4 CONTENT-BASED MEDICAL IMAGE RETRIEVAL

Dr. Weidong Cai[1], Dr. Jinman Kim[1], Prof. David Dagan Feng[1,2]
[1]*BMIT Research Group, School of Information Technologies, The University of Sydney*
[2]*Dept. of Electronic & Information Engineering, The Hong Kong Polytechnic University*

Chapter Contents

4.1 INTRODUCTION

In the past three decades, various medical imaging techniques, as introduced in Chapter 1, have advanced rapidly, providing powerful tools for patient diagnosis, treatment planning, medical reference, and training. Medical image data have been expanded rapidly in *quantity*, *content* and *dimension* – due to an enormous increase in the number of diverse clinical exams performed in digital form and to the large range of image modalities available [1-3]. It has, therefore, led to an increased demand for efficient medical image data retrieval and management. In current medical image databases, images are mainly indexed and retrieved by alphanumerical keywords, classified by human experts. However, purely text-based retrieval methods are unable to sufficiently describe the rich visual properties or features inside the images content, and therefore pose significant limitations on medical image data retrieval [4-6]. The ability to search by medical image content is becoming increasingly important, especially with the current trend toward evidence-based practice of medicine [6, 7].

In this chapter, we present an overview of current content-based medical image retrieval (CBMIR) techniques. We first give an introduction into typical generic content-based image retrieval (CBIR) in Section 4.1.1, including its key components: image feature extraction, similarity comparison, indexing scheme, and interactive query interface, and followed by a short review of the major image visual features, i.e., color, texture, shape, and spatial relationships in Section 4.1.2. Then we briefly address the need for the CBMIR and the related challenges in Section 4.1.3. The major techniques used in the CBMIR are reviewed in details with four different categories, i.e., retrieval based on physical visual features in Section 4.2, retrieval by geometric spatial features in Section 4.3, retrieval by combination of semantic and visual features in Section 4.4, and retrieval based on physiological functional features in Section 4.5. Finally, concluding remarks are drawn in Section 4.6.

4.1.1 Fundamentals of Content-based Image Retrieval (CBIR)

The recent escalating use of digital images in diverse application areas such as medicine, education, remote sensing and entertainment has led to enormous image archives and repositories, and therefore requires for effective image data management and retrieval [8-12]. It is just similar to the rapid increase in the amount of alphanumeric data during the early days of computing, which led to the development of database management systems (DBMS) by organizing data into interrelated collections for convenient information retrieval and storage [11]. Early work on image retrieval was still based on the textual annotation of images, i.e., images were first manually annotated by keywords or descriptive text and organized by topical or semantic hierarchies in traditional DBMS to facilitate easy access based on standard Boolean queries. However, such purely text-based method poses significant limitations in image retrieval. Manual annotation is subjective, time-consuming and prohibitively expensive, and the sheer content volume of very large image databases is simply beyond the manual

indexing capability of human experts. Furthermore, many visual features in images, such as irregular shapes and jumbled textures, are extremely difficult to describe in text. Such text-based approach also limits the scope of searching to that predetermined by the author of the system and leaves no means for using the data beyond that scope.

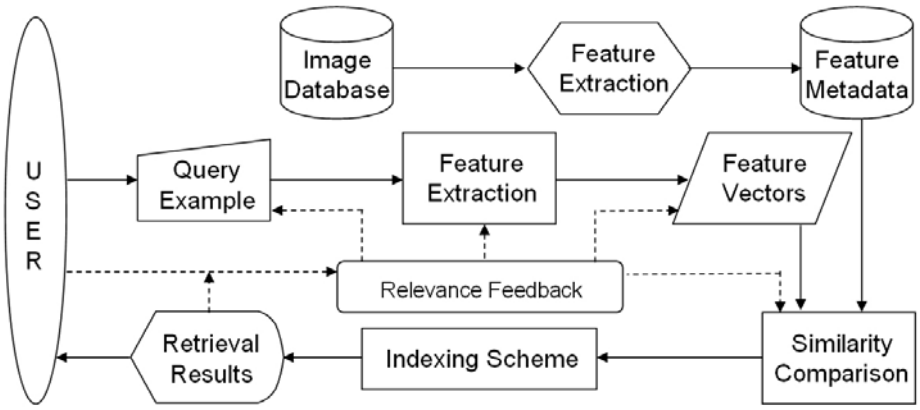

Figure 4.1 General architecture of the CBIR system

In contrast to traditional text-based approaches which perform retrieval only at a conceptual level, the recently developed content-based image retrieval (CBIR) methods support full retrieval by visual content / properties of images, i.e., retrieving image data at a perceptual level with objective and quantitative measurements of the visual content and integration of image processing, pattern recognition and computer vision [9, 10, 13, 14]. A typical CBIR system is depicted in Figure 4.1. Firstly, for each image in the *image database*, its visual contents are extracted – a set of distinguishing features (a *multi-dimensional feature vector*) is pre-computed via *off-line feature extraction* process. The feature vector is then stored in a *feature metadata repository*. To retrieval images, the *user* submits a *query example image* to the system and the example image is then converted into an *internal feature vector* via *on-line feature extraction* process. The similarity or closeness between the feature vector of the user's query image and those feature metadata items is calculated and ranked during *similarity comparison* stage. Retrieval is performed by applying an *indexing scheme* which can be used to support fast retrieval and to make the system scalable to large image databases. If the *retrieval results* in response to the query are not fully satisfactory, the user can give some positive or negative feedback to the system and the modified query can be resubmitted via the interactive *relevance feedback*. Such feedback-retrieval cycle can be repeated until the user is satisfied with the retrieval results. Among various components mentioned above, feature

extraction, similarity comparison, indexing scheme, and interactive query interface are four key issues in any CBIR system, and will be given in more details in the rest of this section.

Feature Extraction – It is the basis and most important component of the CBIR system. In a broad sense, features may include both text-based features (key words, annotations) and visual features (color, texture, shape, spatial relationships). However, since there already exists rich literature on text-based feature extraction in the traditional DBMS and information retrieval research communities, this chapter will be confined to the general visual feature extraction. Representation of images needs to consider which features are most useful for representing the contents of images and which approaches can effectively code the attributes of the images. Due to perception subjectivity, for any given feature, there does not exist only a single best representation, instead, multiple representations are used for characterizing the feature from different perspectives. It is well acknowledged that a good visual feature representation should be invariant to the accidental variance introduced by the imaging process (e.g., the variation of the illuminant of the scene taken in the image). There is a tradeoff, however, between the invariance and the discriminative power of the visual feature, since a very wide class of invariance may lose the ability to discriminate between essential differences [13]. A short review of the major image visual features used in the CBIR will be given in the next section. Moreover, a visual feature can be either global or local. If the feature extraction is applied on the whole image, the derived content features then becomes *global* features. In order to obtain more selective features at a finer resolution, the image is often divided in parts (sub-areas or homogenous regions) before features are computed from each part – that is so called *local* feature extraction. The easiest way to divide an image is to partition the image into equal-size blocks or tiles [13]. Such simple partition, however, does not generate perceptually meaningful or salient regions. A better method is to divide the image into homogenous regions according to some criterion using segmentation algorithms [14]. This is particularly true in medical imaging studies: it is uncommon that a condition or disease will alter an image over its entire spatial extent, more often than not, diagnostic features of interest manifest themselves in local regions. In a screening situation, the radiologist scans the entire image and searches for features that could be associated with disease; in a diagnostic situation, however, the medical expert concentrates on the region of suspected abnormality, and examines its characteristics to decide if such region of interest (ROI) exhibits signs related to a particular disease [15].

Similarity Comparison – To measure the similarity, the general approach is to represent the image features as multidimensional vectors. Selection of metrics has a direct impact on the performance of a CBIR system. A similarity comparison function maps between pairs of feature vectors and a positive real-valued number, which is chosen to be representative of the visual similarity between two images. Therefore, the retrieval result is not a single image but a list of images ranked by their similarities with the query image. Various similarity comparison approaches have been developed for image retrieval based on empirical estimates of the distribution of features, including *Minkowski-Form Distance* (a generic form of the well-known *Euclidean Distance*), *Mahalanobis Distance*, *Quadratic Form* (*QF*) *Distance*,

Proportional Transportation Distance, *Earth Mover's Distance*, *Kullback-Leibler* (*KL*) *Divergence* and *Jeffrey-Divergence* (*JD*). Readers are referred to [13] for a detailed review.

Indexing Scheme – When the number of images in the database is small, a sequential linear search can provide a reasonable performance. However, with large image collections, indexing support for similarity-based queries becomes necessary and can help in avoiding sequential scanning. Index structures ideally filter out all irrelevant images by checking image attributes with the user's query, and retain only relevant images without analyzing the entire database. Retrieved images are ranked in order of similarity to a query. Some commonly used multi-dimensional indexing approaches include: *linear quad-trees* [16], *K-D-B tree* [17], *hB-tree* [18], *R-tree* [19, 20] and its variants *R+-tree* [21], *R*-tree* [22], *X-tree* [23], *TV-trees* [24], and *SS-trees* [25]. Most of these indexing approaches perform reasonably well for a small number of dimensions, however explore exponentially with the increasing of the dimensionality and eventually reduce to sequential searching. One of the commonly used methods for addressing this problem is applying dimension reduction techniques, such as Principal Component Analysis (PCA), an optimal technique that linearly maps input data to a coordinate space such that the axes are aligned to maximumly reflect the variations in the data [26]. Some very good reviews and comparisons of various indexing techniques can be found in [27, 28].

Interactive Query Interface – An interactive retrieval interface allows the user to formulate and modify queries. The ability for the user to express his/her search needs accurately and easily is crucial in any CBIR system. The most appealing paradigm in many ways is *query-by-example*: providing a sample of the kind of output desired and asking the system to retrieve further examples of the same kind. Virtually all current CBIR systems now offer such searching, where the user submits a query image and the system retrieves and displays thumbnails of some closest-matching images in the database. However, the user will not always have an example image on hand. Several alternative query formulation approaches have been proposed, such as *category browsing*, *simple visual feature query*, *feature combination query*, *localized feature query*, *query-by-sketch*, *user-defined attribute query*, *object relationship query*, and *concept queries* [29]. The ability to refine searches online in response to user indications of relevance, known as relevance feedback, is particularly useful for improving the effectiveness of the CBIR systems interactively [30, 31]. The main idea of relevance feedback is to use positive and negative examples from the user to improve system performance. For a given query, the retrieval system firstly returns a list of ranked images based on predefined similarity metrics. The user then labels the retrieved images as relevant to the query as positive examples or not relevant as negative examples. The system will subsequently refine the retrieval results based on the user's feedback by a certain learning algorithm and return a new ranked list of images to the user. This process can continue to iterate until the user is satisfied. The relevance-feedback strategies help to alleviate the semantic gap [14] between low-level visual features and high-level semantic features, since it allows the CBIR systems to learn the user's image perceptions. The learning algorithm usually deals with small training samples (typically less than 20 per round of interaction), asymmetry in training sample (a few negative examples are normally fed back to the system), and real time requirement (it should be fast enough to support real-time user interaction) [32, 33]. The

commonly used relevance-feedback learning algorithms include *genetic algorithms* [34], *weight-based learning approaches* [35], *Bayesian probabilistic methods* [36], and *Support Vector Machines* (*SVM*) [37].

4.1.2 Image Features in CBIR

Clearly, the key to the CBIR framework lies in the feature extraction, in which the quantitative image features, computed automatically, are used to characterize image content. In essence, image features can be classified into *general visual features* and *domain-specific semantic features*. General visual features typically include primitive image information which refers to "the make-up of an image", such as color, texture, shape, and spatial relationships. Domain-specific semantic features, on the other hand, are application dependent and mainly consist of abstract image information which refers to "the meaning of an image", describing high-level image semantic content in specialized domains. This section will briefly introduce the general visual features which can be used in most CBIR applications. Readers are referred to [8-10, 12, 14, 29] for detailed reviews on visual feature extraction. Typically, general visual features can be further classified into two categories: *physical visual features* including color and texture, and *geometric spatial features* such as shape and spatial relationships. The domain-specific semantic features, which can be obtained either by textual annotation or by complex inference procedures based on image visual content, will be covered in Section 4.4, concentrating on the medical domain application – CBMIR.

Color is the most frequently used general visual feature for CBIR due to its invariance with respect to image scaling, translation and rotation, and to its three-color-component values (such as RGB, HSV, CIE L*a*b* or CIE L*u*v*) which make its discrimination potentiality superior to the single-gray intensities of images [8, 9, 13]. There are many color features that have been developed for the CBIR, such as: *Color Histogram* – a most effective color representation with the distribution of the number of pixels for each quantized color bin locating in three different color components [38, 39]; *Color Moments* – very compact color representations with three low-order moments (mean, variance and skewness) for each color component [26, 40]; *Color Coherence Vectors* (*CCV*) – incorporating spatial information into the color histogram (histogram refinement) [41]; *Color Correlogram* – a color descriptor characterizing both color distributions of pixels and the spatial correlation of pairs of colors [42]; and *HDS-S* – a color structure descriptor for capturing local color image structure based on the MPEG-7 HMMD (hue-min-max-difference) color space [43, 44].

Texture is a powerful discriminating visual feature which has been widely used in pattern recognition and computer vision for identifying visual patterns with properties of homogeneity that cannot result from the presence of only a single color or intensity [45]. Texture presents almost everywhere in nature – the size of the image patch, the number of distinguishable grey-level primitives and the spatial relationships between these primitives, are all interrelated elements which characterize a texture pattern [10]. Some commonly used texture features are:

co-occurrence matrices with 14 texture descriptors for capturing the spatial dependence of gray-levels [46]; *Tamura features* with six visual texture components designed in accordance with psychological studies on the human perception of texture [47]; *Run-length matrices* for quantifying the coarseness of texture in specified directions [48]; *Wavelet Transform coefficients* representing frequency properties of texture patterns – including pyramid-structured and tree-structured wavelet transforms [49, 50]; *Gabor Filter* as an orientation and scale tunable edge and bar / line detector [51, 52]; *Wold decomposition* providing perceptual properties with three components – harmonic (repetitiveness), evanescent (directionality), and indeterministic (randomness) [53, 54]; *Markov random field (MRF)* [55-57]; *Fourier power spectrum* [58]; *Fractal dimension* [59]; and *shift-invariant principal component analysis (SPCA)* [60].

Shape can be used to identify an object or region as a meaningful geometric form. To humans, perceiving a shape is to capture prominent / salient elements of the object or region [10]. Therefore, shape features in an image are normally represented after that image has been segmented into objects or regions. Due to the difficulties in fully-automated image segmentation [13] and the variety of ways a given 3D object can be projected into 2D shapes in 2D images, the CBIR based on shape features is considered to be one of the most challenging tasks and has usually been limited to specific applications where objects or regions are readily available [14]. In general, shape representation techniques fall in two broad categories: *boundary-based* and *region-based* approaches. The boundary-based approaches work on the outer boundary of the shape, and the shape descriptors in this category include: *Fourier Descriptor* which describes the shape of an object with the Fourier transform of its boundary [61, 62]; *Turning Function* for comparing both convex and concave polygons [63]; *Finite Element Method (FEM)* with a stiffness matrix and its eigenvectors [64]; *Curvature Scale Space* (*CSS*) [65, 66]; *Chord-Length Statistics* [67]; *Chain encoding* [68]; *Beam Angle Statistics* (*BAS*) [69]; and *Wavelet Descriptor* [70]. For the region-based approaches, the commonly used shape descriptors are: *Invariant Moments* – a set of statistical region-based moments [71, 72]; *Zernike moments* [68]; generalized complex moments [73] and *morphological descriptors* [74].

Spatial relationships between multiple objects or regions in an image usually capture the most relevant and regular part of the information in the image content [10, 13], and are very useful for image retrieval and searching. The spatial relationships can be divided into two classes: *directional* (or *orientation*) *relationships* and *topological relationships*. Directional relationships capture relative positions of objects with respect to each other, such as “left”, “above” and “front”, and are usually calculated through objects’ centroids or bary-centers. Topological relationships describe neighborhood and incidence between objects, like “disjoint”, “adjacency” and “overlapping”, and are calculated via objects’ shapes. The most commonly used approach to describe spatial relationships is the *Attributed-Relational Graph (ARG)* [75, 76] in which objects are represented by graph nodes, and the relationships between objects are represented by arcs between such nodes. Another well-known approach called the *2D strings method* is based on symbolic projection theory, and allows a bidimensional arrangement of a set of objects to be encoded into a sequential structure [77]. In addition to

the ARG and 2D strings methods, *spatial quad-tree* [78] and *symbolic image* [79] are also used to represent spatial relationships.

4.1.3 CBIR in Medical Domain (CBMIR)

In the medical domain, the majority of acquired medical images are currently stored with a limited text-based description of their content. As image databases expand, it is becoming increasingly apparent that these simple text-based descriptions are inadequate for the proper search and retrieval of medical images. As a consequence, valuable diagnostic and prognostic information contained in such databases remains unusable, and it leads to an increased demand for efficient retrieval techniques which can tap the expertise contained in these databases [80]. In previous sections of this chapter, the CBIR has been shown to be a viable alternative to text-based image retrieval, with the abilities to search for an image depending on metrics for comparing image with visual / spatial properties that can match human judgments of similarity. It is therefore very natural to apply such CBIR to medical domain (CBMIR) by retrieving medical images according to their domain-specific image features, providing an important alternative and complement to the traditional text-based retrieval. The potential benefits emanating from the CBMIR range from clinical decision support to medical education and research [6]. Clinical knowledge has shown that visual characteristics of medical images have strong effect on diagnosis [81]. Therefore, diagnosis by comparing past and current medical images associated with pathologic conditions has become one of the primary approaches in case-based reasoning or evidence-based medicine [7], while the clinical decision-making process can be beneficial to find other images of the same modality, the same anatomic region of the same disease [6]. The CBMIR could aid on such diagnosis in the following way: by observing an abnormality in a diagnostic image, the physician can query a database of known cases to retrieve images (and associated textual information) that contain regions with features similar to what is observed in the image of interest. With the knowledge of disease entities that match features of the selected region, the physician can be more confident of the diagnosis, and may be to expand the differential diagnosis to include pathological entities not previously considered. Here the CBMIR provides relevant supporting evidence from prior known cases, offering the physician with training-set examples that are close to his/her decision boundary, along with the correct class labels (proven pathology) for these examples. As such, a less experienced practitioner can benefit from this expertise in that the retrieved images, if visually similar, provide the role of an expert consultant [82-84]. The CBMIR could be used to present cases that are not only similar in diagnosis, but also similar in appearance and in cases with visual similarity but different diagnoses. It would therefore, be useful as a training tool for medical students, residents, and researches to browse and search large collections of disease related illustrations using their visual attributes [6]. The success of CBMIR will open up many new vistas in medical services and research, such as disease tracking, differential diagnosis, non-invasive surgical planning, clinical training, and outcomes research [85].

Albeit the need is clearly identified, developing the CBMIR systems imposes several distinct challenges from the CBIR for general images, and requires a thorough understanding of the nature and requirements of medical images in the following aspects [2, 5, 6, 85-88]: (1) Medical image data are *heterogeneous* in how they are collected, distributed, and displayed. Images are acquired from different modalities and in different settings in terms of position, resolution, contrast, and signal to noise ratio. Inside one modality, the tuning of an imager may lead to significantly varying images, for example, an MRI image may be used for acquisition of completely different anatomical and functional information, a PET or SPECT imaging scan may be conducted for different organ studies with different tracer settings. A hospital can generate thousands of diverse images of different modalities for different clinical studies everyday. Retrieval from these heterogeneous images is much more complex as compared to a single image modality and the existing CBIR approaches hardly address all medical image properties; (2) Most of medical images, except histological, dermatoscopic and endoscopic color images, are intensity-only-images and represented in *gray scale*. The color features frequently used in the CBIR thus cannot be applied; (3) Medical image are usually *low resolution* and *high noise* images, and therefore difficult to automatically analyze for extracting visual features; (4) Many human organs are *soft tissue* or *non-rigid bodies*. The diseased tissue often has *no regular shape* and there is no clear boundary with respect to the surrounding healthy part. Automatic and accurate image segmentation of these organs or lesion regions is difficult to achieve; (5) A large fraction of medical images captures human anatomy that is in nature a *three-dimensional* (*3D*) *structure*, and provides an additional information not available in 2D images. Therefore, it requires an additional complex registration procedure before volumetric image comparisons and image feature extraction; (6) Staging of the disease state and the monitoring of patient progress over time are fundamental to diagnostic and therapeutic decisions and to outcome assessment in long term follow up. The CBMIR system is required to have the ability to define and track *temporal relation* in the medical image sets of a patient taken at different periods, together with the medical history of that patient; and (7) Careful treatment to medical images is required due to patient privacy and other legal constraints. Such *security* and administrative barriers hinder CBIR research within the medical domain [85].

In recent year, various CBMIR approaches have been developed and mainly integrated in research prototypes. In a broad sense, these retrieval techniques can be classified into four different categories according to the key image features used: (1) retrieval based on physical visual features such as color and texture; (2) retrieval based on geometric spatial features such as shape, 3D volumetric features and spatial relationships; (3) retrieval by combination of semantic and visual features including semantic pathology interpretation approaches and generic model-based methods; and (4) retrieval based on physiological functional features such as the dynamic activities of glucose metabolism in human brain images. Details of these techniques will be given in subsequent sections.

4.2 CBMIR BY PHYSICAL VISUAL FEATURES

4.2.1 Retrieval based on Color

As mentioned in Section 4.1.2, color is the most extensively used low-level feature for the CBIR. However, since the majority of medical images are intensity-only-images carrying less information than color images, the color-based retrieval would only be applicable to medical images based on light photography, where color is an inherent feature and any deviations or changes in the color of a particular sample from normal sample can have significant medical implications [4, 89]. Such medical images that could benefit from the color-based CBMIR include: *histological* images, *dermatoscopic* images, *endoscopic* images, and *tongue* images, in which the well-known *color moments* and *color histogram* approaches are often adopted.

Histological images are taken via light microscopy, and can be used for assisting the pathologist to observe and analyze the fine details of biological cells and tissues. These images usually possess unique color signatures, including various subtle changes in color such as jaundice, congestion, changes in exudation and effusion, and pigmentation [4, 90]. In [80], the mean optical density and color moments to RGB components were calculated separately for the cytoplasm, nucleus, and nucleolus, and adopted as specific local color descriptors in a content-based cell image retrieval system. As a significant color descriptor, the histogram of color distribution has been used in retrieving cytohistological breast carcinoma images [91], breast cancer biopsy slide images [92], and microscopic pathology images of prostate, liver, and heart [93]. In [94], the local RGB color histograms were applied as the coarse features for pre-partitioned high-resolution histological images of the gastrointestinal (GI) tract. Since color characteristics in stained tissue images are prominent within these coarse structures, the extracted color histograms make themselves an ideal coarse detector for the iconic content analysis and retrieval, especially when an image database contains a large number of high-resolution images [94-97].

Dermatoscopic imaging is a technique that allows microscopic examination of skin lesions, and has already proven to be an effective tool for analysis of skin erythema [98], evaluation of wound status [99] and early detection of skin cancers [100]. Skin color is produced by a combination of complex mechanisms and utilized as vital information in dermatology to interpret the characteristics of a lesion and the depth in skin at which the lesion exists [101]. In [100], a CBMIR system has been developed as a diagnostic aid to the dermatologists for skin cancer recognition. A 64-dimensional color histogram consisting of 4 uniformly quantized bins for each color channel and two color moments (mean and variance) are extracted from the segmented skin lesion and used as specific local features for retrieval of color variation.

Endoscopic images, on the other hand, are taken by a lighted tube-like instrument with a camera which is placed into the gastrointestinal tract for viewing abnormalities such as the bleeding, growth of tumors, polypoid lesions and ulcerations [102]. A content-based endoscopic image retrieval (CBEIR) system with color clustering approach has recently been presented in [103]. In order to reduce the color sensitivity to noise and the color histogram

dimension, and considering that endoscopic images generally contain only a few dominant colors (such as red, yellow or purple) and HSV color space is most approximate to human perception, the original 24-bit RGB color images are converted into HSV color space, in which Hue, Saturation and Value components have been non-uniformly quantified into 6, 4, 8 levels, respectively. Therefore, endoscopic images' color feature can be represented by a clustered 192-bin (6×4×8) histogram. Based on some preliminary studies of the upper gastrointestinal endoscopic images, [104] proposed an endoscopic image analysis system based on a domain specified color model and the color variations in the endoscopic images of the stomach, for detecting and retrieving abnormal regions such as early gastric cancer and the inflammatory change of the surroundings like rubor, erosion, intestinal metaplasia or atrophic gastritis [104, 105]. Since the most of endoscopic images have redness on the whole due to the influence of *haemoglobin* (a predominant pigment in the gastrointestinal mucosa), the detailed colors of an endoscopic image can be determined by the amount of hemoglobin, i.e., the distribution of blood flow in the mucous membrane, using *IHb* (*Index of Hemoglobin*) [106, 107]. Such IHb can be calculated from the values of the original RGB channels (Vr, Vg, and Vb) [107]: IHB = 32{log2(Vr/Vg)}, and used as a unique color feature descriptor of the gastrointestinal endoscopic images. Similarly, [108] proposed a new color feature extraction approach for the classification of colon status from the colonoscopic images. In general, colonoscopic images contain rich color information associated with various lesions: (1) malignant tumors are normally inflated and inflamed, and the inflammation is usually reddish and more severe in color than the surrounding tissues; (2) benign tumors exhibit less intense hues; (3) redness may specify bleeding; (4) blackness may be treated as deposits due to laxatives; (5) green may be the presence of faecal materials; and (6) yellow relates to pus formation [108]. Based on the above properties, some local features are extracted from the chromatic and achromatic histograms of the image. In the histograms of each image, certain lower and upper threshold values of the regions of interest are selected for the extraction of the quantitative parameters. The color features are defined as follows:

$$\beta_C = \sum_{i=L_1}^{L_2} Hist_C(i) \Big/ \sum_{i=0}^{L-1} Hist_C(i) \tag{1}$$

where β_C is a set of the color features for various image components C={I (Intensity), R(Red), G(Green), B(Blue), H(Hue), S(Saturation)}, $Hist_C(i)$ is the histogram amplitude at level i of a particular color component C, L is the number of gray levels, and L_1 and L_2 are the lower and upper threshold values of the histogram of the region [108].

Tongue diagnosis, i.e., inspection of the tongue, is one of the most important methods in traditional Chinese medicine (TCM) where a physician visually examines the color and other properties of the substance and coating of the tongue. It has been shown to be a simple and non-invasive way to identify the body condition / state and symptoms of the patient and can be further combined with the other three major diagnosis methods (i.e., listening / smelling, inquiry, and palpation) in traditional Chinese medicine for determining the actual disease or deficiency in the patient [89, 109-111]. However, tongue diagnosis in traditional Chinese medicine is usually based on the capacity of the physician's vision for detailed discrimination.

The environmental factors such as the difference of light sources and surrounding illuminations may affect the physician in making accurate diagnosis via observation of the tongue. Therefore, it is vital to build an objective and quantitative tool / system, so called computerized tongue characterization, for acquiring *true-color tongue images* that are invariant to illumination changes, and supporting automatic analysis of tongue properties [111-113]. Various tongue diagnostic systems have been recently developed to support acquisition, processing, analysis, storage and retrieval of tongue images, and provide a useful aid in standardization and automation of tongue diagnosis [111, 114-118]. The key components of these systems normally include a color CCD (charge-coupled device) camera, standardized light sources, a semi-close box / chest with a head / face supporting mechanism, and a computer system with image processing tools and image databases. Some sample tongue images captured from the systems are shown in Figure 4.2 [111, 119, 120].

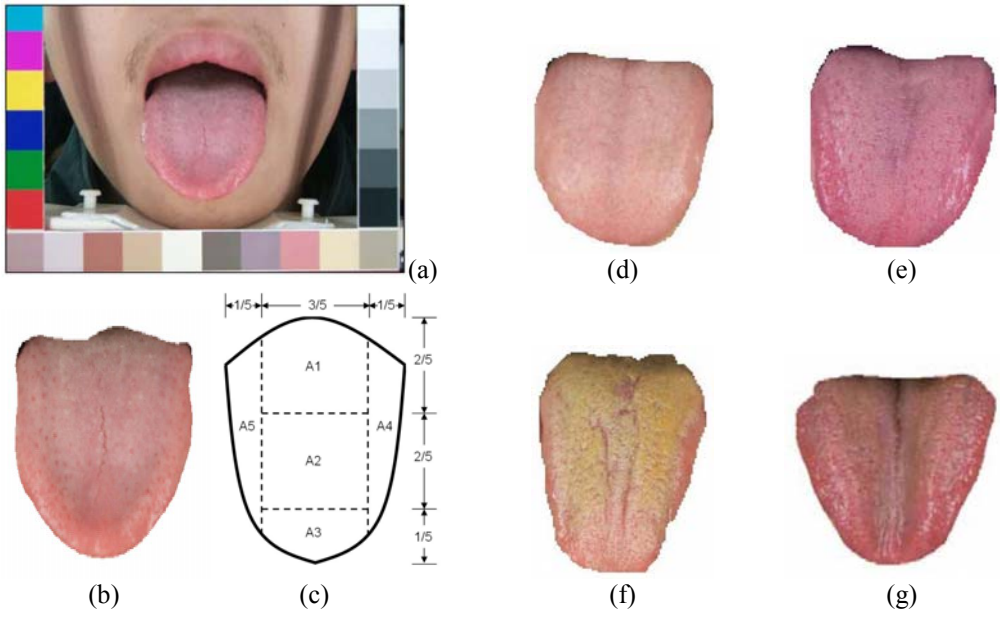

Figure 4.2 Examples of tongue images after color calibration: (a) An original tongue image with standard color bars for color calibration; (b) A segmented normal tongue body image; (c) The partition of a tongue – five sub-regions of a tongue corresponding to the health condition of different organs; (d) A light red tongue with thin white coating; (e) A dark purple tongue with thin white coating; (f) A light white tongue with yellow thick coating; (g) a deep red tongue with brown coating (Courtesy of Prof. L. S. Shen, Beijing University of Technology, China).

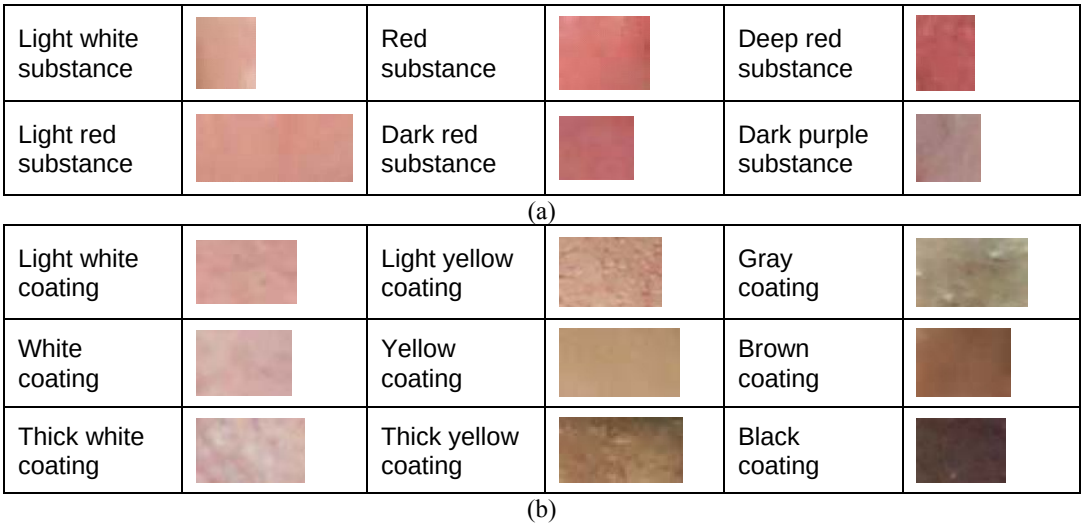

Figure 4.3 Sample images of different fifteen color categories: (a) six color classes of the tongue substance; (b) nine color classes of the tongue coating. (Courtesy of Prof. L. S. Shen, Beijing University of Technology, China).

According to the principles of traditional Chinese medicine, in general, the tongue body can be divided into five sub-regions, each of which corresponds to the health conditions of different organs of the human body, as shown in Figure 4.2(c): regions A1 (root) and A2 (middle / center) represent the health conditions of stomach and spleen; region A3 (tip) corresponds to heart and lungs; and regions A4 (right side) and A5 (left side) indicate the health information of liver and gall bladder [111, 117]. Giving the fact that there are large variations in the color of tongue substance and coating, in [111], a total of fifteen color categories were set up when running color feature extraction – the colors of the tongue substance were classified into six different classes (Figure 4.3a), while the coating colors were divided into nine different classes (Figure 4.3b) [111, 119]. Based on the above five different sub-regions and fifteen color categories, the local color features of tongue images were extracted by estimating the distribution of substance (and coating) of each color category in each sub-region of the segmented tongue body, and have been used for tongue image analysis and retrieval in a tongue image analysis instrument (TIAI) system. Moreover, in order to obtain color features that are more consistent and corresponding to human vision, the color reproduction was conducted for image capturing, transferring and displaying [121]. Some other systems for tongue image analysis and retrieval include: a tongue-computing model (TCoM) system for the diagnosis of appendicitis using a total of 22 local metrics in four color spaces (RGB, CIEYxy, CIELUV and CIELAB) with color moments (mean and variance) for different sub-regions of the tongue body [114]; a vision-based tongue diagnosis system using the local

(block-size) RGB color mean metric [115]; a tongue diagnosis supporting system based on quantized color class labels [116]; a computerized tongue examination system (CTES) using a color relaxation approach with decision boundaries for HSV color space [117]; and a digital tongue inspection system (DigiTIS) based on RGB color histogram approach [118].

For generic CBMIR, i.e., retrieving varied medical images acquired from various image modalities, including color medical images mentioned above, the color-based retrieval is vital and needs to be included in the systems. [122] simply used a different quantization of the *HSV space* with 6 Hues, 3 Saturations and 3 Values for retrieving color medical images in CasImage [123]. [43], on the other hand, proposed a new color structure descriptor called *HDS-S* based on the MPEG-7's HMMD (hue, min, max, diff) for retrieval of varied medical images from ImageCLEFmed [124] collection. In HMMD color model, the hue component is from HSV and the min and max components are taken from the maximum and minimum values in the RGB model, respectively. The diff component is the difference between min and max. Based on the MPEG-7 standard [44], the HMMD space was unevenly quantized into 184 bins in the 3-component HDS (hue, diff, sum) color space. Then a HDS-S color structure descriptor was extracted from quantized HDS histogram based on an 8×8 sliding window – each color bin contains the number of window positions for which there is at least one pixel falling into the bin under consideration. Since the proposed HDS-S can discriminate between images with the same global color distribution but differing local color structures, it could benefit the retrieval of color medical images which tend to have similar overall color but differing color structure corresponding to intrinsic image content [43].

4.2.2 Retrieval based on Texture

Most of medical images acquired and displayed in gray scale are often highly textured, and consequently, examination of medical images usually requires interpretation of tissue appearance, i.e., the local intensity variations, based on different texture properties such as smoothness, coarseness, regularity, and homogeneity [1,125]. Since texture acquires such distinguished importance, it is becoming one of the most commonly used characteristics in medical image analysis, classification and retrieval [1,6,126]. Among various texture descriptors introduced in Section 4.1.2, the *co-occurrence matrices approach* and the *Gabor filters* are mainly adopted in the CBMIR.

The well-known co-occurrence matrices approach for texture feature representation explores the gray level spatial dependencies of texture by constructing co-occurrence matrices based on different orientation and distance between image pixels and extracting meaningful statistics from the matrices as texture representations [8]. In the co-occurrence matrices approach, given a distance d at an orientation angle θ, $p_{(d,\,\theta)}(l_1, l_2)$, the (l_1, l_2) coefficient of the corresponding matrix $\mathbf{P}_{(d,\,\theta)}$, is the co-occurrence count or probability of going from a grey level l_1 to another grey level l_2 with an inter-sample spacing of d along the axis making an angle θ with the x axis. If the number of distinct gray levels in the quantized image is L, then the co-occurrence matrix

P will be of size $L \times L$. For computational efficiency, the number of gray levels can be reduced by binning, i.e., a simple procedure in which the total range of values is divided by a smaller amount – the required number of bins, thus "shrinking" the co-occurrence matrix. Different co-occurrence matrices can be constructed by mapping the gray-level probabilities based on the spatial relations of pixels at different angular directions specified by θ while scanning the image according to the distance d [15, 46, 87, 127]. On their own, these co-occurrence matrices do not provide any measure of texture that can easily be used as texture descriptors. The information in the matrices needs to be further extracted as a set of feature values. In [46], totally 14 second-order statistic quantities called *Haralick texture features* are computed out of the coefficients. In order to allow images of different sizes to be compared, before extracting these features, all co-occurrence matrices are normalized by dividing each coefficient in a matrix by the sum of all elements. The formulas for calculating these texture features are listed in Table 1.

Table 1. The 14 Haralick texture features

Feature	Equation	Feature	Equation
F_1: (Energy) Angular Second Moment (ASM)	$\sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1}[p(l_1,l_2)]^2$	F_8: Sum Entropy	$-\sum_{k=0}^{2(L-1)} p_{x+y}(k)\log[p_{x+y}(k)]$
F_2: Contrast	$\sum_{k=0}^{L-1} k^2 \sum_{\substack{l_1=0 l_2=0 \\ \lvert l_1-l_2\rvert=k}}^{L-1}\sum p(l_1,l_2)$	F_9: Entropy	$-\sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1} p(l_1,l_2)\log[p(l_1,l_2)]$
F_3: Correlation	$\dfrac{\sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1} l_1 l_2 p(l_1,l_2) - \mu_x\mu_y}{\sigma_x\sigma_y}$	F_{10}: Difference Variance	Variance of p_{x-y}
F_4: Sum of Squares (Variance)	$\sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1}(l_1-\mu_f)^2 p(l_1,l_2)$	F_{11}: Difference Entropy	$-\sum_{k=0}^{L-1} p_{x-y}(k)\log[p_{x-y}(k)]$
F_5: Inverse Difference Moment	$\sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1}\dfrac{1}{1+(l_1-l_2)^2} p(l_1,l_2)$	F_{12}: Information Measure-A of Correlation	$\dfrac{HXY - HXY_1}{\max\{HX, HY\}}$
F_6: Sum Average	$\sum_{k=0}^{2(L-1)} k p_{x+y}(k)$	F_{13}: Information Measure-B of Correlation	$(1-\exp[-2(HXY_2 - HXY)])^{1/2}$
F_7: Sum Variance	$\sum_{k=0}^{2(L-1)}(k-F_6)^2 p_{x+y}(k)$	F_{14}: Maximal Correlation Coefficient	$\sqrt{\text{second largest eigenvalue of } Q}$

Notations: μ_x, μ_y, σ_x and σ_y are the means and standard deviations of C_x and C_y, respectively;

$p_x(l_1) = \sum_{l_2=0}^{L-1} p(l_1,l_2)$; $p_y(l_2) = \sum_{l_1=0}^{L-1} p(l_1,l_2)$; $p_{x+y}(k) = \sum_{\substack{l_1=0 l_2=0 \\ l_1+l_2=k}}^{L-1}\sum p(l_1,l_2)$ where k=0,1,2,…,2(L-1);

$p_{x-y}(k) = \sum_{\substack{l_1=0 l_2=0 \\ \lvert l_1-l_2\rvert=k}}^{L-1}\sum p(l_1,l_2)$ where k=0,1,2,…,L-1; $Q(l_1,l_2) = \sum_{k=0}^{L-1}\dfrac{p(l_1,k)p(l_2,k)}{p_x(k)p_y(k)}$; $HXY = F_9$;

$HXY_1 = -\sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1} p(l_1,l_2)\log[p_x(l_1)p_y(l_2)]$; and $HXY_2 = -\sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1} p_x(l_1)p_y(l_2)\log[p_x(l_1)p_y(l_2)]$

The co-occurrence matrices approach which provides enough discrimination power for various texture appearances has been frequently used in the CBMIR [4-6]. [128] proposed a texture-based image retrieval system for normal anatomical regions presented in CT studies of the chest and abdomen by using local and global co-occurrence texture descriptors. To extract global-level features, twenty co-occurrence matrices for the pre-segmented image were firstly constructed for four different orientations (θ as horizontal 0°, vertical 90°, and two diagonals 45° and 135°) and five displacements (d=1,2,3,4,5). Then ten Haralick features were computed for each of the twenty matrices. The obtained twenty values for each feature were further averaged and recorded as a mean-based feature vector for the corresponding image. For local (pixel)-level feature extraction, the co-occurrence matrix construction was based on setting a 5×5 neighborhood window for each pixel within the segmented region without the orientation and displacement settings, i.e., only a single co-occurrence matrix for each pixel was produced rather than for each choice of θ and d. Therefore, for each co-occurrence matrix (each pixel), the same set of Haralick features was calculated. In [127], the above ten Haralick features and eleven texture descriptors extracted by run-length matrices approach [48, 125] were calculated and stored in a texture dictionary for classification and retrieval of various human organs tissues including backbone, heart, kidney, liver and spleen. In [129], various texture descriptors were computed based on gray tone co-occurrence matrices method associated with regions of interest and used as key image content descriptors in a so-called AttributeMatch retrieval system for MRI head images. It is shown that the Haralick descriptors are also effective for the characterization of intrinsic texture features in the HRCT images of lung [82], and can help on retrieval and tracking of Alzheimer's disease from MRI head images [130]. In [131], a tool for texture extraction called TextEx was developed to index and retrieve CT and MRI images based on six texture descriptors extracted by the co-occurrence matrices method, supporting tissue identification of brain, spine, heart, lung, breast, adiposity, muscle, liver and bone. Besides the applications in texture-based retrieval of CT and MRI images, the co-occurrence matrices approach has also been used in retrieving ultrasound images which contain various granular texture layouts [132, 133], mammography images with benign and malignant masses [134], tongue images with various texture layouts at different sub-regions of the tongue body [114,117,121], and dermatoscopic images with differential texture structures specific to skin lesions [100]. For texture-based retrieval of color endoscopic images, a color-co-occurrence model [135] has been set up to integrate color and texture characteristics for efficient image retrieval. The original RGB color space is first converted to HSV space and re-binned to 192 levels (H:6×S:4×V:8). Then twelve color-co-occurrence matrices with size of 192 × 192 were constructed for four different orientations (0, $\pi/4$, $\pi/2$, $3\pi/4$) and three different distances (1,3,9). Finally, two statistics (mean and standard derivation) of four Haralick descriptors (CONtrast, ENerGy, CORrelation and ENTropy) were computed from each matrix to form a textual feature vector $<\mu_{CON}, \sigma_{CON}, \mu_{ENG}, \sigma_{ENG}, \mu_{COR}, \sigma_{COR}, \mu_{ENT}, \sigma_{ENT}>$ [135]. [43] and [88] have reported the use of the co-occurrence matrices approach for retrieving large collections of medical images from ImageCLEFmed and images from World Wide Web medical image atlases, such as X-ray of the chest with enlarged heart, sagittal or frontal views of the head MRI, chest CT with micro nodules, abdominal CT with liver blood vessels,

angiogram of aorta, microscopic images of leukemia, Alzheimer's disease, bacterial meningitis, skin lesions, etc. Unlike another well-known *Tamura features* method [47] which only provides visually meaningful texture properties, the co-occurrence matrices approach allows detecting some abnormalities in medical images that are beyond human limited appreciation of complexity and otherwise difficult to determine by other texture extraction methods, and provides valuable information about the medical images that may not be visible to the human eye [127]. In the above CBMIR systems, a subset of Haralick texture features is normally selected and adopted, based on experiments, exhibited the best performance among the total 14 features – the most efficient and frequently used descriptors in the CBMIR are: *Energy*, *Entropy*, *Contrast*, *Inverse Difference Moment*, *Correlation*, and V*ariance*.

The Gabor filters have been commonly adopted as powerful edge / line / bar detectors with orientation and scale (frequency) tunable properties, and their statistics in the image or image parts (regions) are often used to characterize the underlying texture information, while achieving minimum joint 2D uncertainties in both spatial and frequency domains [52]. Since the Gabor filters allow one to choose arbitrary orientation and scale, and considering that textural images are usually distinguishable with orientation and scale features, the Gabor filters have also widely used to extract texture features from images for the CBIR [51, 52, 136]. A 2D Gabor function is a Gaussian-modulated sinusoid and can be defined as:

$$g(x,y)=\left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jW_x\right] \tag{2}$$

where σ_x and σ_y are the standard deviations of the Gaussian shape envelop along the x and y direction, respectively, and W_x is the modulation frequency of the filter. A set of self-similar Gabor filters is then generated via scaling (m) and orientation (n):

$$g_{mn}(x,y)=a^{-m}g(x',y') \tag{3}$$

where $a>1$, $x'=a^{-m}(x\cos\theta+y\sin\theta)$, $y'=a^{-m}(-x\sin\theta+y\cos\theta)$, $\theta=n\pi/K$, m =0,1,…, S-1, n =0,1,…, K-1, S and K are the total number of scales and that of orientations, respectively. Based on the obtained Gabor filters, given an image $I(x,y)$, its Gabor transform is defined to be:

$$W_{mn}(x,y)=\int I(x,y)g_{mn}^{*}(x-x_1,y-y_1)dx_1dy_1 \tag{4}$$

where * indicates the complex conjugate. Then the mean μ_{mn} and the standard deviation σ_{mn} of the magnitude $|W_{mn}|$ form a feature vector, $<\mu_{00}, \sigma_{00},\ldots, \mu_{mn}, \sigma_{mn},\ldots, \mu_{S\text{-}1\ K\text{-}1}, \sigma_{S\text{-}1\ K\text{-}1}>$, representing the texture features of the image. Here, the whole image is decomposed at S scales and K orientations by using the Gabor filters and ready for texture-based indexing and retrieval.

The applications of the Gabor filters in the CBMIR have been reported and covered in various medical image categories such as MRI cardiac images [87], liver CT images [137], histological images of ten organs [138] and gastrointestinal tract [95], mammography images [134], and very large collections of varied medical images [43, 122]. In [87], a bank of 42 Gabor filters with angular bandwidths of 30° and frequency bandwidths of one octave was used to compute the mean and standard deviation of the magnitude response of each heart

MRI image for extracting feature vectors of myocardium texture, supporting CBMIR queries such as "*given one vertical slice, find out slides corresponding to a given instant in time along one cardiac cycle* (*in particular the end of systole or the end of diastole*)". Such myocardium texture-based retrieval approach lies in the fact that the contraction of the myocardium is correlated to the fineness of its texture presented in the image: (1) the more the myocardium will be contracted, the finer its texture will be; and (2) the contraction of the myocardium corresponds to a reduction of its volume and its fibers then lie more closely [87]. A CBMIR system for liver CT images based on Gabor texture was presented in [137], allowing retrieving different types of CT findings and manifestations such as low attenuation with infiltration, lipiodol retention, and multi-focal nodular type. However, all the hepatic pathology bearing patches and other abnormal regions need to be manually delineated before Gabor feature extraction. For high-resolution histological image retrieval, since texture patches in histological images are normally not homogenous, it is difficult to directly use global Gabor features to image indexing and retrieval. Instead, the original image is usually divided into sub-images or blocks, with the hypothesis that texture patterns within a block are homogenous. The Gabor filters are then used to extract texture features for these patterns. Finally, a histological image can be represented with finite number of feature vectors each corresponding to one block [95, 138]. In [95], total fifteen semi-fine Gabor features based on 64×64 window-size were used for retrieving histological images of six gastrointestinal tract organs. For the CBMIR system presented in [138], a Gabor filter bank with eighteen filters of three scales and six orientations is calculated as texture features to index and retrieve histological images from ten categories: adrenal, heart, kidney, liver, lung, pancreas, spleen, testis, thyroid and uterus. The Gabor features have also been adopted for generic CBMIR, i.e., retrieving varied medical images acquired from various image modalities. [122] investigated the potential of using a bank of real circularly symmetric Gabor filters with three scales and four orientations for retrieving varied medical images from CasImage teaching files collections. The resultant 12 Gabor filters have been shown that give good coverage of the frequency domain and little overlap between filters. However, the performance of these Gabor features has not yet been compared with other characteristics such as those based on the co-occurrence matrices approach or wavelet filters. Another generic CBMIR study in [43] concluded that a Gabor filter bank with two scales and four orientations gave the best retrieval performance for retrieving varied medical images from ImageCLEFmed collection.

4.3 CBMIR BY GEOMETRIC SPATIAL FEATURES

4.3.1 Retrieval based on Shape

Shapes found in medical images express different characteristics for different anatomy. Some possess readily identifiable shapes, such as brain, heart, lungs, kidneys, and several bones, while others can be arbitrary, such as lesions or tissues. However, disease processes can affect

the structure of organs and cause deviation from their expected shapes. Even abnormal entities tend to demonstrate differences in shape between benign and malignant conditions [15]. Furthermore, shapes can undergo non-rigid deformation over time, over the progression of diseases, or from patient to patient. Therefore, shape information becomes one of the most important and effective criteria in characterizing many pathologies identified by medical experts, and medical image retrieval by shape features appears promising for quickly finding the same anatomic region of the same disease, and can be beneficial to support certain disease diagnoses. Many CBIR techniques based on shape features which exhibit and retain different shape characteristics have been developed, as mentioned in Section 4.1.2. Some of them seem to be well-fitted to certain CBMIR applications, while several new approaches have been proposed for particular shapes that populate the databases in some specific medical imaging studies.

The well-known *Fourier Descriptors (FDs)* represent the shape in a frequency domain with the Fourier transformation of its boundary signatures, and can be used to discriminate different shapes – the lower frequency FDs contain information about the general shape, while the higher frequency FDs hold information about smaller details of the shape [62]. The most general form of representation of a contour (2D shape) can be just a closed sequence of N successive boundary points or pixels (x_i, y_i), where i=0, 1,2,…,N-1. Three main signatures of the contour are defined as: *curvature*, *centroidal distance*, and *complex coordinate* functions. The curvature function at a point i along the contour is defined as the rate of change in tangent direction of the contour: $C_i=d\theta_i/d_i$, where θ_i is the *turning function* of the contour, i.e., $\theta_i=\arctan(y_i'/x_i')$, here $y_i'=dy_i/d_i$ and $x_i'=dx_i/d_i$. The centroidal distance function expresses the distance of the boundary points from the centroid (x_c, y_c) of the shape: $R_i=[(x_i-x_c)^2+(y_i-y_c)^2]^{1/2}$. And the complex coordinate function can be obtained by simply representing the coordinates of the boundary points as complex numbers: $Z_i=(x_i-x_c)+j(y_i-y_c)$. Through Fourier transforms of the above three functions, three sets of complex coefficients, i.e., FDs, can then be generated. In order to achieve rotation invariance (i.e., contour encoding is irrelevant to the choice of the reference point), only the amplitudes of the complex coefficients are used and the phase components are discarded. Scale invariance is achieved by dividing the amplitudes of the coefficients by the amplitudes of DC descriptor or the first non-zero frequency coefficient. And the translation invariance is obtained directly from the contour representation [13, 62]. The FDs of the curvature are:

$$FD_C = \left[\left|F_1\right|, \left|F_2\right|, ..., \left|F_{M/2}\right|\right] \tag{5}$$

and the FDs of the centroidal distance are:

$$FD_R = \left[\frac{\left|F_1\right|}{\left|F_0\right|}, \frac{\left|F_2\right|}{\left|F_0\right|}, ..., \frac{\left|F_{M/2}\right|}{\left|F_0\right|}\right] \tag{6}$$

where F_i denotes the ith component of Fourier transform coefficients. The FDs of the complex coordinate are:

$$FD_Z = \left[\frac{|F_{-(M/2-1)}|}{|F_1|}, \ldots, \frac{|F_{-1}|}{|F_1|}, \frac{|F_2|}{|F_1|}, \ldots, \frac{|F_{M/2}|}{|F_1|} \right] \quad (7)$$

where F_1 is the first non-zero frequency component used for normalizing the transform coefficients.

The study in [139] has shown that the FDs-based retrieval approach can achieve better performance for the spine X-ray images with vertebral shapes which effectively described various pathologies identified by medical experts as being consistently and reliably found in the image collection. The shapes of cervical and lumbar vertebrae were firstly segmented from the digitized spine X-ray images, and each shape was then outlined by the coarse radiologist-marked 9-point model which is widely used in the vertebra morphometry community. The 9-point shapes were further made dense to 36 points by linear interpolation. The resulted dense segmented vertebra boundary shapes can be used to extract Fourier descriptors as shape attributes. The nature of spine X-ray images with gray scale and offering very little in terms of texture for the anatomy of interest makes this an ideal application area for the CBMIR by Fourier descriptors. [82] presented a CBMIR system for lung HRCT images, in which FDs were extracted from different pathology bearing regions with seven types of diseases: centrilobular emphysema, paraseptal emphysema, sarcoid, invasive aspergillosis, broncheitasis, eosinophilic granuloma, and idiopathic pulmonary fibrosis identified by radiologists. The local shape attributes using FDs have been demonstrated to significantly improve retrieval performance in the domain of HRCT images of the lung over a purely global approach. The shape-based retrieval by FDs has also been proven to be an effective tool for pathology images and can assist physicians in differential diagnosis of lymphoproliferative disorders [140]. Various nucleus shapes segmented from original leukocyte images were characterized through Fourier descriptors, and used for efficient retrieval of different types of leukocytes (band neutrophils, lymphocytes, monocytes, and polymorphonuclear leukocytes) and three particular disorders, i.e., chronic lymphocytic leukemia (CLL), follicular center cell lymphoma (FCC), and mantle cell lymphoma (MCL) [140].

Recalling the character of FDs: lower frequency FDs describe the general shape property while higher frequency FDs reflect shape details, i.e, FDs will become larger if the object shape becomes more complex and rough. It appears that such FDs will be ideal to help in performing retrieval of mammography images and distinguishing accurately between benign masses and malignant tumours, due to the following observation: benign masses are normally round or macro-lobulated in appearance, and well-circumscribed with smooth contours; malignant tumors, on the other hand, usually possess rough, jagged or irregular boundaries, including micro-lobulations, spicules, and concavities [141, 142]. [141] investigated the use of various shape descriptors such as FDs, compactness, invariant moments, acutance measure and chord-length statistics to discriminate between benign and malignant breast tumors. Giving that some benign lesions may have a speculated appearance and round / well-defined malignant lesions do exist, an evaluation study of using the above different shape factors to distinguish between circumscribed and speculated tumors has also been conducted in [141].

The results showed that FDs method gave higher accuracy for circumscribed / speculated classification, however, less accuracy than acutance measure for benign / malignant classification. Overall acutance measure achieved best performance for breast cancer classification and retrieval. *Acutance measure*, as presented in [141, 143], is a descriptor of the sharpness or change in density across a mass margin, and can be obtained by first computing the sum of differences $d(j)$ along the normal to each boundary point j=0,1,…,N-1, where N is the number of boundary points of the region:

$$d(j) = \sum_{i=1}^{n_j} \frac{f(i) - b(i)}{2i} \tag{8}$$

here $f(i)$ and $b(i)$, are pixels along the normal inside and outside the region, respectively. Here i=1,2,…,n_j, and ,n_j is the number of pixel pairs along the normal used to calculate the differences for the jth boundary pixel, and is limited to a predefined maximum value n_{max}. The acutance measure A can then be derived by normalizing $d(j)$ over all boundary pixels:

$$A = \frac{1}{d_{\max}} \left[\frac{1}{N} \sum_{j=0}^{N-1} \frac{d^2(j)}{n_j} \right]^{1/2} \tag{9}$$

where d_{max} is a normalization factor dependent upon the maximum grey-level range and n_{max}, such that A is within the range (0, 1]. In [141], acutance measure was used as a descriptor of edge strength or diffusion of a breast tumor or mass into the surrounding regions, in which low value indicated malignant tumors while high value implied benign masses.

Unlike FDs and acutance measure, *invariant moments* can be computed from a region's boundary or silhouette: the former is more sensitive to high-frequency edge details, while the latter is less sensitive to noise and is an indicator of gross shape [141]. The 2D (p+q)th-order central moments μ_{pq} of a density distribution function $f(x,y)$ are derived as:

$$\mu_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q f(x, y) \tag{10}$$

where (x_c, y_c) is the center of the region, and the summation is over all pixels in the region boundary. The central moments can be further normalized for scale invariant: $\eta_{pq} = \mu_{pq} / \mu_{00}^r$, where γ=(p+q+2)/2. A set of seven low-order, central moments invariant to translation, rotation and scale can be obtained [144, 145]:

$$M_1 = \eta_{20} + \eta_{02}$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})\left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})\left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right]$$

$$M_6 = (\eta_{20} - \eta_{02})\left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] + 4\mu_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$
$$M_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})\left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right]$$
$$- (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})\left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] \quad (11)$$

The invariant moments have been adopted for retrieving HRCT lung images [82] and mammography images [134]. The evaluation study in [134] showed that the shape-based retrieval by the invariant moments can be used to classify breast tumors and has achieved similar performance as FDs. In [146], a shape-based retrieval system for thermal medical images has been developed by using a set of combinations of invariant moments:

$$\beta_1 = \frac{\sqrt{M_2}}{M_1},\ \beta_2 = \frac{M_3\mu_{00}}{M_1M_2},\ \beta_3 = \frac{M_4}{M_3},\ \beta_4 = \frac{\sqrt{M_5}}{M_4},\ \beta_5 = \frac{M_6}{M_1M_4},\ \beta_6 = \frac{M_7}{M_5}. \quad (12)$$

Such combinations can be found to achieve invariance not only to translation, rotation and scale but also to contrast [147]. Each thermal image is then characterized by the above six invariant moments $\Phi=[\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6]$. Image retrieval is performed by finding those images whose invariant moments are closest to the ones calculated for a given query image, and the results have shown this approach was very robust for thermal images of arms, neck, lower back, dorsal view, and legs [146, 148].

Other shape-based retrieval techniques used in the CBMIR include: retrieval of MRI heart images based on the simple *turning function* approach [149]; retrieval of angiogram and MRI images with *mass / centroid / dispersion* features [150]; retrieval of dental radiograph images using the *finite element method (FEM)* and *eigendecomposition* approach [151]; retrieval of mammogram images using *Zernike Moments* [152], and *morphological* features for fast search of tumor shapes [153]; retrieval of pathology images based on *integrated region matching (IRM)* measure [154]; retrieval of fMRI brain images using *concentric circles (CC)* features extracted from Wavelet-decomposed ROI subsets [81]; retrieval of varied medical images using MPEG-7's *Contour Shape Descriptor* [155], *Salient Point Detector* [156], *optimal Dynamic Time Warping* (*DTW*) approach [157], *convex hull model* [158] and *similarity flooding* approach [159].

4.3.2 Retrieval by 3D Volumetric Features

The majority of medical images capture human anatomy that is in nature a 3D structure. The 3D medical images can be used for non-destructive inspection of the body and its component regions *in vivo* and *in vitro*, supporting execution and monitoring of interventions, and providing quantitative measurements to determine if an abnormality is present by comparison with normal controls in diagnosis and treatment planning [1, 3, 160]. However, the feature extraction and retrieval of 3D medical images in most of the existing CBMIR systems so far are still performed based on 2D slices – simply following the way of the conversional image retrieval in 2D domain, and departing from their originally obtained 3D form. It is believed

that the more accurate CBMIR with more discriminating power can be achieved if we can take full advantage of the information available in 3D spatial domain by performing retrieval of these medical images based on their 3D volumetric features [161].

[162] proposed a CBMIR approach based on *3D iMSP* (*ideal midsagittal plane*) features for retrieving 3D CT neuroimages of hemorrhage (blood), bland infarct (stroke) and normal brains. Giving the following basic observation of neuroradiologic imaging: normal human brains exhibit an approximate bilateral symmetry which is often absent in pathological brains, a symmetry detector was firstly constructed to automatically extract an iMSP – a virtual geometric plane about which the 3D anatomical structure captured in the given brain image presents maximum bilateral symmetry [163]. The basic idea is to find where the MSP is supposed to be for a given 3D brain image, especially for pathological brains where the anatomic MSP is often distorted (shifted or bent) due to large lesions. Automatically locating and retrieving possible lesions (such as bleeds, stroke, tumors) can therefore, be conducted by detecting asymmetrical regions with respect to the extracted iMSP. After the iMSP is aligned in the middle of each 3D volumetric image, a stack of cross-sectional 2D slices was temporarily derived as basic image units, from which three types of asymmetry features were then calculated to quantify and capture the statistical distribution difference of various brain asymmetries: global statistical properties; measures of asymmetry of halved and quartered brains; and local asymmetrical region-based properties. These features were calculated from the original image unit with its iMSP aligned, the difference image of the original image and its mirror reflection with respect to iMSP, the thresholded difference image, and the original image masked by the thresholded binary image. For each image unit, a feature vector with 48 image features was constructed, including the means, the variances, and the X and Y gradients of the intensity images at different regions and under various Gaussian smoothing, scaling and thresholding [162].

The above retrieval approach, although taking advantage of the 3D iMSP characteristics, provides global query and access on entire 2D-slice images, instead of local retrieval based on segmented 3D lesions. Since the 3D medical images are usually acquired from different modalities in different hospitals, each 3D neuroradiology scan may start and end at different portions of the brain with different angles or along different axes. This requires that 3D image data sets need to be properly registered and segmented before existing CBMIR techniques can be directly used for meaningful results [164-166]. In [166], a content-based retrieval of 3D MRI brain images with the user-selected 3D anatomical structure defined as *volume of interest* (*VOI*) in a given reference image has been developed. The retrieval process was turned into a deformable registration process using global to local affine followed by free-form transformations for the best matching between the user-defined VOI and 3D images in the database and correlation was used as a measure for morphological differences. This retrieval method, however, relies on the VOI definition to be a simple local subdivision of the image, rather than the segmented region. [167] proposed a CBMIR approach for 3D fMRI images using specific *3D concentric sphere* features extracted from pathological VOIs. Since tumors or lesions in medical images are often considered to be homogeneous regions for simplicity, the VOIs which formed by the voxels with the same value were firstly segmented, followed by

a three-step feature extraction process of these VOIs. In the first step, the center of mass V of the VOI was computed. In step 2, using V as the center, a series of 1 … k concentric spheres in 3D domain was constructed with regular increments of their radius. In the last step, for each sphere, the k dimensional feature vectors $\mathbf{F}_s$ and $\mathbf{F}_r$ were constructed, representing the fraction of the sphere occupied by the VOI and the fraction of the VOI occupied by the sphere, respectively. The obtained 3D concentric sphere feature vectors actually map the entire VOI to a specific point in the k-dimensional space. To project the characterization vector to a lower dimensionality space, the well-known KL transform or the closely related Singular Value Decomposition (SVD) can be adopted. Furthermore, for a given training set and classes of VOIs, experimentation can be conducted as an optimal classification process to obtain an appropriate value for the each increment of radius of the concentric sphere [81, 167].

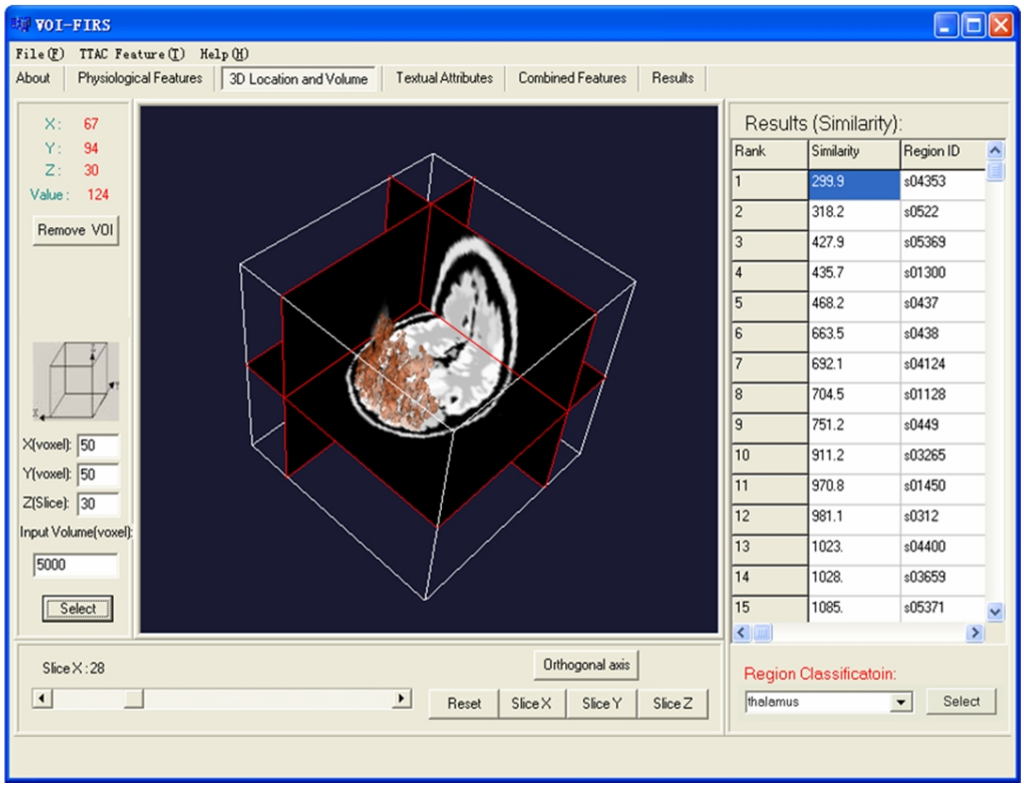

Figure 4.4 The GUI of "Query by visual features" exemplified with the rendered surface corresponding to the user-selected "prefrontal lobes" in the orthogonal Zubal slices [161]

In [161], a 3D VOI-based retrieval approach for multidimensional dynamic brain PET images was proposed and integrated into a prototype VOI-based functional image retrieval system (VOI-FIRS). For each dynamic image data set, VOIs which contain various physiological characteristics related to the local cerebral glucose consumption were firstly segmented based on a domain knowledge-based classification process, while a set of VOI functional and physiological features can be extracted (more details can be found in Section 4.5). For the extraction of VOI visual features, an anatomical standardization procedure of the 3D stereotactic surface projection transformation in the NEUROSTAT package [168-170] was adopted to deform the 3D 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. A set of transformation library files was then created for the dynamic PET images and applied to the corresponding segmentation results for warping the segmented images into the same standardized image frame of reference. After that, the centroid moment for each VOI and the corresponding volumes (i.e., the total number of voxels in the VOI) were estimated and indexed into the database, respectively. The VOI volumetric features can be retrieved by measuring the absolute differences between the two volumes. In the similarity measure of the VOI location, the digital *Zubal phantom* [171] with fully labelled phantom was transformed into the standard coordinate system as with dynamic PET images. The similarity of the VOI location was measured based on the 3D spatial distance (in voxels) between the user-defined point from the Zubal atlas and the centroids of the VOIs indexed in the database.

The ability for the physician to express their search needs and to navigate their search results accurately and easily is crucial in CBMIR systems [172], especially for retrieval of multi-dimensional medical image by volumetric features. In [161], the VOI-FIRS was responsive to the user interaction and in the retrieval process. As one of the most important interface functionalities of query components, the “Query by VOI visual features” GUI (see Figure 4.4) can support user-selection of the VOI location from the transformed Zubal phantom which used as the standard brain atlas (center-window). The user allows to navigate in the 3D viewing space (rotation, scaling, and translation) and to change the viewing planes of saggital, coronal, and transaxial slices, providing both the conventional and 3D orthogonal views. The location feature can be selected either by selecting a point on the Zubal phantom in the center display zone, setting voxel coordinates numerically via input panel at the left-middle, or on the labelled Zubal structure at the bottom-right. The volume of the VOI can also be set numerically, which works independently to the location feature. The GUI of the output folder / window for retrieved result displaying is illustrated in Figure 4.5. Each thumbnail result in the left part (Figure 4.5a - d) is presented as an *active display zone*, instead of a still index image, and can be individually navigated in orthogonal planes. The similarity indices from the retrieved VOIs are listed below the images. Any of the retrieved images can be enlarged and the VOI surface rendered, as shown in the right display zone in Figure 4.5e. 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 dynamic functional behaviour and size apparent in a malignant brain tumor case. The query was formulated to include a kinetic curve as a

functional feature derived from an existing tumor VOI in the database, and 2000 voxels as the volume parameter, with equal weights applied to the two selected features. The highest ranked result shown in Figure 4.5a and enlarged in Figure 4.5e 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 behaviour and user-defined volume, which is in agreement with the similarity indices for the query features shown below the images [161].

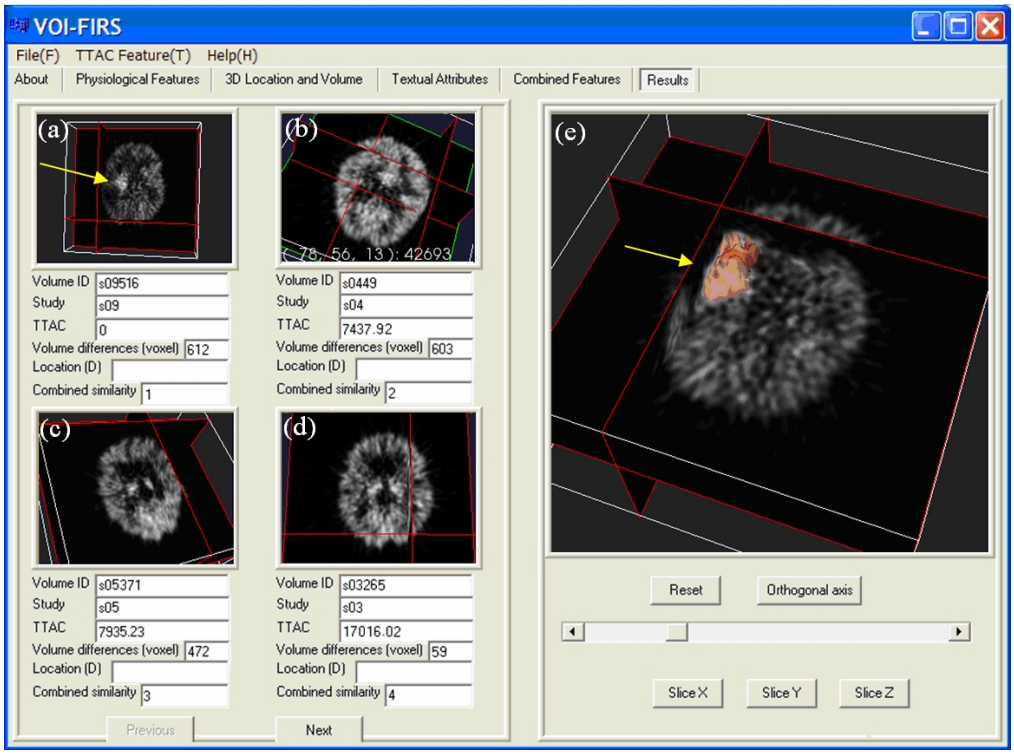

Figure 4.5 The GUI of the query results in the VOI-FIRS [161]

4.3.3 Retrieval by Spatial Relationships

Spatial relationships are an important piece of medical knowledge since a physician's mental model of the patient includes understanding not only the shape, size and boundaries of organs or lesions but also their spatial extents with other adjacent structures in the human body. Medicine thus critically relies on knowing where body structures are located and their locations relative to other structures [173, 174]. Giving that the abnormalities are defined as

gross deviations from the anatomical models, the spatial relationships of body structures are often critical to the diagnosis, prognosis, and mechanics of human disease. For example, spatial content in terms of relationships in surgical or radiation therapy of brain tumors is very decisive because the location and related adjacent structures of a tumor have profound implications on a therapeutic decision [175]. Low-level features cannot always capture or describe these complex scenarios. Various spatial relationships have therefore, been modelled and used as content features for the CBMIR to support complex image queries like "*find out all image cases demonstrating the invasion of an adenoma into the sphenoid sinus*" [173] or "*retrieving all images having dangerous tumor on left lung above and approximately touching another tumor*" [176], and assist the physician to understand and integrate the complex relationships between patient symptoms, diagnostic image features, and underlying disease pathology [174, 176-179].

In [177], a CBMIR system for MRI and CT images has been developed based on Attributed Relational Graphs (ARGs) representations holding features of objects (regions) and relationships between them. In an ARG descriptor, the objects are represented by graph nodes, and the relationships between objects are constructed by arcs between such nodes. Both nodes and arcs are labelled by the attribute values of the objects features and the relationship properties, respectively. The individual objects are described by four attributes: *Size* (*s*) – computed as the size of the area it occupies; *Perimeter* (*p*) – calculated as the length of its bounding contour; *Roundness* (*r*) – computed as the ratio of the smallest to the largest second moment; and *Orientation* (*o*) – defined as the angle between the horizontal direction and the aixis of elongation which is the axis of least second moment. The spatial relationships between objects are described by three properties: *Relative distance* (*rd*) – calculated as the minimum distance between their surrounding contours; *Relative orientation* (*ro*) – defined as the angle with the horizontal direction of the line connecting the centers of mass of the objects; and *Relative position* (*rp*) – taking values (-1, 0, 1) corresponding to objects which are the one inside the other (-1), outside each other (0) or, the second inside the first one (1) respectively. The images were firstly pre-segmented to obtain polygonal approximations of object contours, followed by shape corrections or deleting insignificant segments by experts. The segmented objects were further classified into predefined classes corresponding to normal or abnormal anatomical structures such as ventricle, hematoma, and tumor. The individual objects were then represented by 5-dimensional feature vectors of the normalized form $<s, p, r, 1+\cos(o), 1+\sin(o)>$, while relative orientations were defined by $<rd, 1+rp, 1+\cos(o), 1+\sin(o)>$ [177].

The spatial relationship models can be also used to retrieve microscopic tissue images [174]. Conventional region or image level search algorithms always assume that the regions or images consist of uniform pixel feature distributions. However, complex tissue images normally contain many pixels and regions that have different feature characteristics. Two images with similar regions may have very different interpretations if the regions have different spatial arrangements [1, 174]. Therefore a visual grammar of spatial relationships may help on describing these scenarios. In [174], the spatial relationships between all region pairs in a tissue image were represented by a $n \times n$ region relationship matrix:

$$\mathbf{R} = \left\{ \left\{ r_{ij}, d_{ij}, \theta_{ij} \right\} | i, j = 1,...,n, \forall i \neq j \right\} \quad (13)$$

where $r_{ij}=\pi_{ij}/\pi_i$ is the ratio of the common perimeter to the perimeter of the first region, π_i and π_j are the perimeters of the first and second region, respectively, and π_{ij} is the common perimeter between two regions; $d_{ij}=\|v_i - v_j\|$ is the distance between the centroids, v_i and v_j are the centroids of the first region and second region, respectively; θ_{ij} is the angle between the horizontal (column) axis and the line joining the centroids. Here each region pair was assigned a degree of their spatial relationship using fuzzy class membership functions Ω_c with $c \in$ {DIS, BOR, INV, SUR, NEAR, FAR, RIGHT, LEFT, ABOVE, BELOW}. Then the degree of membership of regions i and j to class c can be defined as $\Omega_c(r_{ij}, d_{ij}, \theta_{ij})$. The above class functions were divided into three relationship groups: (1) perimeter-class – DISjoined, BORdering, INVaded_by, and SURrounded_by; (2) distance-class – NEAR and FAR; and (3) orientation-class – RIGHT, LEFT, ABOVE, and BELOW, since multiple relationships may be used to represent a region pair, for example, BORdering from ABOVE, or INVaded_by from LEFT. Based on these second-order region relationships, the higher-order region relationships were constructed for more complex combined relationship representation.

Since the medical image content is very rich in terms of properties, characteristics, salient objects and spatial relationships which heavily related to medical knowledge, it is of great interest to retrieve medical image by combining spatial relationships with high-level medical domain knowledge. In [178], a knowledge-based approach to retrieve medical images by feature and content with spatial and temporal constructs was proposed with a four-layered semantic image model representing the spatial, temporal and evolutionary nature of medical objects. [176] developed a Medical Image Management System (MIMS) based on a global description of medical image to achieve an efficient retrieval process which integrated a high level of precision (especially in terms of spatial relations) required by the medical domain. A spatial knowledge based (SKM) has been integrated into the MIMS to provide coherent and effective objectivity of interpretation at different facets (or views) of medical images. Notwithstanding the spatial relationships approaches are included in above two studies, considering the overall works more focus on the integration of high-level medical domain knowledge in the CBMIR, more details of these two studies will be given in Section 4.4.2.

4.4 CBMIR BY COMBINATION OF SEMANTIC AND VISUAL FEATURES

4.4.1 Retrieval by Semantic Pathology Interpretation

The ultimate goal for the CBMIR is to find medically meaningful similar cases. However these cases may or may not present medical images that are similar in the usual visual sense. Sometimes similar visual features on images may not imply similar diagnosis or symptoms, and vice versa. It thus appears that combining visual features in the image with domain knowledge to reach the right subset of relevant cases in the database is a key to the success of the CBMIR.

Sometimes medical images derived from a specific organ are similar visually and normally differ only in small details which could be disregarded by untrained eyes but such domain-specific subtle differences may be of pathological significance [94]. One such domain is medical radiology, for which the clinically useful information consists of gray level variations in highly localized regions of the image [180]. For example, for HRCT images of the lung, the number of pathology bearing pixels as a fraction of all the image pixels is so small that it is not possible to use global signatures for image characterization. Moreover, the unclear boundaries between the pathology bearing pixels and the surrounding healthy part are difficult to discern and such *pathology bearing regions* (*PBRs*) are unlikely to be automatically extracted. Since medical images frequently give rise to ambiguity in interpretation and in diagnosis, current retrieval techniques using the primitive image characteristics such as color, texture, and shape, are likely to be insufficient for some domain-specific CBMIR systems. One approach to remedy such limitation is to associate high-level semantic information with low-level visual image data, albeit such approach is still controversial in terms of subjectiveness [6].

In [180], a *physician-in-the-loop* approach for content-based retrieval of HRCT lung images was developed and integrated into a CBMIR system called ASSERT (Automatic Search and Selection Engine with Retrieval Tools). To archive an image into the database, a physician spent a few seconds to delineate the PBRs and any relevant anatomical landmarks. The information regarding the pathology of the lung may reside as much in the location of each PBR with respect to the anatomical markers as it does in the characteristics of the PBR. Four major perceptual categories used by the physician for detecting pathology in the image include: (1) linear & reticular opacities; (2) nodular opacities; (3) diffuse regions of high attenuation (high-density areas); and (4) diffuse regions of low attenuation (low-density areas). After the extraction of PBRs, a lung region extraction algorithm was used to determine the boundary of the lungs. For feature extraction of the image, the system computed attribute vectors that characterize the PBRs individually and the part of the image that consists of just the entire lung region. Each PBR was characterized by a set $\mathbf{A}_{PBR}$ of commonly used physical and spatial features such as texture, shape, and other gray-level attributes. However, more importantly, the PBR also included another set $\mathbf{B}_{PBR}$ of attributes that measure the perceptual categories used by physicians for identifying and interpreting pathology in HRCT lung images. For $\mathbf{A}_{PBR}$, a large number of general-purpose attributes was calculated in total 255. While this did provide the user an exhaustive characterization of a PBR, in light of efficient indexing and searching, only a small subset of these attributes can be used for database indexing and retrieval. Therefore, a greedy search algorithm called SFS (sequential forward selection) was applied to reduce the dimensionality of the feature space while retaining the ability to accurately classify each image as belonging to its associated disease pattern, resulting only 12 general features in $\mathbf{A}_{PBR}$, such as the area, contrast, entropy, edginess histogram of the PBR. For $\mathbf{B}_{PBR}$, 14 pathology attributes for measuring four major perceptual categories, such as shape of bronchial wall, curvature of adjacent fissure, and adjacent artery size, were computed. Finally, a 26-dimentional mixture pathology feature vector $\mathbf{F}_{PBR} = < \mathbf{A}_{PBR}, \mathbf{B}_{PBR} >$ can be formed as a PBR feature descriptor that was maximally discriminatory with regard to the different diseases [180]. In the above PBR pathology interpretation approach, the physician is an integral part of

the whole CBMIR system, in the sense that it is the physician who delineates the PBRs with semantic pathology interpretation. Taking into account the physician's perceptual categories which correlate strongly with the various lung diseases, [181] proposed a better alternative to the above method. It extracted only those general features that measured the presence or the absence of the various perceptual categories that the physicians use for disease diagnosis. That is, general features were firstly used to describe the perceptual categories and discriminate between the different perceptual categories. After determining what perceptual categories were present in a PBR, the user could determine the disease of the PBR. An advantage of such hierarchical approach is that it is now easier to decide what features are needed to characterize the PBRs [181]. To support semantic integration and knowledge exchange in medical radiology domain, a CBMIR framework called "evolutionary system for semantic exchange of information in collaborative environments" (Essence) was developed and reported in [182], extracting and managing visual content of lung pathologies. In Essence, an XML-based shared ontology was developed based on the common knowledge from expert radiologists and information from two well-known references [183, 184]. Physicians were able to build their personalized semantic search criteria by customizing the degrees of satisfaction of features to existing semantic terms, and by adding new semantic terms to existing perceptual categories. The system also supported refining the shared ontology by adapting the assignment of semantic terms to image features based on individuals' preferences [182].

[94] presented an intelligent I-Browse system which combined low-level image processing technology with high-level semantic interpretation for retrieving histological images of gastrointestinal (GI) tract – a range of histological images originating from six organs along the GI tract, i.e., esophagus, stomach, small intestine (small bowel), large intestine (large bowel), appendix, and anus. Before semantic features extraction, two sets of relevant histological features (also called semantic labels) in GI tract images were firstly defined by consulting with histopathologists: (1) The 15 coarse feature labels represented by different upper-case letters – such as Lumen (L), Mucosa (M), Submucosa (S), Muscularis Externa (E), and Serosa / Adventitia (A), providing an overall structural description of the image content that is important for the later reasoning procedures; and (2) The 63 fine feature labels indicated by different numbers (1, 2, …, 63) – such as Adipose-tissue #1, Lumen #33, Stomach - junction of lumen and foveolae #63, for distinguishing different visual appearances within each coarse region. To enable the automatic analysis and interpretation, the original microscopy image was firstly partitioned as a set of subimage, each of which has 64×64 pixels inside. For each subimage, via the *Visual Feature Detector* (VFD), a set of coarse features (normalized color and gray-level histograms) was extracted and passed to a 3-layer MLP neural network for classifying the subimage into one of the 15 coarse feature classes, i.e., assigning a coarse feature label to the subimage. Using the VFD, a set of semi-fine features (means, standard deviations of the gray and color level, and Gabor filters) was also computed and used to classify the subimage into one of the 63 fine feature classes using a Bayes minimum risk classifier via marking a fine feature label to the subimage. Since each fine class actually corresponds to a coarse class, another coarse label result can be obtained for the subimage from the above semi-fine features whose classified fine class result can be matched

into one of the coarse labels. Therefore, each subimage was with two letters and one number, representing two coarse feature labels and one fine feature label, respectively, and the whole microscopy image can be presented with three matrices (the label map). With these label matrices, the *Semantic Analyzer* (SA) went through an iteration process to refine and correct the fine histological feature label of each subimage according to the histological context in the *Knowledge Base* (KB) and may produce a set of hypotheses on the labels associated with subimages by the *Hypothesis Generator* (HG), if those labels were deemed erroneously detected by the extracted coarse and semi-fine features. Based on the hypotheses, a number of fine features were invoked to extract and confirm the visual features within the suspected regions. Such analysis-and-detection cycle iterated until the semantic analyzer found a coherent result and no further change was needed. The refined final label map was then used to construct the semantic content representation structure called *Papillon* – a codename used in I-Browse system, and can be used to automatically generate the textual annotation for the image in the database with the *Annotation Generator* (AG). When the query is submitted by free text (natural language), the *Free Text Analyzer* (FTA) will extract the information in the query and convert it into the Papillon. The Papillon actually bridges information from different media (image and text), linking together the SA, AG, and FTA components in the system. Therefore, when the query is issued by sample image or free text, their semantic content in the Papillon can be used for the retrieval [94, 95, 185-189]. Some other semantic-description-based CBMIR systems include: a *property concept frame* (*PCF*) based approach for retrieval of histopathology images [190] and a brain CT image retrieval system based on the *hierarchical medical image description model* [191].

4.4.2 Retrieval based on Generic Models

In light of the heterogeneous forms presented in medical images, most of the CBMIR systems are task-specific, i.e., limited to a particular modality, organ, or diagnostic study and, hence, usually not directly transferable to other medical applications [4, 6, 192]. Medical knowledge arises from anatomic and physiologic information, requiring regional features to support diagnostic queries. However, interpretation of medical images depends on both image and query context. Since the context of queries is unknown when images are entered into the database, especially the number and kind of image features could be subject continuous evolution, the CBMIR scheme must be generic and flexible [4, 192, 193].

To conduct efficient medical image retrieval, there is a need for building comprehensive data models capturing the structured abstracts about images, and supporting more sophisticated query predicates based not only on primitive image characteristics, but also on generic semantic features with the inclusion of knowledge. [178] developed a CBMIR system called KMeD, in which a knowledge-based semantic temporal image model was included and contained four layers: (1) the Raw Data Layer (RDL) with image collection; (2) the Feature and Content Layer (FCL) with different features extracted from the image content, including

shape, spatial relationship characteristics, and temporal features; (3) the Schema Layer (SL) represents entities and relationships (spatial, temporal and evolutionary) among objects based on the features in the previous layer; and (4) the Knowledge Layer (KL) contains hierarchical structures called *Type Abstraction Hierarchies* (TAH) for classifying shape and spatial relationship features. The FCL contains: shape features such as type, area, volume, diameter, length, and circumference; spatial relations features between a pair of object including orthogonal relations (i.e., East, South, SouthEast, etc.) and containment relations (i.e., Invades, Contains, etc.). In SL, the image objects can be represented as visual entities with textual attributes and visual attributes, while multiple versions of an object over a period of time (for example, the stages undergone by a tumor during the cancer process of a particular patient) can be linked to form a stream entity for that time period. Moreover, the evolutionary object constructs for evolution, fusion, and fission, while temporal relation object constructs the temporal relationships between peer objects and between an object and its super- or aggregated type. The TAH in the KL has been designed as hierarchical structures so that at higher levels of abstraction, more generalized concepts are specified (i.e., a wider range of feature values are used), and at lower levels of abstraction, more specific concepts are described (i.e., a narrower range of feature values) [173, 178, 194]. In KMeD system, the knowledge-based temporal, evolutionary, and spatial features extracted from the images were classified and captured in the image data model and stored in tables. For example, the query "*Retrieve all image cases demonstrating a pituitary gland microadenoma which evolved into a macroadenoma with suprasellar extension pressing against the optic chiasm*" can be given by searching a pituitary gland – microadenoma containment relationship table, then following the evolutionary path that leads to a macroadenoma and selecting the instances from the outside contact relationship table between a macroadenoma and the optic chiasm [173].

Existing CBIR systems hardly address all medical image properties since medical images have different acquisition parameters, modalities, and specific noise characteristics for each imaging system. As such, the development of global features that can represent an entire medical image database seems to be practically infeasible. Aiming to provide efficient retrieval of generic medical images with coherent and effective objectivity of interpretation at different facets (or views) of medical images, [195] proposed a global description of medical images, in which a hyper-spaced image data model was constructed [175, 196]. The data model was structured as a multi-spaced form in which each space contains a set of features (contextual, physical, spatial, and semantic), and considering *the medical image as a composition of contextual and content feature spaces*. The contextual space collects the general data attached to the image without taking its visual content into account: (1) *The independent context* (e.g., the medical specialty, the patient name, the acquisition date, etc.) – since this part has no impact on the image description and due to patient privacy and other legal constraints, it needs careful treatment and can be managed separately; (2) *The pseudo-independent context* (e.g., the patient's age, gender, the image quality, etc.) – this part is vital for the CBMIR since it contains very important background knowledge and may help on determining methods to be used to construct and compare image content features. For example, the age of the patient is a determinant factor when considering organ shapes; and (3) *The*

dependant context (e.g., the image type, the incidence – sagittal / coronal / axial / others, the scene described as a triplet <title, organ, alteration>, diagnostic report, voice report, etc.) – it can significantly help on the image description. For instance, with the help from natural language processing or voice segmentation, the diagnostic report can be used to clarify missing factors in some situations like "*describing a lung x-ray of a person remains incomplete and insignificant if we ignore that he smokes*". The content space, on the other hand, provides a global image description and can be used for various query types. In general, a medical image is considered to be composed of a set of *salient image objects* in three different forms: (1) *Anatomic Organ* (AO) – presents the medical organs found in the image such as the brain, lungs, hand, etc. It gathers a set of medical regions, and is also called *Organ of Interest* (OOI); (2) *Medical Region* (MR) – describes the internal structure of the AO such as the left ventricle and the right lobe. It allows one to locate any anomaly and is also referred as *Region of Interest* (ROI); and (3) *Medical Sign* (MS) – concerns either medical anomalies (such as tumor, fracture and lesion) identified and detected by physicians, or unidentified (or variable) objects found in the image. Sometimes it is referred as *pathology bearing region* (PBR). Each salient object (AO, MR, or MS) is projected on the following subspaces: (1) *The physical subspace* – contains low-level physical properties of the image content such as various global or local color and texture features which can be extracted manually, semi-automatically, or automatically depending on the contextual space (image type, format, quality, etc.), and may be used later to analyze other subspaces. Moreover, the physical analysis can be achieved based on the pseudo-independent and dependent contexts. For example, the patient's age is a determinant factor when considering the medical organ shape. The type of the image determines the appropriate color extraction approaches; (2) *The spatial subspace* – holds middle-level geometric features of salient objects such as the shape and spatial relationships features; (3) *The semantic subspace* – concerns high-level semantic properties of salient objects. The objective of the semantic subspace is to integrate high-level features of objects and relations judged primordial by medical users for image description. However, such semantic feature analysis may require human intervention since explicit semantic objects must be recognized. The semantic subspace is usually described manually by the user due to the fact that the medical domain is very complex and each term may have several meanings depending on the context. Medical signs can be codified by some existing, albeit controversial, labeling codes for disease classification such as ICD-10 (International Classification of Diseases 10th Revision) [197] or Unified Medical Language System (UMLS) [198]. The above hyper-spaced image data model has been integrated into a prototype called MIMS (Medical Image Management System) [176, 196, 199-201].

[192] presented a general structure for content-based image retrieval in medical applications (IRMA) based on a generic multi-step approach including: categorization of the entire image; registration with respect to prototypes; extraction and query-dependent selection of local features; hierarchical blob / object representation; and image retrieval. To cope with the complex medical knowledge, the IRMA split the whole retrieval process into seven consecutive steps, while each step represents a higher level of image abstraction, reflecting an increasing level of image content understanding: (1) *Image Categorization* (based on global

features) – it can determine the imaging modality and its body orientation as well as the examined body region and biological system for each image entry with a detailed hierarchical coding scheme [202] to supplement the existing standard like DICOM; (2) *Image Registration* (in geometry and contrast) – diagnostic inferences derived from medical images are normally deduced from an incomplete but continuously evolving model of normality [4]. Therefore, the registration was based on prototype images defined for each category by experts with medical prior knowledge or by statistical analysis, in which the prototypes can be used for determination of parameters for rotation, translation, scaling, and contrast adjustment; (3) *Feature Extraction* (using local features) – it derived various local image descriptions with either category-free or category-specific approach. Like the global features for categorization, the number of local feature images is extensible; (4) *Feature Selection* (category and query dependent) – the separation of feature selection from feature extraction enables the former task to be retrieval-dependent. It can integrate both the image category and query context into the abstraction process with pre-computed set of adequate features. For instance, the retrieval of radiographs with respect to bone fractures or tumors can be conducted using a shape-based or texture-based feature set, respectively; (5) *Indexing* (multi-scale blob-representation) – it provided an abstraction of the previously generated and selected image features, resulting in a compact image description via clustering of similar image parts into regions described by invariant moments as "blobs". Thereafter, the blob-representation of the image is adjusted with respect to the parameters determined in the registration step, yielding a multi-scale "blob-tree"; (6) *Identification* (incorporating prior knowledge) – it provided linking of medical a-priori knowledge to certain blobs generated during the indexing step. Therefore it is the fundamental basis to introduce high-level image understanding by analyzing regional or temporal relationships between the blobs; and (7) *Retrieval* (on abstract blob-level) – it was performed by searches in the hierarchical blob-structures. This retrieval step requires online computations while all other steps can be performed automatically in batch mode at image entry time (offline computation). The above multi-step approach has been applied to the IRMA database of radiographs (consisting of medical images of six major body regions taken from daily routine), narrowing the gap between the semantic imprint of images and any alphanumerical description that is always incomplete [192, 193].

Some other CBMIR systems that provide varied medical images retrieval include I^2C (Image Indexing by Content) [203], COBRA (Content-Based Retrieval Architecture) [204], ImageEngine [205], and MedGIFT [122, 206, 207]. In particularly, the MedGIFT with the integration of the GNU Image Finding tool (GIFT) [208], the Multimedia Retrieval Markup Language (MRML) [209], and CasImage provides an open source framework of reusable components for a variety of CBMIR systems to foster resource sharing and avoid costly redevelopment.

4.5 CBMIR BY PHYSIOLOGICAL FUNCTIONAL FEATURES

The CBMIR techniques introduced in this chapter so far are mainly designed for anatomical images which capture human anatomy at different levels and primarily provide structural information. Unlike those anatomical images, functional / molecular images such as PET and SPECT allow the *in vivo* study of physiological and biochemical processes, providing functional information previously not available – that is what distinguishes medical images most from other types of general images [86, 210, 211]. Physiological function can be estimated at the molecular level by observing the behaviour of a small quantity of an administered substance 'tagged' with radioactive atoms. Images are formed by the external detection of gamma rays emitted from the patient when the radioactive atoms decay. Glucose metabolism, oxygen utilisation, and blood flow in the brain and heart can be measured with compounds labelled with carbon (^{11}C), fluorine (^{18}F), nitrogen (^{13}N), and oxygen (^{15}O), which are the major elemental constituents of the body. Existing CBMIR approaches may not be optimal when applied to functional images due to their unique characteristics with regard to the inherent knowledge of the disease state as it affects the physiological and biochemical processes before the morphological change of the body. Such quantitative physiological information inside the functional image content is unlikely to be retrieved by common image retrieval techniques using color, texture, and shape features. Color is not captured in the imaging process and functional images are usually acquired and displayed in grayscale, or pseudo-color. Therefore, the color feature is unlikely to be applicable to functional images. Texture is likely to be confounded by the statistical noise in functional images. Shape is also unlikely to be relevant to function. In deed, function is likely to result in changes in apparent shape during acquisition as the tracer redistributes. It appears that the development of CBMIR for functional images should take into account the specific physiological functional features [161, 212].

An early study on content-based retrieval of dynamic PET functional images has been reported in [212]. Based on this work, [161] recently developed a new VOI-based retrieval system for multidimensional dynamic functional [^{18}F]*2-fluoro-deoxy-glucose* (FDG) brain PET images which are widely used to determine the local cerebral metabolic rate of glucose (LCMRGlc) and to depict the glucose consumption and energy requirements of various structural and functional components in human brain. In dynamic functional imaging studies, the prior knowledge has the form of tracer kinetic model to a time series of PET tracer uptake measurements. Such functional information can be defined in terms of a mathematical model $\mu(t|p)$ (where t=1,2,…, T are discrete sampling times of the uptake measurements while the number of conventional scan time interval T is 22, and p is a set of the model parameters), whose parameters describe the delivery, transport, and biochemical transformation of the tracer. The input function for the model is the plasma time activity curve (PTAC) obtained from serial blood samples. Reconstructed PET images provide the tissue time activity curve (TTAC), or the output function, denoted by $\mathbf{f}(t)$ for every voxel in the image. Application of the model on a voxel-by-voxel basis to measured PTAC and TTAC data using certain rapid

parameter estimation algorithms [213, 214] yields physiological parametric images. In [161], a four-dimensional fuzzy c-means cluster analysis [215, 216] was used to construct VOI functional groups consisting of voxels that have similar kinetic behaviors. The physiological TTACs were firstly extracted for each of the N nonzero voxels in the image to form the kinetic feature vector comprising the voxel values at the dynamic time sequence of tracer uptake measurements. After applying the optimal image sampling schedule (OISS) technique [217, 218], for the dynamic FDG brain PET image study based on the five-parameter FDG model, the dimension of TTAC vectors were reduced from 22 to 5, while increasing the signal-to-noise ratio of the individual image frames for better clustering output. The fuzzy c-means cluster analysis was then applied to assign each of the N feature vectors to a set number C of distinct cluster groups and minimized the objective function J:

$$J = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{P} D_{ij}^{2} \tag{13}$$

where P ($1 \le P \le \infty$) 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 ith feature vector in the cluster j. The similarity measure between the ith feature vector $\mathbf{f}_i(t)$ and the cluster centroid $\bar{\mathbf{f}}_{c_j}(t)$ of the jth cluster group c_j was computed using the Euclidean distance:

$$D_{ij} = \left[\sum_{t=1}^{T} s(t) \left(\mathbf{f}_i(t) - \bar{\mathbf{f}}_{c_j}(t) \right)^2 \right]^{1/2} \tag{14}$$

where $s(t)$ is a scale factor of time point 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 frames with longer scan time durations which contain more reliable data. The minimization of J was achieved by iteratively updating the u_{ij}

$$u_{ij} = \left[\sum_{k=1}^{C} \left[\frac{D_{ij}}{D_{ik}} \right]^{\frac{2}{P-1}} \right]^{-1} \tag{15}$$

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

$$\bar{\mathbf{f}}_{c_j}(t) = \frac{\sum_{i=1}^{N} u_{ij}^{P} \mathbf{f}_i(t)}{\sum_{i=1}^{N} u_{ij}^{P}} \tag{16}$$

Therefore a probabilistic weighting was assigned to every voxel i representing it to be likely a member of each cluster j. For any voxel, the sum of the assigned membership degrees was 1.0. The procedure was terminated when the convergence inequality

$$\max_{ij} \left\{ \left| u_{ij}^{m+1} - u_{ij}^{m} \right| \right\} < \varepsilon \tag{17}$$

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. From the derived clustered results, the region-growing algorithm [145] was applied to the voxels in each cluster to construct the VOIs for grouping the voxels that were spatially connected and separating the different structures that may have been classified into a cluster due to the similarity of voxel's kinetic behaviour. The TTAC feature vectors extracted from the VOIs were indexed as physiological functional features and used as a key query method in the proposed VOI-based functional image retrieval system (VOI-FIRS) [161].

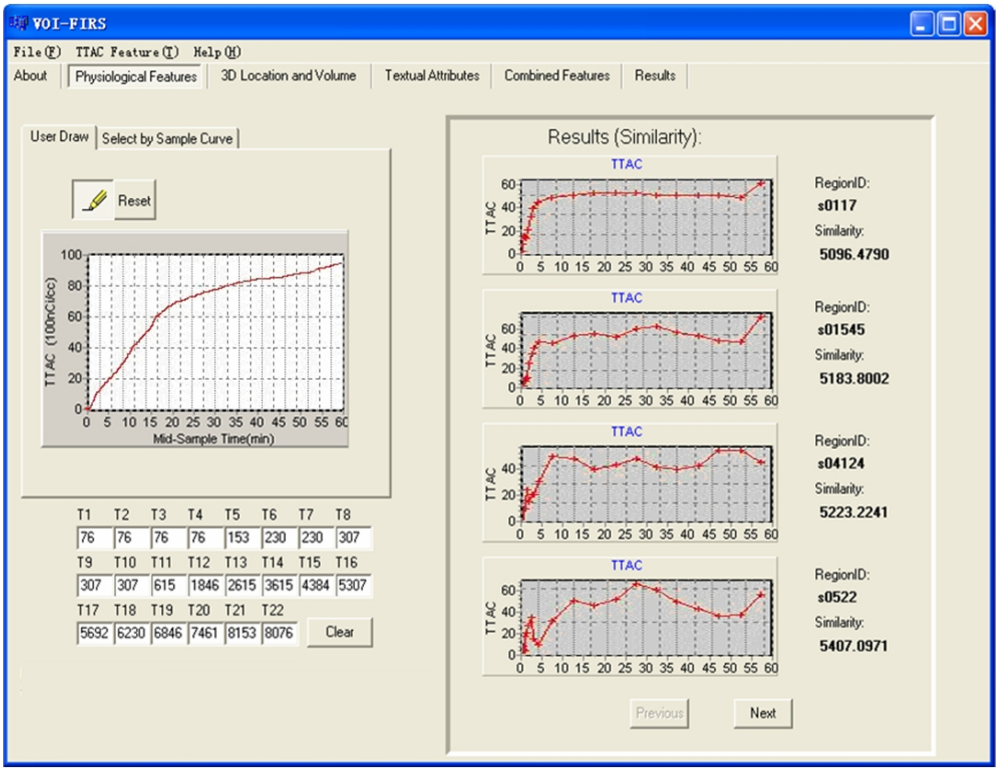

Figure 4.6 The GUI of "Query by physiological features" shows the user-drawn TTAC curve (left panel) and the retrieved VOIs with their TTAC curves and similarity indices (right panel) [161]

Figure 4.6 shows the GUI of the query component "Query by functional and physiologic features" in VOI-FIRS. The user allows to manually draw the TTAC feature curve with the labelled grid, or to select from a list of predefined sample TTACs if needed. Once the selection has been made, the TTAC curve can be manually adjusted for individual TTAC sampling

points. As the TTAC curve is concentrated in the early temporal frames, the drawn curve can be zoomed for closer inspection. An example of dynamic image retrieval based on physiological functional feature is illustrated in Figure 4.7. The sample TTAC, which approximates a pattern found in gray matter of dynamic brain FDG-PET images (as shown in Figure 4.7a), and the 3D location of the "right thalamus" selected from the labelled structures in the Zubal phantom panel (Figure 4.7b) were set as the query features. Weighting was set to 50% for the functional feature and 50% for the 3D volumetric location feature (see Section 4.3.2). The highest ranked retrieved VOI is shown in Figure 4.7c, where the query identified a VOI representing the right thalamus. Figure 4.7d presents the top ranked result from changing the location feature to "left thalamus". The result demonstrated that retrieving based on the combination of functional features and the spatial properties of the dynamic PET images in the 3D volumetric location feature, VOIs with the user-defined kinetic TTAC characteristics can be successfully identified, which may have not been possible from the functional feature alone [145, 196].

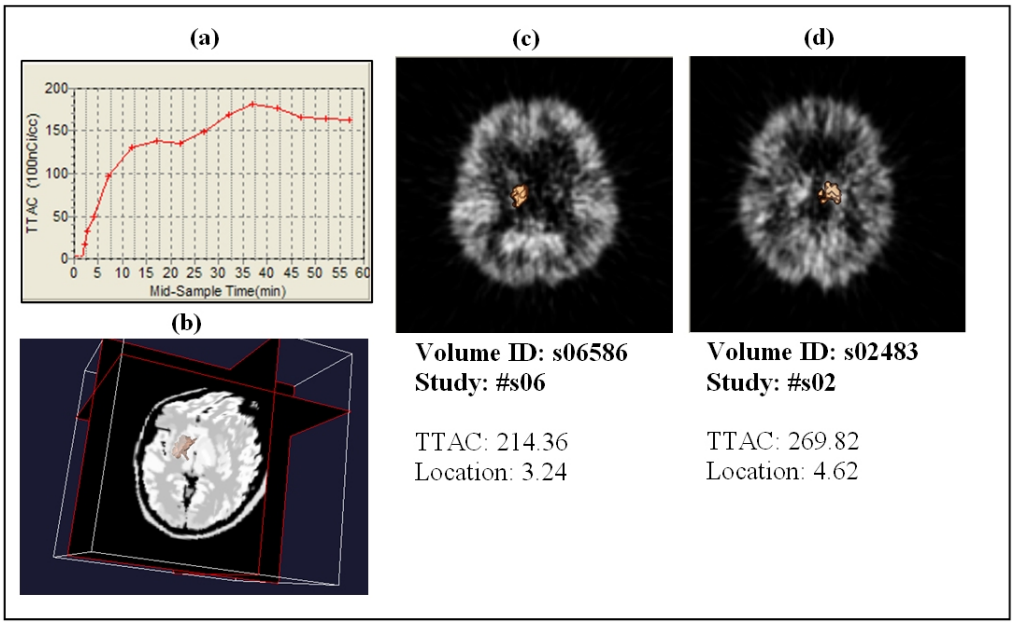

Figure 4.7 (a) The sample TTAC curve; (b) Selection of a location (right thalamus) in the standard atlas (Zubal) panel; (c) Retrieved result with the combination query features of (a) and (b); (d) Retrieved result with (a) and a different location feature (left thalamus) as the query feature [161]

4.6 SUMMARY

This chapter introduced the content-based image retrieval (CBIR) and its key components including image feature extraction, similarity comparison, indexing scheme, and interactive query interface. The need for the CBIR in medical domain (CBMIR) and its related challenges were discussed, and followed by a detailed review on the current major CBMIR techniques in four different categories: retrieval based on physical visual features (color and texture); retrieval based on geometric spatial features (shape, 3D volumetric features and spatial relationships); retrieval by combination of semantic and visual features (semantic pathology interpretation and generic models); and retrieval based on physiological functional features. The success of the CBMIR could open up many new vistas in medical services and research, such as disease tracking, differential diagnosis, non-invasive surgical planning, clinical training, and outcomes research.

4.7 EXERCISES

1. Describe the mechanism of content-based image retrieval technique.
2. What are the primary differentiating factor between CBIR and CBMIR?
3. Texture as a visual feature has been successfully applied in numerous CBMIR systems, i.e., used in MRI head images and HRCT images of the lung. What is an image texture and what are its attributes that enables content-based retrieval?
4. Why can 3D volumetric features be used in the CBMIR? What are the advantages and disadvantages of 3D volumetric features versus 2D shape features?
5. Give an example of the CBMIR application in clinical decision support.
6. What are the advantages and disadvantages of combining semantic and visual features in CBMIR? How does combing these two components exceed expected results from just using one component?

ACKNOWLEDGEMENT

This work is partially supported by ARC and PolyU/UGC grants.

BIBLIOGRAPHY AND REFERENCES

1. I. Bankman (Ed.), *Handbook of Medical Imaging Processing and Analysis*, San Diego: Academic Press, 2000.
2. H.K. Huang, *PACS and Imaging Informatics: Basic Principles and Applications*, Wiley-Liss, 2nd edition, 2004.
3. J. Duncan, N. Ayache, "Medical image analysis: Progress over two decades and the challenges ahead", *IEEE Trans. PAMI.*, vol.22, no.1, pp.85-106, 2000.
4. H.D. Tagare, C. Jaffe, J. Duncan, "Medical image databases: a content-based retrieval approach", *J. Am. Med. Inform. Assoc.*, vol. 4, no. 3, pp184-198, 1997.
5. L.H.Y. Tang, R. Hanka, H.H.S. Ip, "A review of intelligent content-based indexing and browsing of medical images", *Health Informatics J.*, vol.5, pp40-49, 1999.
6. H. Müller, N. Michoux, D. Bandon, A. Geissbuhler, "A review of content-based image retrieval systems in medical applications – clinical benefits and future directions", *Int. J. Med. Info.*, vol.73, pp.1-23, 2004.
7. J. Boissel, M. Cucherat, E. Amsallem, P. Nony, M. Fardeheb, W. Manzi, M. Haugh, "Getting evidence to pre-scribers and patients or how to make EBM a reality", *Proc. Med. Info. Europe Conf.*, France, 2003.
8. Y. Rui, T. S. Huang, S.-F. Chang, "Image retrieval: past, present, and future", *Proc. Int. Symposium on Multimedia Information Processing*, Taipei, Taiwan, 1997.
9. A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, "Content-based image retrieval at the end of the early years", *IEEE Trans. PAMI.*, vol.22, no.12, pp.1349-1380, 2000.
10. A. D. Bimbo, *Visual Information Retrieval*, San Mateo, CA: Morgan Kauffman Publishers, 1999.
11. V. Castelli, L.D. Bergman (Eds.), *Image Databases: Search and Retrieval of Digital Imagery*, New York: John Wiley & Sons, 2002.
12. D. Feng, W.C. Siu, H.J. Zhang (Eds.), *Multimedia Information Retrieval and Management: Technological Fundamentals and Applications*, Berlin: Springer, 2003.
13. F. Long, H. Zhang, D. Feng, "Fundamental of content-based image retrieval", pp.1-26, in D. Feng, W.C. Siu, H. Zhang (Eds.), *Multimedia Information Retrieval and Management: Technological Fundamentals and Applications*, Berlin: Springer, 2003.
14. Y. Rui, T.S. Huang, S.F. Chang, "Image retrieval: Current techniques, promising directions and open issues", *J. Vis. Comm. Image Rep.*, vol.10, pp.39-62, 1999.
15. R. M. Rangayyan, *Biomedical Image Analysis*, CRC Press, 2005.
16. J. Vendrig, M. Worring, A. Smeulders, "Filter image browsing: exploiting interaction in retrieval", *Proc. Viusl'99 Information and Information System*, 1999.
17. J. Robinson, "The k-d-B-tree: A search structure for large multidimensional dynamic indexes", *Proc. SIGMOD Conf.*, Ann Arbor, April 1981.
18. D. Lomet, B. Salzberg, "A robust multimedia-attribute search structure", *5th Int. Conf. Data Eng.*, pp.296-304, 1989.
19. T. Brinkhoff, H. Kriegel, B. Seeger, "Efficient processing of spatial joins using R-trees", *Proc. ACM SIGMOD*, pp.237-246, 1993.
20. A. Guttman, "R-tree: A dynamic index structure for spatial searching", *Proc. ACM SIGMOD*, pp.47-57, 1984.
21. T. Sellis, N. Roussopoulos, C. Faloutsos, "The R^+-tree: A dynamic index for multidimensional objects", *Proc. 12th VLDB*, pp.507-518, 1987.
22. N. Beckmann, H. Kriegel, R. Schneider, B. Seeger, "The R*-tree: An efficient and robust access method for points and rectangles", *Proc. ACM SIGMOD*, pp.322-331, 1990.
23. S. Berchtold, D.A. Keim, H.-P. Kriegel, "The X-tree: An index structure for high-dimensional data", *Prod. 22nd Int. Conf. on Very Large Data Bases*, pp.28-39, Bombay, India, 1996.
24. K.-I. Lin, H.V. Jagadish, C. Faloutsos, "The TV tree: An index structure for high-dimensional data", *VLDB J.*, vol.3, no.4, pp.517-549, 1994.
25.] D. A. White, R. Jain, "Algorithms and strategies for similarity retrieval", *Technical Report VCL-96-01*, Visual Computing Laboratory, University of California, San Diego, 1996.

26. M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, P. Yanker, "Query by image and video content: The QBIC system", *IEEE Computer*, vol.28, no.9, pp.23-32, Sept. 1995.
27. D. White, R. Jain, "Similarity indexing: Algorithms and performance", *Proc. SPIE Storage and Retrieval for Image and Video Databases*, 1996.
28. R. Ng, A. Sedighian, "Evaluating multi-dimensional indexing structures for images transformed by principal component analysis", *Proc. SPIE Storage and Retrieval for Image and Video Databases*, 1996.
29. Y.A. Aslandogan, C.T. Yu, "Techniques and systems for image and video retrieval", *IEEE Trans. KDE*, vol.11, no.1, pp.56-63, Jan/Feb, 1999.
30. Y. Rui, T.S. Huang, M. Ortega, S. Mehrotra, "Relevance feedback: A power tool for interactive content-based image retrieval", *IEEE Trans. Circ. Sys. Video Tech.*, vol.8, no.5, pp.644-655, 1998.
31. J. Huang, S. Kumar, M. Metra, "Combining supervised learning with color correlograms for content-based image retrieval", *Proc. ACM Multimedia '95*, pp.325-334, Nov. 1997.
32. R. Torres, A. Falcao, "Content-based image retrieval: Theory and applications", *Revista de Informatica Teorica e Aplicada*, 13(2), pp.161-185, 2006.
33. X.S. Zhou, T.S. Huang, "Relevance feedback in image retrieval: A comprehensive review", *Multimedia Systems*, vol.8, pp.536-544, 2003.
34. C. Lopez-Pujalte, V. Bote, F. Anegon, "Order-based fitness functions for genetic algorithms applied to relevance feedback", *J. Am. Soc. Info. Sci. and Tech.*, vol.54, no.2, pp.152-160, 2003.
35. Y. Rui, T.S. Huang, M. Ortega, S. Mehrotra, "A power tool in interactive content-based image retrieval", *IEEE Trans. Cir. Sys. Video Tech.*, vol.8, no.5, pp.644-655, 1998.
36. I. Cox, M. Miller, T. Minka, T. Papathomas, P. Yianilos, "The Bayesian image retrieval system, PicHunter: Theory, implementation, and psychophysical experiments", *IEEE Trans. Image Processing*, vol.9, no.1, pp.20-37, 2000.
37.] S. Tong, E. Chang, "Support vector machine active learning for image retrieval", *Proc. 9th ACM Int. Conf. Multimedia*, pp.107-118, NY, USA, 2001.
38. J. Huang, S. R. Kumar, M. Metra, W. J. Zhu, R. Zabith, "Spatial color indexing and applications", *Int. J. Computer Vision*, vol.35, no.3, pp.245-268, 1999.
39. G. Pass, R. Zabith, "Comparing images using joint histograms", *Multimedia Systems*, vol.7, pp.234-240, 1999.
40. W. Niblack et al., "Querying images by content using color, texture, and shape", *SPIE Conf. Storage and Retrieval for Image and Video Database*, vol.1908, pp.173-187, April 1993.
41. G. Pass, R. Zabith, "Histogram refinement for content-based image retrieval", *IEEE Workshop on Applications of Computer Vision*, pp.96-102, 1996.
42. J. Huang, *et al.*, "Image indexing using color correlogram", *IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp.762-768, Puerto Rico, June 1997.
43. P. Howarth, A. Yavlinsky, D. Heesch, S. Rüger, "Medical image retrieval using texture locality and colour", *CLEF2004 – LNCS*, vol.3491, pp.740-749, 2005.
44. B.S. Manjunath, J.R. Ohm, "Color and texture descriptors", *IEEE Trans. Cir. Sys. Video Tech.*, vol.11, pp.703-715, 2001.
45. J. R. Smith, S.-F. Chang, "Automated binary texture feature sets for image retrieval", *Proc. ICASSP*, Atlanta, 1996.
46. R. M. Haralick, K. Shanmugam, I. Dinstein, "Texture features for image classification", *IEEE Trans. Sys., Man., and Cyb.*, vol.SMC-3, no.6, 1973.
47. H. Tamura, S. Mori, T. Yamawaki, "Texture features corresponding to visual perception", *IEEE Trans. Sys., Man., and Cyb.*, vol.Smc-8, no.6, June 1978.
48. X. Tang, "Texture information in Run-length matrices", *IEEE Trans. Image Processing*, vol.7, no.11, pp.1602-1609, Nov. 1998.
49. A. Laine, J. Fan, "Texture classification by wavelet packet signatures", *IEEE Trans. PAMI.*, vol.15, no.11, pp.1186-1191, Nov. 1993.
50. T. Chang, C.C.J. Kuo, "Texture analysis and classification with tree-structured wavelet transform", *IEEE Trans. Image Processing*, vol.2, no.4, pp.429-441, October 1993.

51. J. G. Daugman, "Complete discrete 2D Gabor transforms by neural networks for image analysis and compression", *IEEE Trans. ASSP.*, vol.36, pp.1169-1179, July 1988.
52. B. S. Manjunath, W. Y. Ma, "Texture features for browsing and retrieval of image data", *IEEE Trans. PAMI.*, vol.18, no.8, pp.837-842, Aug. 1996.
53. J. Francos, "Orthogonal decompositions of 2D random fields and their applications in 2D spectral estimation", N. K. Bose, C. R. Rao (Eds.), *Signal Processing and its Applications*, pp.20-27, North Holland, 1993.
54. F. Liu, R. W. Picard, "Periodicity, directionality, and randomness: Wold features for image modeling and retrieval", *IEEE Trans. PAMI.*, vol.18, no.7, July 1996.
55. P. P. Ohanian, R. C. Dubes, "Performance evaluation for four classes of texture features", *Pattern Recognition*, vol.25, no.8, pp.819-833, 1992.
56. C. Shyu, C. Brodley, A. Kak, A. Kosaka, A. Aisen, L. Broderick, "ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases", *Comput. Vis. Image Understand*, 75(1-2):111-132, 1999.
57. J. Mao, A. K. Jain, "Texture classification and segmentation using multiresolution simultaneous autoregressive models", *Pattern Recognition*, vol.25, no.2, pp.173-188, 1992.
58. J. Weszka, C. Dyer, A. Rosenfild, "A comparative study of texture measures for terrain classification", *IEEE Trans. Sys., Man., and Cyb.*, vol.SMC-6, no.4, 1976.
59. A. P. Pentland, "Fractal-based description of natural scenes", *IEEE Trans. PAMI.*, vol.6, no.6, pp.661-674, 1984.
60. C. Chatfield, A. Collins, *Introduction to multivariate analysis*, London: Chapman & Hall, 1983.
61. E. Persoon, K. Fu, "Shape discrimination using Fourier descriptors", *IEEE Trans. Sys., Man., and Cyb.*, vol.7, pp.170-179, 1977.
62. H. Kauppinen, T. Seppnaen, M. Pietikainen, "An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classification", *IEEE Trans. PAMI.*, vol.17, no.2, pp.201-207, 1995.
63. E. M. Arkin, L. Chew, D. Huttenlocher, K. Kedem, J. Mitchell, "An efficiently computable metric for comparing polygonal shapes", *IEEE Trans. PAMI.*, vol.13, no.3, 1991.
64. A. Pentland, R.W. Picard, S. Sclaroff, "Photobook: Content-based manipulation of image databases", *Int. J. Computer Vision*, 1996.
65. S. Abbasi, F. Mokhtarian, J. Kittler, "Enhancing CSS-based shape retrieval for objects with shallow concavities", *Image and Vision Computing*, vol.18, no.3, pp.199-211, 2000.
66. F. Mokhtarian, S. Abbasi, "Shape similarity retrieval under Affine transforms", *Pattern Recognition*, vol.35, no.1, pp.31-41, 2002.
67. Z. You, A.K. Jain, "Performance evaluation of shape matching via chord length distribution", *Computer Vision, Graphics, Image Processing*, vol.28, pp.185-198, 1984.
68. A.K. Jain, *Fundamentals of Digital Image Processing*, New York: Prentice Hall, 1986.
69. N. Arica, F. Vural, "BAS: A perceptual shape descriptor based on the beam angle statistics", *Pattern Recognition Letters*, vol.24, no.9-10, pp.1627-1639, 2003.
70. G. C.-H. Chuang, C.-C. J. Kuo, "Wavelet descriptor of planar curves: Theory and applications", *IEEE Trans. Image Processing*, vol.5, no.1, pp.56-70, Jan. 1996.
71. M. K. Hu, "Visual pattern recognition by moment invariants, computer methods in image analysis", *IRE Trans. Info. Theory*, 8, 1962.
72. L. Yang, F. Algregtsen, "Fast computation of invariant geometric moments: A new method giving correct results", *Proc. IEEE Int. Conf. on Image Processing*, 1994.
73. Y.S. Kim, W.Y. Kim, "Content-based trademark retrieval system by using visually salient feature", *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp.307-312, 1997.
74. L. Prasad, "Morphological analysis of shapes", *CNLS Research Highlights*, Los Alamos National Laboratory, Los Alamos, NM, July 1997.
75. D. H. Ballard, D. M. Brown, *Computer Vision*, Englewood cliffs, N.J.: Prentice Hall, 1982.
76. S.-K. Chang, *Principles of Pictorial Information Systems Design*, Englewood Cliffs, N.J.: Prentice Hall Int'l Editions, 1989.

77. S. K. Chang, Q. Y. Shi, C. Y. Yan, "Iconic indexing by 2-D strings", *IEEE Trans. PAMI.*, vol.9, no.3, pp.413-428, May 1987.
78. H. Samet, "The quadtree and related hierarchical data structures", *ACM Computing Surveys*, 16(2):187-260, 1984.
79. V. N. Gudivada, V. V. Raghavan, "Design and evaluation of algorithms for image retrieval by spatial similarity", *ACM Trans. Info. Sys.*, vol.13, no.2, pp.115-144, April 1995.
80. M.E. Mattie, L. Staib, E. Stratmann, H.D. Tagare, J. Duncan, P.L. Miller, "PathMaster: Content-based cell image retrieval using automated feature extraction", *J. Am. Med. Informatics Assoc.*, vol.7, no.4, pp.404-415, 2000.
81. Q. Wang, V. Megalooikonomou, D. Kontos, "A medical image retrieval framework", *Proc. IEEE Workshop on Machine Learning for Signal Processing (MLSP'05)*, pp.233-238, 2005.
82. C. Brodley, A. Kak, C. Shyu, J. Dy, L. Broderick, A.M. Aisen, "Content-based retrieval from medical image databases: A synergy of human interaction, machine learning and computer vision", *Proc. the 10th National Conf. on Artificial Intelligence*, pp.760-767, Orlando, FL, USA, 1999.
83. A. Marchiori, C. Brodley, J. Dy, C. Pavlopoulou, A. Kak, L. Broderick, A.M. Aisen, "CBIR for medical images – An evaluation trial", *Pro. IEEE Workshop on Content-Based Access of Image and Video Libraries*, pp.89-93, 2001.
84. I. El-Naqa, Y. Yang, N.P. Galatsanos, R.M. Nishikawa, M.N. Wernick, "A similarity learning approach to content-based image retrieval: Application to digital mammography", *IEEE Tran. Med. Imaging*, 23(10):1233-1244, 2004.
85. S.T.C. Wong, "CBIR in medicine: still a long way to go", *Proc. IEEE Workshop on Content-Based Access of Image and Video Libraries*, pp.114, 1998.
86. S.T.C. Wong, H.K. Huang, "Design methods and architectural issues of integrated medical image data base systems", *Computerized Medical Imaging and Graphics*, vol.20, no.4, pp.285-299, 1996.
87. T. Glatard, J. Montagnat, I.E. Magnin, "Texture based medical image indexing and retrieval: Application to cardiac imaging", *Proc. the 6th ACM SIGMM Int. Workshop on Multimedia Info. Retrieval*, pp.135-142, 2004.
88. M. M. Rahman, P. Bhattacharya, B. C. Desai, "A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback", *IEEE Trans. Info. Tech. Biomed.*, vol.11, no.1, pp.58-69, 2007.
89. C.H. Li, P.C. Yuen, "Regularized color clustering in medical image database", *IEEE Trans. Med. Imag.*, vol.19, no.11, pp1150-1155, 2000.
90. S. Tamai, "The color of digital imaging in pathology and cytology", *Digital Color Imaging in Biomedicine*, no.2, pp.61-66, 2001.
91. H.K. Choi, H.G. Hwang, M.K. Kim, T.Y. Kim, "Design of the breast carcinoma cell bank system", *Proc. the 6th Int. Workshop on Enterprise Networking and Computing in Healthcare Industry (HEALTHCOM'04)*, pp.88-91, 2004.
92. F. Schnorrenberg, C.S. Pattichis, C.N. Schizas, K. Kyriacou, "Content-based retrieval of breast cancer biopsy slides", *Tech. and Health Care*, vol.8, pp.291-297, 2000.
93. L. Zheng, A.W. Wetzel, J. Gilbertson, M.J. Becich, "Design and analysis of a content-based pathology image retrieval system", *IEEE Trans. Info.0 Tech. Biomed.*, vol. 7, no. 4, pp.249-255, Dec. 2003.
94. H. L. Tang, R. Hanka, H.S. Ip, "Histological image retrieval based on semantic content analysis", *IEEE Trans. Info. Tech. Biomed.*, vol. 7, no. 1, pp.26-36, March 2003.
95. R.W.K. Lam, K.T. Cheung, H.S. Ip, L.H.Y. Tang, R. Hanka, "An iconic and semantic content based retrieval system for histological images", *VISUAL 2000- LNCS*, vol.1929, pp.384-395, 2000.
96. L. H. Tang, R. Hanka, R. Lan, H. H. S. Ip, "Automatic semantic labelling of medical images for content-based retrieval", *Proc. the Int. Conf. Artificial Intelligence, Expert Systems and Applications (EXPERSYS 1998)*, pp.77-82, Virginia Beach, VA, USA, 1998.
97. L. H. Tang, R. Hanka, H. H. S. Ip, R. Lam, "Extraction of semantic features of histological images for content-based retrieval of images", *Proc. the IEEE Symp. Computer-Based Medical Systems (CBMS 2000)*, Houston, TX, USA, 2000.
98. M. Nischik, C. Forster, "Analysis of skin erythema using true-color images", *IEEE Trans. Med. Imag.*, vol.16, no.12, pp711-716, 1997.
99. G.L. Hansen, E.M. Sparrow, J.Y. Kokate, K.Y. Leland, P.A. Iaizzo, "Wound status evaluation using color image processing", *IEEE Trans. Med. Imag.*, vol.16, no.2, pp78-86, 1997.

100. M. M. Rahman, B. c. Desai, P. Bhattacharya, "Image retrieval-based decision support system for dermatoscopic images", *Proc. the 19th IEEE Symp. on Computer-Based Medical Systems*, pp.285-290, 2006.
101. M. Nishibori, "Problems and solutions in medical color imaging", *Proc. Second Int. Symp. on Multispectral Imaging and High Accurate Color Reprocuction*, pp.9-17, 10-11 October, 2000.
102. B.V. Dhandra, R. Hegadi, M. Hangarge, V.S. Malemath, "Analysis of abnormality in endoscopic images using combined HIS color space and watershed segmentation", *The 18th Int. Conf. on Pattern Recognition*, 2006.
103. S. Xia, W. M, Y. Yan, X. Chen, "An endoscopic image retrieval system based on color clustering method", *Third Int. Symp. on Multispectral Image Processing and Pattern Recognition, Proc. the SPIE*, vol.5286, pp.410-413, 2003.
104. K. B. Kim, S. Kim, G. H. Kim, "Analysis system of endoscopic image of early gastric cancer", *IEEE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, vol.E89-A, no.10, pp.2662-2669, 2006.
105. U. Honmyo, A. Mitsui, A. Murakami, S. Mizumoto, I. Yoshinada, M. Maeda, S. Yamamoto, S. Shimada, "Mechanisms producing color change in flat early gastric cancers", *Endoscopy*, vol.29, pp.366-371, 1997.
106. T. Ogihara, H. Watanabe, M. Namihisa, N. Sato, "Display of mucosal blood flow function and color enhancement based on blood flow index (IHb color enhancement", *Clinical Gastroenterology*, vol.12, pp.109-117, 1997.
107. S. Tsuji, N. Sato, S. Kawano, T. Kamada, "Functional imaging for the analysis of the mucosal blood hemoglobin distribution using electronic endoscopy", *Gastrointest Endosc*, vol.34, pp.332-336, 1988.
108. M. P. Tjoa, S. M. Krishnan, "Feature extraction for the analysis of colon status from the endoscopic images", *BioMedical Engineering OnLine*, vol.2, no.9, pp1-17, 2003.
109. Z.L. Chen, *Research on Tongue Diagnosis*, Shanghai Science and Technology Publishing House, Shanghai, China, 1982.
110. G. Maciocia, *Tongue Diagnosis in Chinese Medicine*, Eastland Press, 1995.
111. L.S. Shen, B.G. Wei, Y.H. Cai, X.F. Zhang, Y.Q. Wang, "Image analysis for tongue characterization", *Chinese J. Electronics*, vol.12, no.3, pp317-323, 2003.
112. N.M. Li, et al. "The contemporary investigations of computerized tongue diagnosis", in: *The Handbook of Chinese Tongue Diagnosis*, Shed-Yuan Publishing, Peking, pp1315-1317, 1994.
113. B.G. Wei, Y.H. Cai, X.F. Zhang, L.S. Shen, "Recent progresses in analysis of tongue manifestation for traditional Chinese medicine", *Chinese J. Biomed. Eng.*, vol.14, no.2, pp55-64, 2005.
114. B. Pang, D. Zhang, K. Wang, "Tongue image analysis for appendicitis diagnosis", *Info. Sci.*, vol.175, pp160-176, 2005.
115. C.H. Li, P.C. Yuen, "Tongue image matching using color content", *Pattern Recognition*, vol.35, pp407-419, 2002.
116. Y.G. Wang, Y. Zhou, J. Yang, Q. Xu, "An image analysis system for tongue diagnosis in traditional Chinese medicine", *CIS 2004, LNCS 3314*, pp1181-1186, 2004.
117. C.C. Chiu, "A novel approach based on computerized image analysis for traditional Chinese medical diagnosis of the tongue", *Computer Methods and Programs in Biomedicine*, vol.61, pp77-89, 2000.
118. J.H. Jang, J.E. Kim, K.M. Park, S.O. Park, Y.S. Chang, B.Y. Kim, "Development of the digital tongue inspection system with image analysis", *Proc. the Second Joint EMBS/BMES Conf.*, pp1033-1034, Houston, TX, USA, October 23-26, 2002.
119. A.M. Wang, *Research on Image Analysis for Tongue Characterization*, PhD Dissertation, Beijing Polytechnic University, Beijing, 2001.
120. X.F. Zhang, *The Primary Study of Classification and Recognition on Tongue Manifestation in Traditional Chinese Medicine*, PhD Dissertation, 2006.
121. B.G. Wei, *Research on Color Reproduction and Analysis of Texture, Shape and State of Tongue for Traditional Chinese Medicine*, PhD Dissertation, 2005.
122. H. Muller, A. Rosset, J.-P. Vallee, A. Geissbuhler, "Integrating content-based visual access methods into a medical case database", *Proc. the Medical Informatics Europe Conf.*, pp.480-485, St. Malo, France, 2003.
123. A. Rosset, H. Muller, M. Martins, N. Dfouni, J.-P. Vallee, O. Ratib, "Casimage project – a digital teaching files authoring environment", *J. Thoracic Imaging*, vol.19, no.2, pp.1-6, 2004.

124. P. Clough, H. Muller, M. Sanderson, "The CLEF 2004 cross-language image retrieval track", *Proc. 5th Workshop Cross-Language Evaluation Forum, CLEF 2004*, vol.3491, pp.597-613, Bath, U. K., Sep. 2004.
125. D. Xu, A.S. Kurani, J.D. Furst, D.S. Raicu, "Run-length encoding for volumetric texture", *Proc. VIIP* 2004.
126. T.M. Lehmann, M.O. Guld, D. Keysers, T. Deselaers, H. Schubert, B. Wein, K. Spitzer, "Similarity of medical images computed from global feature vectors for content-based retrieval", *Proc. KES2004 – LNAI*, vol. 3214, pp.989-995, 2004.
127. D. S. Raicu, J. D. Furst, D. Channin, D. Xu, A. Kurani, S. Aioanei, "A texture dictionary for human organs tissues'classification", *Proc. 8th World Multiconf. on Sys. Cyb. and Info.(SCI2004)*, USA, July 18-21, 2004.
128. W. Tsang, A. Corboy, K. Lee, D. Raicu, J. Furst, "Texture-based image retrieval for computerized tomography databases", *Proc.18th IEEE Symp. on Computer-Based Medical Systems (CBMS'05)*, pp.593-598, 2005.
129. S. Orphanoudakis, C. Chronaki, D. Vamvaka, "I2Cnet: Content-based similarity search in geographically distributed repositories of medical images", *Computerized Medical Imaging and Graphics*, vol.20, no.4, pp.193-207, 1996.
130. P.A. Freeborough, N.C. Fox, "MR image texture analysis applied to the diagnosis and tracking of Alzheimer's disease", *IEEE Trans. Med. Imag.*, vol.17, no.3, pp.475-479, June 1998.
131. J.C. Felipe, A.J.M. Traina, C. Traina, "Retrieval by content of medical images using texture for tissue identification", *Proc. 16th IEEE Symp. on Computer-Based Medical Systems (CBMS'03)*, pp.175-180, 2003.
132. D.M. Kwak, B.S. Kim, O.K. Yoon, C.H. Park, J.U. Won, K.H. Park, "Content-based ultrasound image retrieval using a coarse to fine approach", *Ann. N.Y. Acad. Sci.*, vol. 980, pp.212-224, 2002.
133. M.A. Sheppard, L. Shih, "Efficient image texture analysis and classification for prostate ultrasound diagnosis", *Proc. IEEE Conf. Computational Systems Bioinformatics*, pp.7-8, 2005.
134. H. Alto, R.M. Rangayyan, J.E. Leo Desautels, "Content-based retrieval and analysis of mammographic masses", *J. Electronic Imaging*, vol.14, no.2, pp.023016-1-17, 2005.
135. S. Xia, D. Ge, W. Mo, Z. Zhang, "A content-based retrieval system for endoscopic images", *Proc. the 27th Annual Conf. the IEEE EMBS*, pp.1720-1723, Shanghai, China, Sep. 1-4, 2005.
136. A. K. Jain, F. Farroknia, "Unsupervised texture segmentation using Gabor filters", *Pattern Recognition*, vol.24, no.12, pp.1167-1186, 1991.
137. C.G. Zhao, H.Y. Cheng, Y.L. Huo, T.G. Zhuang, "Liver CT-image retrieval based on Gabor texture", *Proc. the 26th Annual Conf. the IEEE EMBS*, pp.1491-1494, San Francisco, CA, USA, Sept. 1-5, 2004.
138. D. Zhao, Y. Chen, H. Correa, "Statistical categorization of human histological images", *Proc. IEEE Int. Conf. on Image Processing (ICIP'05)*, pp.628-631, 2005.
139. S. Antani, D.J. Lee, L.R. Long, G.R. Thoma, "Evaluation of shape similarity measurement methods for spine X-ray images", *J. Vis. Commun. Image R.*, vol.15, pp.285-302, 2004.
140. D. Comaniciu, D. Foran, P. Meer, "Shape-based image indexing and retrieval for diagnostic pathology", *Proc. the 14th Int. Conf. on Pattern Recognition*, pp.902-904, 1998.
141. R. M. Rangayyan, N. M. El-Faramawy, J. E. Leo, O. A. Alim, "Measures of acutance and shape for classification of breast tumors", *IEEE Trans. on Med. Imag.*, vol.16, no.6, pp.799-810, 1997.
142. S. Ciatto, L. Cataliotti, v. Distante, "Nonpalpable lesions detected with mammography:Review of 512 consecutive cases", *Radiol.*, vol.165, no.1, pp.99-102, 1987.
143. R. M. Rangayyan, S. G. Elkadiki, "Algorithm for the computation of region-based image edge profile acutance", *J. Electron. Imag.*, vol.4, no.1, pp.62-70, 1995.
144. M. K. Hu, "Visual pattern recognition by moment invariants", in J. K. Aggarwal, R. O. Duda, A. Rosenfeld, *Computer Methods in Image Analysis*, IEEE Computer Society, Los Angeles, CA, 1977.
145. R. C. Gonzalez, R. E. Woods, *Digital Image Processing*, New Jersey: Prentice-Hall, 2002.
146. S. Zhu, G. Schaefer, "Thermal medical image retrieval by moment invariants", *Int. Symposium on Biological and Medical Data Analysis – LNCS*, vol.3337, pp.182-187, 2004.
147. S. Maitra, "Moment invariants", *Proc. the IEEE*, vol.67, pp.697-699, 1979.
148. G. Schaefer, S.Y. Zhu, S. Ruszala, "Visualisation of medical infrared image databases", *Proc. the 27th Annual Conf. of the IEEE EMBS*, pp.634-637, Shanghai, China, Sep. 1-4, 2005.

149. G. P. Robinson, H. D. Targare, J. S. Duncan, C. C. Jaffe, "Medical image collection indexing: shape-based retrieval using KD-trees", *Comput. Vis. Graphics Image Proces.*, vol.20, no.4, pp.209-217, 1996.
150. A.J.M. Traina, A.G.R. Balan, L.M. Bortolotti, C. Traina Jr., "Content-based image retrieval using approximate shape of objects", *Pro. 17th IEEE Symposium on Computer-Based Medical Systems (CBMS'04)*, pp.91-96, 2004.
151. W. Zhang, S. Dickinson, S. Sclaroff, J. Feldman, S. Dunn, "Shape-based indexing in a medical image database", *Proc. IEEE Workshop on Biomed. Image Analysis*, pp.221-230, 1998.
152. J. Felipe, M. Ribeiro, E. Sousa, A. Traina, C. Traina Jr, "Effective shape-based retrieval and classification of mammograms", *Proc. the 2006 ACM Symp. on Applied Computing*, pp.250-255, 2006.
153. P. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, Z. Protopapas, "Fast and effective retrieval of medical tumor shapes", *IEEE Trans. KDE.*, vol. 10, no. 6, Nov./Dec. 1998.
154. J.Z. Wang, "Pathfinder: Multiresolution region-based searching of pathology images using IRM", *J. Am. Med. Informatics Asso. (AMIA) – Proc. the AMIA Annual Symp., vol. 2000 Symposium Suppl.*, pp.883-887, Los Angeles, CA, USA, Nov. 2000.
155. C. Ng, G. Martin, "Content-description interfaces for medical imaging", *Technical Report CS-RR-383*, Covertry, UK, 2001.
156. W. Liu, Q. Tong, "Medical image retrieval using salient point detector", *Proc. the 2005 IEEE EMBS 27th Annual Conference*, pp.6352-6355, Shanghai, China, Sept. 1-, 2005.
157. S. Chu, S. Narayanan, C.-C. Kuo, "Efficient rotation invariant retrieval of shapes with applications in medical databases", *Pro. 19th IEEE Symp. on Computer-Based Medical Systems (CBMS'06)*, pp.673, 678, 2006.
158. P.A. Mlsna, N.M. Sirakov, "Intelligent shape feature extraction and indexing for efficient content-based medical image retrieval", *Pro. 6th IEEE Southwest Symp. on Image Analysis and Interpretation*, pp.172-176, 2004.
159. B. Fischer, C. Thies, M.O. Guld, T.M. Lehmann, "Content-based image retrieval by matching hierarchical attributed region adjacency graphs", *Proc. SPIE – Medical Imaging: Image Processing*, vol.5370, pp598-606, 2004.
160. J. K. Udupa, G. T. Herman, *3D Imaging in Medicine*, CRC Press, 2000.
161. J. Kim, W. Cai, D. Feng, H. Wu, "A new way for multidimensional medical data management: volume of interest (VOI)-based retrieval of medical images with visual and functional features", *IEEE Trans. Info. Tech. Biomed.*, vol. 10, no. 3, pp.598-607, July, 2006.
162. Y. Liu, N.A. Lazar, W.E. Rothfus, F. Dellaert, A. Moore, J. Schneider, T. Kanade, "Semantic-based biomedical image indexing and retrieval", in: *Trends and Advances in Content-Based Image and Video Retrieval*, L. Shapiro, H. Kriegel, nd R. Veltkamp, Eds., Springer, 2004.
163. Y. Liu, R. T. Collins, W. E. Rothfus, "Robust midsagittal plane extraction from normal and pathological 3-D Neuroradiology images", *IEEE Trans. Med. Imag.*, vol.20, no.3, pp.175-192, 2001.
164. Y. Liu, W. E. Rothfus, T. Kanade, "Content-based 3D Neuroradiologic image retrieval: Preliminary results", *Proc. IEEE Workshop Content-based Access of Image and Video Libraries, in Conjunction with Int. Conf. Computer Vision*, pp.91-100, Bombay, India, Jan. 1998.
165. J. Declerck, G. Subsol, J-P Thirion, N. Ayache, "Automatic retrieval of anatomical structures in 3d medical images", *Technical Report 2485*, INRIA, Sophia-Antipolis, France, 1995.
166. A. Guimond, G. Subsol, "Automatic MRI database exploration and applications", *Int. J. Pattern Recog. Artif. Intell.*, vol.11, pp.1345-1365, 1997.
167. V. Megalooikonomou, H. Dutta, D. Kontos, "Fast and effective characterization of 3D region data", *Proc. IEEE Int. Conf. on Image Processing (ICIP'02)*, pp.421-424, 2002.
168. S. Minoshima, K. L. Berger, K. S. Lee, M. A. Mintun, "An automated method for rotational correction and centering of three-cimensional functional brain images", *J. Nucl. Med.*, vol.33, pp.1579-1585, 1992.
169. S. Minoshima, R. Koeppe, M. Mintun, K. Berger, S. Taylor, K. Frey, D. Kuhl, "Automated detection of the intercommissural line for stereotactic localization of functional brain images", *J. Nucl. Med.*, 34:322-329, 1993.
170. S. Minoshima, R. A. Koeppe, K. A. Frey, D. E. Kuhl, "Anatomic standardization: Linear scaling and nonlinear warping of functional brain images", *J. Nucl. Med.*, vol.35, pp.1528-1537, 1994.
171. I. G. Zubal, C. R. Harrell, E. O. Smith, Z. Rattner, G. Gindi, P. B. Hoffer, "Computerized three-dimensional segmented human anatomy", *Med. Phys.*, vol.21, pp.299-302, 1994.

172. T.M. Lehmann, B. Plodowski, K. Spitzer, B.B. Wein, H. Ney, T. Seidl, "Extended query refinement for content-based access to large medical image databases", *Proc. SPIE – Medical Imaging: PACS and Imaging Informatics*, vol. 5371 (15), pp.90-98, 2004.
173. W.W. Chu, A.F. Cardenas, R.K. Taira, "KMED: A knowledge-based multimedia medical distributed database system", *Info. Sys.*, vol. 19, no. 4, pp.33-54, 1994.
174. S. Aksoy, G. Marchisio, C. Tusk, K. Koperski, "Interactive classification and content-based retrieval of tissue images", *Proc. SPIE – Applications of Digital Image Processing XXV*, vol. 4790, pp.71-81, 2002.
175. R. Chbeir, F. Favetta, "A global description of medical image with high precision", *Proc. IEEE Int. Symp. on Bio-Informatics and Biomedical Engineering (BIBE'2000)*, pp.289-296, Washington, D. C. USA, Nov. 8-10, 2000.
176. R. Chbeir, Y. Amghar, A. Flory, "MIMS: A prototype for medical image retrieval", *Proc. 6th Int. Conf. of Content Based Multimedia Information Access*, RIAO2000, April 2000, pp.846-861, 2000.
177. E. Petrakis, "Content-based retrieval of medical images", *Int. J. Comput. Res.*, vol. 11, no. 2, pp.171-182, 2002.
178. W.W. Chu, C.C. Hsu, A.F. Cardenas, R.K. Taira, "Knowledge-based image retrieval with spatial and temporal constructs", *IEEE Trans. KDE*, vol.10, no.6, 1998.
179. P.M. Willy, K.H. Küfer, "Content-based medical image retrieval (CBMIR): An intelligent retrieval system for handling multiple organs of interest", *Proc. the 17th IEEE Symp. on Computer-Based Medical Systems (CBMS'04)*, pp.103-108, 2004.
180. C.R. Shyu, C.E. Brodley, A.C. Kak, A. Kosaka, A.M. Aisen, L.S. Broderick, "ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases", *Computer Vision and Image Understanding*, vol. 75, no. 1/2, Jul./Aug., pp.111-132, 1999.
181. C.R. Shyu, C. Pavlopoulou, A.C. Kak, C.E Brodley, "Using human perceptual categories for content-based retrieval from a medical image database", *Computer Vision and Image Understanding*, vol. 88, pp.119-151, 2002.
182. A.S. Barb, C.R. Shyu, Y.P. Sethi, "Knowledge representation and sharing using visual semantic modeling for diagnostic medical image databases", *IEEE Trans. Info. Tech. Biomed.*, vol. 9, no. 4, pp.538-553, Dec. 2005.
183. E. Stern, S. Swensen, *High Resolution CT of the Chest: Comprehensive Atlas*, 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2000.
184. W. Webb, N. Muller, D. Naidich, *High-Resolution CT of Lung*, Philadelphia, PA: Lippincott-Raven, 1996.
185. K.K.T. Cheung, R.W.K. Lam, H.H.S. Ip, R. Hanka, L.H.Y. Tang, G. Fuller, "An object-oriented framework for content-based image retrieval based on 5-tier architecture", *Proc. Asia-Pacific Software Eng. Conf. 99*, pp.174-177, Takamatsu, Japan, Dec. 7-10, 1999.
186. R.W.K. Lam, H.H.S. Ip, K.K.T. Cheung, L.H.Y. Tang, R. Hanka, "A multi-window approach to classify histological features", *Proc. Int. Conf. Pattern Recognition*, vol.2, pp.259-262, Barcelona, Spain, Sept. 2000.
187. K.K.T. Cheung, R. Lam, H.H.S. Ip, L.H.Y. Tang, R. Hanka, "A software framework for combining iconic and semantic content for retrieval of histological images", *VISUAL 2000- LNCS,* vol.1929, pp.488-499, 2000.
188. L.H. Tang, R. Hanka, H.H.S. Ip, K.K.T. Cheung, R. Lam, "An intelligent system for integrating semantic and iconic features for image retrieval", *Proc. Computer Graphics International*, pp.240-245, 2001.
189. L. Tan, R. Hanka, H. Ip, K. Cheung, R. Lam, "Integration of intelligent engines for a large scale medical image database", *Proc. 13th IEEE Symp. on Computer-based Medical Systems*, pp.235-240, 2000.
190. M.C. Jaulent, C.L. Bozec, Y. Cao, E. Zapletal, P. Degoulet, "A property concept frame representation for flexible image content retrieval in histopathology databases", *Proc. the Annual Symp. of the Am. Soc. Med. Informatics (AMIA)*, pp.379-383, Los Angeles, CA, USA, 2000.
191. H. Shao, W.C. Cui, L. Tang, "Medical imge description in content-based image retrieval", *Proc. the 27th Annual Conf. the IEEE EMBS*, pp.6336-6339, Shanghai, China, Sep. 1-4, 2005.
192. T.M. Lehmann, M.O. Guld, C. Thies, B. Fischer, K. Spitzer, D. Keysers, H. Ney, M. Kohnen, H. Schubert, B.B. Wein, "Content-based image retrieval in medical applications", *Methods Inf. Med*. no. 4, pp. 354-361, 2004.
193. D. Keysers, J. Dahmen, H. Ney, B. Wein, T. Lehmann, "A statistical framework for model-based image retrieval in medical applications", *J. Electronic Imaging*, vol.12, no.1, pp.59-68, 2003.
194. C. Hsu, W. Chu, R. Taira, "A knowledge-based approach for retrieving images by content", *IEEE Trans. KDE.*, Aug. 1996.

195. R. Chbeir, F. Favetta, "A global description of medical imaging with high precision", *IEEE Trans. Sys., Man, and Cyb.* – Part B: Cybernetics, vol. 33, no. 5, pp.752-757, Oct. 2003.
196. R. Chbeir, Y. Amghar, A. Flory, L. Brunie, "A hyper-spaced data model for content and semantic-based medical image retrieval", *Proc. ACS / IEEE Int. Conf. Computer Systems and Applications*, pp.161-167, 2001.
197. International Classification of Diseases 10th Revision. http://www.who.int/classifications/icd/en/
198. B.L. Humphreys (Ed.), *UMLS Knowledge Sources – First Experimental Edition Documentation*, Bethesda, MD: National Library of Medicine, 1990.
199. R. Chbeir, Y. Amghar, A. Flory, "System for medical image retrieval the MIMS model", *Proc. The 3rd Int. Conf. Visual (VISUAL'99), LNCS 1614*, pp.37-42, Amsterdam, The Netherlands, June 1999.
200. S. Atnafu, R. Chbeir, L. Brunie, "Content-based and metadata retrieval in medical image database", *Proc. the 15th IEEE Symp. on Computer-Based Medical Systems (CBMS 2002)*, pp.327-332, 2002.
201. R. Chbeir, S. Atnafu, L. Brunie, "Image data model for an efficient multi-criteria query: A case in medical databases", *Proc. 14th Int. Conf. Scientific and Statistical Database Management (SSDBM'02)*, pp.165-174, 2002.
202. T. Lehmann, H. Schubert, D. Keysers, M. Kohnen, B. Wein, "The IRMA code for unique classification of medical images", *Proc. SPIE*, vol.5033, pp.440-451, 2003.
203. S. C. Orphanoudakis, C. Chronaki and S. Kostomanolakis, "I2C: A System for the Indexing, Storage, and Retrieval of Medical Images by Content", *Med. Informatics*, vol. 19, no. 2, pp.109-122, 1994.
204. E. Ei-Kwae, H. Xu, M. Kabuka, "Content-based retrieval in picture archiving and communication systems", *J. of Digital Imaging*, vol.13, no.2, pp.70-81, 2000.
205. H. Lowe, I. Antipov, W. Hersh, C. Smith, M. Mailhot, "Automated semantic indexing of imaging reports to support retrieval of medical images in the multimedia electronic medical record", *Meth Info Med*, 38(303-7) 1999.
206. H. Muller, C. Lovis, A. Geissbuhler, "The medGIFT project on medical image retrieval", *Proc. the 15th IEEE Symp. on Computer-Based Medical Systems (CBMS2002)*, pp.321-326, 2002.
207. [M9] H. Muller, A. Rosset, J.P. Vallee, A. Geissbuhler, "Comparing feature sets for content-based image retrieval in a medical case database", *Proc. SPIE 2004*, 2004.
208. http://ww.gnu.org/software/gift/
209. http://www.mrml.net/
210. D. Feng, "Information technology applications in biomedical functional imaging", *IEEE Trans. Info. Tech. Biomed.*, vol.3, no.3, pp.221-230, 1999.
211. D. Feng, D. Ho, H. Iida & K. Chen, "Techniques for Functional Imaging", an invited chapter contributing to *Medical Imaging Techniques and Applications*, Edited by: C. T. Leondes, in 'Gordon and Breach International Series in Engineering, Technology and Applied Science', *Gordon and Breach Science Publishers*, pp.85-145, 1997.
212. W. Cai, D. Feng and R. Fulton, "Content-Based Retrieval of Dynamic PET Functional Images", *IEEE Trans. Info. Tech. Biomed.*, vol. 4, no. 2, pp.152-158, 2000.
213. S. C. Huang, M. E. Phelps, E. J. Hoffman, K. Sideris, C. 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.
214. 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, pp. 697-710, 1995.
215. J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithm*, Norwell, MA: Kluwer, 1981.
216. J. Kim, W. Cai, D. Feng, S. Eberl, "An objective evaluation framework for segmentation techniques of functional positron emission tomography studies", *IEEE NSS-MIC Conf.*, vol.5, pp.3217-3221, Oct. 16-22, 2004.
217. X. Li, D. Feng, 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.
218. D. Feng, W. Cai, 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, pp.361, 1998.

4.9 INDEX