sets for parametric PET imaging," *Biosystems*, vol. 71, pp. 81–92, 2003.
- [20] J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. Norwell, MA: Kluwer, 1981.
- [21] J. Kim, W. Cai, D. Feng, and S. Eberl, "An objective evaluation framework for segmentation techniques of functional positron emission tomography studies," in *IEEE Proc. Nucl. Sci. Symp. Med. Imag. Conf.*, vol. 5, Oct. 16–22, 2004, pp. 3217–3221.
- [22] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [23] 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.
- [24] D. Feng, W. Cai, and R. Fulton, "An optimal image sampling schedule design for cerebral blood volume and partial volume correction in neurologic FDG-PET studies," *Aust. N. Z. J. Med.*, vol. 28, p. 361, 1998.
- [25] 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. 32, no. 4, pp. 249–256, 1985.
- [26] X. L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 8, pp. 841–847, Aug. 1991.
- [27] S. C. Huang, M. E. Phelps, E. J. Hoffman, K. Sideris, C. J. Selin, and D. E. Kuhl, "Non-invasive determination of local cerebral metabolic rate of glucose in man," *Amer. J. Physiol.*, vol. 238, pp. e69–e82, 1980.
- [28] D. Feng, D. Ho, K. Chen, L.-C. Wu, J.-K. Wang, R.-S. Liu, and S.-H. Yeh, "An evaluation of the algorithms for determining local cerebral metabolic rates of glucose using positron emission tomography dynamic data," *IEEE Trans. Med. Imag.*, vol. 14, no. 4, pp. 697–710, Dec. 1995.
- [29] 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.
- [30] 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.
- [31] 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.

- [32] 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.
- [33] J. J. Patrick, *SQL Fundamentals*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [34] W. Schroeder, K. Martin, and W. Lorensen, *The Visualization Toolkit: An Object-Oriented Approach to 3-D Graphics*. Englewood Cliffs, NJ: Prentice-Hall, 1996.



Jinman Kim (S'02–M'06) received the B.S. (honors) degree in computer science and technology from the University of Sydney, Sydney, Australia, in 2001, where he is pursuing the Ph.D. degree in information technologies.

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

He was a Postdoctoral Research Associate at the Centre for Multimedia Signal Processing (CMSP), Hong Kong Polytechnic University, Hong Kong. He joined the School of Information Technologies, University of Sydney, Sydney, Australia as a Lecturer, where he is now a Senior Lecturer. His research inter-

ests 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 (EECS) from the Shanghai JiaoTong University, Shanghai, China, in 1982, and the M.S. degree in biocybernetics and the Ph.D. degree in computer science from the University of California, Los Angeles (UCLA), in 1985 and 1988, respectively.

He has worked as an Assistant Professor at the University of California, Riverside, after which he joined, at the end of 1988, as Lecturer, Senior Lecturer, Reader, Professor, and Head of Department of

Computer Science/School of Information Technologies, University of Sydney, Sydney, Australia. He is currently an Associate Dean of the Faculty of Science, University of Sydney; Honorary Research Consultant, Royal Prince Alfred Hospital, Australia; Chair-Professor of Information Technology, Hong Kong Polytechnic University, Hong Kong; Advisory Professor, Shanghai JiaoTong University; Guest Professor, Northwestern Polytechnic University, CA, Northeastern University and Tsinghua University, China. His research area is biomedical and multimedia information technology. He has published over 400 scholarly research papers, pioneered several new research directions, made a number of landmark contributions in his field with significant scientific impact and social benefit.

He is a fellow of ACS, ATSE, HKIE, and IEE. He is the Founder and Director of the Biomedical And Multimedia Information Technology Research Group, and the Chairman of the IFAC-TC-BIOMED. He is also the Special Area Editor of IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE. He was the recipient of the Crump Prize for Excellence in Medical Engineering.

Hao Wu (S'04) received the B.S. degree in computer science from Wuhan University, Wuhan, China, in 2001, and the masters degree in information technology from the University of Sydney, Sydney, Australia, in 2004.

His research interests include image processing and content-based image retrieval.