

---

## 25. Mining Knowledge in Computer Tomography Image Databases

Daniela Stan Raicu

**Summary:** This chapter presents our research results obtained for texture extraction, classification, segmentation, and retrieval of normal soft tissues in Computed Tomography studies of the chest and abdomen. The texture extraction step consists of various texture methods applied to the collection of tissue data in order to derive a set of features characterizing the best the visual perception of texture. The classification step involves different data mining learning models used to automatically map similar texture features to the same type of tissues, and produce a set of rules that can be used for automatic classification and annotation of unlabelled image data. When the classification approach is applied at the local (pixel) level, it can be used for automatic segmentation. Each pixel will receive a label through the classification rules and connected pixels having the same labels will form a region or segment in the corresponding image. This type of segmentation will have a significant impact on the current research efforts for providing *automatic context* (i.e. that the cursor is hovering over “liver” in CT images). The image retrieval step consists of the selection of the best similarity metric and the best texture feature representation in order to retrieve the most similar images with the organ query image. Since there is no similarity measure known to perform the best for the CT modality, we compare eight metrics and three different feature representations, and show how the selection of a similarity metric affects the texture-based retrieval. Furthermore, since our work deals with normal tissues, the system proposed here can be considered as *a first step for supporting the clinical decision-making process*.

### 25.1 Introduction

The human body is an extremely complex system. Images are by far the most efficient way to obtain, interpret, and manage information about complex systems. Physicians increasingly rely on images to understand the human body and to intervene in the processes of human illness and injury. The use of images to manage information about biologic and medical processes is certain to grow, not only in clinical medicine but also in the biomedical imaging research efforts that support it [1].

Advances in biomedical imaging over the years have been driven principally by

the continuous evolution of imaging technologies. The ever-present radiographic film that has been the basis of image management for almost 100 years is being displaced by new digital imaging modalities such as:

- 1) Computed Tomography (CT)
- 2) Magnetic Resonance (MR)
- 3) Nuclear Medicine (NM)
  - a. Emission Computed Tomography (ECT) with compounds releasing positrons (Positron Emission Tomography [PET])
  - b. single photons (Single Photon Emission Computed Tomography [SPECT])
- 4) Ultrasound (US)
- 5) Digital Radiography (DF)
- 6) Computed Radiography (CR) using storage phosphor imaging plates or film digitizers
- 7) digital angiography (DA)
- 8) MR Spectroscopy (MRS)
- 9) Electron Emission Radiography (EMR).

The explosion of the medical imaging technologies has generated mountains of data; depending on the size of the institution, a radiology department can perform between 100 and 5,000 examinations daily, generating a myriad of images, patient data, report text, findings, and recommendations [2]. Digital image management systems are under development now to handle these images in digital form. These systems are termed Picture Archiving and Communication Systems (PACS) and they are based on the integration of different technologies that form a system for image acquisition, storage, transmission, processing and display of images for their analysis and further diagnosis. Availability of digital data within the PACS raises a possibility of health care and research enhancements associated with manipulation, processing and handling of data by computers, that is a basis for computer-assisted radiology development.

In general, radiology data is well organized but poorly structured, and structuring this data prior to knowledge extraction is an essential first step in the successful mining of radiological data [2]. Furthermore, when compared to text, radiology images are enormous in size and highly variable over time. Another challenge is that the image data itself is contained within PACS systems which are in constant use and quite difficult to mine for image data while in use as clinical systems. Therefore, image processing and data mining techniques are necessary for structuring, classification and retrieval of image data. For instance, the development of Content-based Image Retrieval (CBIR) systems and their integration into the PACs systems will have many potential applications in three large domains: education, research, and diagnosis. In the domain of education, a teacher could query for images with specific anatomical regions in order to provide visual similar examples to students. Research can benefit from the CBIR systems: by including visual features directly in medical studies, new correlations between the visual nature of a case and its diagnosis or textual description can be found. Finally, the diagnostics will be the hardest but most important application

for image retrieval. For domains such as evidenced-based medicine or case-based reasoning it is essential to supply relevant, similar cases for comparison based on special visual image features that model the visual detection of a medical doctor.

Besides the technology push, the recent advances in the medical imaging have also been driven by the concept of biologic/clinical pull [1]. For instance, in our days, Computed Tomography imaging is used extensively in radiation oncology as a feedback mechanism to shape and guide external beams of radiation so that cancer treatments are optimized in real time during delivery of the radiation. This is an exceptional development in clinical medicine for designing plans for cancer and guiding the delivery of radiation.

The continuous evolution of imaging techniques, the accelerating advances in computer technology, and the innovations in information networking set the stage for major advances in medical image data mining and its contributions to health care. This chapter presents the potential contributions of data mining and image processing to the field of radiology; in particular, we discuss how computer aided diagnosis (CAD) systems can play a major and important role in early detection, diagnosis, and computerized treatment planning in cancer radiation therapy.

## **25.2 Mining Computed Tomography Data: Classification, Segmentation, and Retrieval**

In this section, we present an overview of our research work and contributions to the process of classification [3], segmentation [4], and retrieval [5] of normal tissues in Computed Tomography studies of the chest and of the abdomen.

### **25.2.1 Image Classification: Related Work**

Tissue classification has been largely limited to specific pathologies, typically within a single organ or tissue. Karkanis et al. [6] proposed a scheme that uses textural descriptors based on second order gray level statistics and employs a multilayer feed forward neural network to discriminate among normal and cancer regions in colonoscopic images. Chabat et al. [7] proposed an automated technique for differentiation between obstructive lung diseases using a supervised Bayesian classifier on the basis of statistical textural descriptors of thin-section CT images. Sluimer et al. [8] also looked at the automatic differentiation of normal from abnormal lung tissue in HRCT of the lungs and found the best classifier to be k-nearest neighbors when using multi-scale filter bank for texture analysis. Fortson et al. [9] applied several Gaussian Maximum Likelihood classifiers to capture the large variety of tissues types inherent to Scleroderma in HRCT of the lung as well. Cios et al. [10] proposed a semi-automatic procedure based on decision trees for analyzing single photon emission computed tomography (SPECT) images of a human heart and classifying the images into several categories. Albrecht et al. presented in [11] a pattern classification method

that combines the classical perceptron algorithm with simulated annealing in order to recognize focal liver tumors in CT images. Wolf et al. [12] studied the hierarchical classification of brain tumors by CT-image Walsh spectra.

Our work will provide texture extraction and classification of normal tissue across a wide range of organs and tissues in Computed Tomography (CT) scans. The texture extraction step will consist of various texture methods applied to the collection of tissue data in order to derive a set of features (numerical descriptors) characterizing the best the visual perception of texture. The classification step will involve different learning models that will automatically map similar texture descriptors to the same type of tissues and thus, produce a dictionary of texture values for normal tissues in CT scans.

The automatic classification of normal tissues will contribute to the development of new approaches for computerized treatment planning in cancer radiation therapy. Before radiation therapy begins, it is important to precisely locate the cancer, but also accurately determine normal areas of the body through which radiation will pass such that the dose of radiation the cancer and nearby normal tissues will be exposed to will be accurately identified. This will enable radiation oncologists to significantly reduce side effects while improving the ability to deliver a curative radiation dose to cancer-containing areas and minimizing the radiation dose to normal tissue.

### **25.2.2 Image Segmentation: Related Work**

The previous approach for normal tissue classification when applied at the local (pixel) level can be used for automatic segmentation. Each pixel will receive a label through the classification rules and connected pixels having the same labels will form a region or segment in the corresponding image. This type of segmentation will have a significant impact on the current research efforts for providing *automatic context* (i.e. that the cursor is hovering over “liver” in CT images). Accurate tissue segmentation and classification of normal CT structures will allow the radiologist to invoke specific tools named *context sensitive tools* during the interpretation process. These tools may represent further image processing and decision support tools or may invoke tissue specific reporting and annotation tools that should speed radiologic reporting while at the same time promote the acquisition of structured information about the images. In the absence of automatic tissue segmentation and classification the navigation of structured reporting hierarchies presents too much of a delay to be practicably implemented.

There are a large number of texture-based segmentation algorithms in the literature. Texture segmentation usually involves the combination of texture feature extraction techniques with a suitable segmentation algorithm. Among the most popular feature extraction techniques used for texture segmentation are Gabor filters and wavelets transforms [13, 14, 15]. Among the most commonly used segmentation algorithms based on these features are clustering techniques [16, 17], region growing and split-and-merge [18, 19].

Segmentation using the traditional techniques previously mentioned requires considerable amounts of expert interactive guidance. In the medical imaging field, deformable models are commonly applied because of their capability to capture the irregular shapes and shape deformations found in anatomical structures. The deformable model that has attracted the most attention to date is popularly known as “snakes” [20] and it has been used for different applications such as the segmentation of the heart from cardiac imagery, of neural tissue textures, and of the bone and cartilage in clinical knee Magnetic Resonance Imaging (MRI). However, the application of snakes and other similar deformable contour models to extract regions of interest is not without limitations. One of the limitations is that snakes were designed as interactive (semi-automatic) models. In order to increase the speed, accuracy, and consistency of the segmentation, automatic segmentation is a desirable, albeit very difficult, long-term goal.

Our proposed approach for segmentation does not require any selection of initial points in order to perform the organ segmentation; moreover, it can be used as an automated first-pass segmentation that can be followed by the snake algorithm.

### 25.2.3 Image Retrieval: Related Work

In medicine to date, virtually all Picture Archiving and Communications Systems (PACS) retrieve images simply by *textual* indices based on patient name, technique, or some-observer-coded text of diagnostic findings [21]. Fields of text tags, such as patient demographics, diagnostic codes (e.g. ICD-9, American College of Radiology diagnostic codes), image view-plane (e.g. sagittal, coronal, etc) and so on usually are the first handles on this process. This textual approach, however, may suffer from considerable observer variability, high cost of manual classification and manipulation of images by medical experts, and failure to fully account for *quantitative relationships of medically relevant structures within an image* that are visible to a trained observer but not codable in conventional database terms.

In the radiology domain, a study [22] was performed to measure the level of inter- and intra-observer agreement and to evaluate the causes of variability in radiologists' descriptions and assessments of sonograms of solid breast masses. The findings of the study showed the lack of uniformity among observers' use of descriptive terms that produced inconsistent diagnoses even though the appearance of masses was described accordingly to a lexicon that was proposed in an earlier benchmark study. Since radiologists themselves rely on visual texture to detect and describe breast lesions on ultrasound images, a Content-Based Image Retrieval (CBIR) system using automatically extracted *quantitative texture features* could have been used to address this established clinical weakness of the diagnostic process and also complement the radiologists' perceptive abilities.

According to Muller et al. [23], only six research projects currently aim at creating CBIR for general medical applications, namely: I<sup>2</sup>C [24], COBRA [25], IRMA [26, 27], KMeD [28], MedGIFT [29], and Image Engine [30]. Most

applications for CBIR focus on the analysis of a specific anatomical structure from images produced in radiology, pathology and cardiology departments. Glatard et al. [31] introduced a CBIR system which uses Gabor filters [32] extracted from segmented cardiac Magnetic Resonance Imaging (MRI) to perform clinically relevant queries on large image databases that do not require user supervision. Mueller et al. [23] compares changes in texture analysis with Gabor filters and the performance of variations in feature space in relation to color (grey level) quantization. Brodley et al. [33] introduced a CBIR system for the retrieval of CT lung images; their proposed system, ASSERT, uses several features (such as co-occurrence texture features, Fourier descriptors and moments), and relies on expert interaction with the system in addition to various machine learning and computer vision techniques. Zheng et al. [34] designed and analyzed a CBIR system for pathology using four types of image features and the dot product as a similarity metric. Wei et al. [35] proposed a CBIR system for the mammography imaging modality using the co-occurrence texture signatures as global features.

In our proposed approach, we use both global-level [36] and local-level co-occurrence texture features to retrieve normal anatomical regions produced by Computed Tomography as the imaging modality. Since there is no similarity measure known to perform the best for the CT modality, we compare eight metrics and three different feature representations, and show how the selection of a similarity metric affects the texture-based retrieval. Furthermore, since our work deals with normal tissues, the system proposed here can be considered as *a first step for supporting the clinical decision-making process*. Same as the normal workflow in medicine, in order to find out if a new case is normal (non-pathological) or not, the new case will be compared with the existing normal cases from a database doing dissimilarity retrieval as opposed to similarity retrieval. The *distance to normality* of the new case along with the knowledge of the medical specialist will determine the placement of the case either in the normal database or in a pathological database; for the later situation, more specialized computer-aided diagnosis tools or image retrieval systems focusing on the specific pathology can be applied further for evaluation and clinical decision making.

## 25.3 Materials and Methods

### 25.3.1 Image Database

Data Acquisition: Our preliminary results are based on data extracted from two normal CT studies from Northwestern Memorial Hospital (NMH) PACS. The data consists of multiple, serial, axial computed tomography images derived from helical, multi-detector CT abdominal and chest acquisitions using a HiSpeed CT/i scanner (GE Medical Systems, Milwaukee, WI); the imaging protocol parameters are: 120 kVp, 60-120 mA (depending on body size), 480 mm for the field-of-view

(FOV) and 0.9375 for the voxel size. The images were taken at the same time for each of the patients, leading to two time points in the data; the patient positioning in the images was FFS (Feet – First - Supine), one of the patient-space coordinate system conventions used in DICOM (Digital Imaging and Communications in Medicine) standard format. The images were transferred via Ethernet to a nearby computer workstation in DICOM format of size 512 by 512 and having 12-bit gray level resolution. An automated software agent (DICOM Valet, ETIAM, Rennes, France) attached to the DICOM Storage Service Class Provider (WinSCP, ETIAM, Rennes, France) performed de-identification according to DICOM Supplement 55.

Image Segmentation for the global tissue classification and retrieval tasks:

Using the Active Contour Models (ACM) algorithm [37], we segmented five organs from 344 2-D axial slices: heart and great vessels, liver, renal and splenic parenchyma, and backbone. We used the ACM segmentation algorithm because it allowed us segment regions with complex shapes, and once several initial points were selected on the boundary, the algorithm calculated automatically the boundary of each of the region of interest.

The main steps involved in our proposed approaches for classification, segmentation, and retrieval are:

- 1) For each segmented region of interest from the image database, we calculate a set of ten Haralick texture features at both global and pixel-level; therefore, each organ or pixel is represented as a vector with ten elements that will be further used for comparing the similarity among the images/organs.

- 2) Once the features are calculated, they can be represented either as a mean-based vector, binned histogram or texture signature depending on the level of granularity considered. Furthermore, a preprocessing step is applied: the features are normalized such that the differences in their scales do not influence the similarity results or the classification results. Then, the normalized texture features are used for the three proposed tasks.

- 3) In the case of the classification task, a decision tree classification model is used to derive a set of classification rules both at the pixel and at the global level. The classification results are evaluated with respect to four performance metrics: sensitivity, specificity, precision, and accuracy. The global rules can be used further for organ tissue classification and annotation while the pixel (local) rules can be used for the CT image segmentation task. Pixels with the same classification labels and being adjacent will form connected components and thus, the regions of interest within the corresponding CT images. In the case of retrieval, eight measures are calculated between the query and all the other images from the database. As a response to a specific query, the system will display the most similar images with the query image. The retrieval performance is evaluated using the precision and recall metrics and each image is considered a query image; therefore, a total of 344 queries are performed and evaluated using the two performance metrics.

### 25.3.2 Texture Features

In medical image processing, texture is especially important, because it is difficult to classify human body organ tissues using shape or gray level information. This is because of the uncertainty introduced by the unlimited variability in organ shape distortion and the potential absolute gray level variability due to the imaging device. While gray levels purely describe point-wise properties of images, texture uses these gray levels to derive some notion of spatial distribution of tonal variations, surface orientation and scenic depth. Furthermore, contrary to the discrimination of morphologic information (shape, size), there is evidence that the human visual system has difficulties in the discrimination of textural information that is related to higher-order statistics or spectral properties in an image. Consequently, texture analysis can potentially augment the visual skills of the radiologist by extracting features that may be relevant to the diagnostic problem but they are not necessary visually extractable [38].

Several approaches have been applied towards the analysis and characterization of texture within medical images including fractal dimension, run-length encoding, discrete wavelet transform, and co-occurrence matrices. While there has not been any conclusive study to prove the superiority of one method over the other methods of capturing texture, we choose to use the Haralick co-occurrence [36] method because it is a well known, established method that has been proven to correlate well with what experts generally look for in texture features. Also, it has been used successfully to produce good results in classification studies of normal tissues in CT images of chest and abdomen [4].

The Haralick co-occurrence texture model and its texture descriptors capture the spatial dependence of gray-level values and texture structures within an image [36]. There are many statistics that can be used; however, due to the redundancy and the high correlation in these statistics, only ten statistics are advocated for feature representation in this application. We are using the following ten descriptors (given by Equations (25.1) through (25.10), where  $P$  is the normalized co-occurrence matrix,  $(i,j)$  is the pair of gray level intensities  $i$  and  $j$ , and  $M$  by  $N$  is the size of the co-occurrence matrix):

$$Entropy = -\sum_i \sum_j P[i, j] \log P[i, j] \quad (25.1)$$

$$Energy = \sum_i \sum_j P^2[i, j] \quad (25.2)$$

$$Contrast = \sum_i \sum_j (i - j)^2 P[i, j] \quad (25.3)$$

$$Homogeneity = \sum_i \sum_j \frac{P[i, j]}{1 + |i - j|} \quad (25.4)$$

$$SumMean = \frac{1}{2} \sum_i \sum_j (i * P[i, j] + j * P[i, j]) \quad (25.5)$$



$$Variance = \frac{1}{2} \sum_i^M \sum_j^N ((i - \mu)^2 P[i, j] + (j - \mu)^2 P[i, j]) \quad (25.6)$$

$$Maximum\_Probability = \underset{i, j}{Max} P[i, j] \quad (25.7)$$

$$Inverse\_Difference\_Moment = \sum_i^M \sum_j^N \frac{P[i, j]}{|i - j|^k} \quad (25.8)$$

$$Cluster\_Tendency = \sum_i^M \sum_j^N (i + j - 2\mu)^k P[i, j] \quad (25.9)$$

$$Correlation = \sum_i^M \sum_j^N \frac{(i - \mu)(j - \mu)P[i, j]}{\sigma^2} \quad (25.10)$$

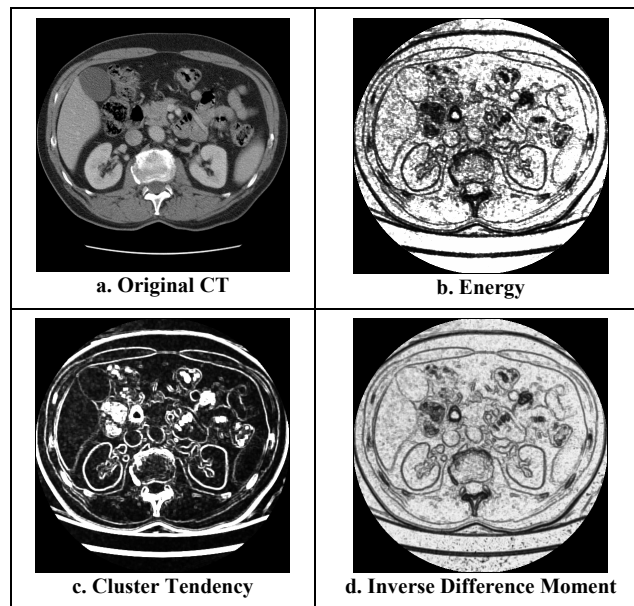
These descriptors are calculated at both local (pixel) and global (organ) level depending on the tasks to be used for and the fundamental structures present in the images. Pixel level properties are calculated to be able to isolate regional properties within an image, while global-level features summarize the whole image and represent it as one entity.

Global-level feature representation: To compute global-level features, the normalized co-occurrence matrices are calculated in four directions ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ ) and five displacements ( $d = 1, 2, 3, 4, 5$ ) generating twenty matrices per segmented image. These rotations and displacements are only in-plane since the images being considered are only 2-dimensional axial slices. The ten Haralick features are calculated for each of the twenty matrices and then, the twenty values are averaged and recorded as a *mean-based feature vector* for the corresponding segmented image [39].

Pixel-level feature representation: To compute pixel-level features, a small neighborhood is considered for each pixel within the segmented region. The size of the neighborhood has to be large enough in order to get enough samples to produce statistically significant texture features at the pixel-level and small enough in order to capture the local property of the texture and not to introduce multiple textures within the pixel neighborhood. Therefore, we choose a neighborhood of size 5 by 5 as a tradeoff between the level of locality and the statistical significance of the results; the choice of this size is also partially motivated by the good classification accuracy obtained for the classification of pixels and regions of soft tissues when using texture features calculated within a 5 by 5 neighborhood [4].

Once the neighborhood size is determined, a co-occurrence matrix is calculated for each neighborhood within the corresponding region. While co-occurrence matrices are normally defined for a fixed distance and direction when calculated at the global level, for the pixel-level approach, we do not calculate the co-occurrence along fixed directions and displacements. Instead we consider all pixel pairs within that neighborhood such that there will be enough samples (pairs) for calculating the co-occurrence matrix in order to produce statistically

significant results. Thus, our implementation produces a single co-occurrence matrix for each pixel rather than for each choice of distance and direction. Then, for each co-occurrence matrix (each pixel), we calculate ten Haralick features which can be related to specific characteristics in the image. Figure 25.1 (b-d) illustrates the image representations of different pixel-level texture features for the original CT image from Figure 25.1 (a).



**Fig. 25.1:** Image Representation for the pixel-level texture features

From the pixel-level data, we derive different representations for the texture co-occurrence features: 1) *mean vector-based data*, 2) *binned histogram data* and 3) *texture signatures*. These vectors are the possible representations of the texture features at the pixel-level and they will be evaluated in order to study the effect of the feature space on the choice of the similarity metric and thus, on the retrieval results.

The *mean vector-based data representation* consists of the average of the normalized pixel-level data for each region such that the texture representation of that corresponding region is a vector instead of a set of vectors given by the pixels' vector representations within that region.

The *binned histogram data representation* consists of texture values grouped within equal-width bins. The number of bins and their placement are important parameters as they determine how crudely, or how well, the underlying probability distribution (obtained by quantizing the responses into bins and normalizing such that the sum over all bins is unity) is: too many number of bins will overfit the data and introduce noise while less number of bins will make the

binning crude. In our experimental results, a number of 256 equal-size bins produced the best results for this representation.

The *texture signature representation* is the clustered representation of the normalized local level data obtained using a k-d tree clustering algorithm [40]. The k-d tree clustering algorithm is chosen because 1) it does not require any assumptions about the data set; 2) it is computational efficient; and 3) it allows clusters of unequal size and thus, it will eliminate the limitation of the binned histogram representation. The k-d tree algorithm iteratively divides the data space using predefined stopping criteria. In our approach, we implement two stopping criteria: the first criterion was to establish a minimum variance within the subset to be divided to prevent creating redundant clusters and over-splitting; the second stopping criterion is used to enforce a minimum cluster size as a percentage of the original data set and to maintain a significant size within the clusters and to prevent outliers from uncontrollably growing the tree. By varying both the variance and minimum cluster size different texture signatures are obtained. In our experimental results obtained using the directed Hausdorff distance, a variance equal to 10% and a cluster size equal to 20% of the parent cluster's variance and size, respectively, produced the best retrieval results for this representation.

### 25.3.3 Classification Model

Decision tree classifier: There are many classifiers that can be used to discriminate among the organ tissue classes in the feature space. In our preliminary work, we evaluated a decision tree classifier because: 1) it does not make any assumptions of the distribution of the data; 2) it has a relatively faster learning speed than other classification methods, while still producing classification accuracy comparable with those methods; and 3) it has a good ability to generate decision rules that can be easily understood, interpreted, and used to annotate different tissues in future CT scans. The implementation of our decision tree was based on the Classification and Regression Trees (C&RT) approach and used the SPSS Answer Tree 3.0 software. From the decision tree, a set of the most important decision rules were generated to be used for classification of the regions, and to derive the most relevant texture descriptors for specific organs. To evaluate the performance of the classifier, we calculated four metrics on the regions of interest in the testing set: sensitivity, specificity, precision, and accuracy.

The C&RT tree is constructed by splitting subsets of the data set using all descriptors as predictors to create two child nodes repeatedly, beginning with the entire data set. The best predictor is chosen using the Gini impurity index which works by choosing a split at each node such that each child node is more pure than its parent node:

$Gini(S) = 1 - \sum_{i=1}^c p_i^2$ , where S is the data set to be split, c is the number of classes and  $p_i$  is the probability of class i within the data set S. A total pure node

is a node for which the Gini index is equal to zero. The goal is to produce subsets of the data which are as homogeneous as possible (producing pure nodes in the tree) with respect to the class label. For each split, each predictive descriptor is evaluated to find the best cut point (our descriptors being continuous predictors) based on improvement score or reduction in impurity. Then, the predictors are compared, and the predictor with the best improvement is selected for the split. The process repeats recursively until one of the stopping rules is triggered: 1) the maximum tree depth,  $d$ , has been reached; 2) there is no significant predictive descriptor left to split the node; 3) the number of cases in the terminal node is less than the minimum number,  $np$ , of cases for parent nodes; 4) if the terminal node were to split, the number of cases in one or more child nodes would be less than the minimum number,  $nc$ , of cases for child nodes; and 5) minimum change in impurity,  $imp$ , is reached. Depending on the values set for the parameters ( $d$ ,  $np$ ,  $nc$ ,  $imp$ ), a different tree will be obtained; the ‘best’ tree will be chosen to be the one with the highest classification accuracy.

Evaluation metrics: In order to select the ‘best’ parent and ‘best’ child, and thus the ‘best’ decision tree for our data, the following four performance metrics have to be maximized: 1) sensitivity (the ratio between true positives and total positives), 2) specificity (the ratio between true negatives and total negatives), 3) precision (the ratio between true positives, and the summation of true positives and false positives), and 4) accuracy (the ratio between the summation of the true positives and negatives and the total number of samples). For example, if we are interested in measuring the classification performance for ‘heart and great vessels’ class, a true positive is a tissue region classified as ‘heart and great vessels’ when the original class label (the label given by a human expert) is ‘heart and great vessels’; a true negative is a tissue region correctly classified as ‘non ‘heart and great vessels’’, a false positive is a tissue region classified as ‘‘heart and great vessels’ when it is actually a ‘non-‘heart and great vessels’’, total positives is the total number of ‘heart and great vessels’, and total negatives is the total number of ‘non-heart and great vessels’. The same definitions apply for the other tissue types.

Decision rules: The decision tree can be generated at both local and global feature level. At the local level, its decision rules can be used to classify each pixel within the CT images. Pixels with the same classification labels and being adjacent will form connected components and thus, produce the segmentation of the regions of interest within the corresponding CT images.

Once the optimal decision tree has been constructed, it is a simple matter to convert it into an equivalent set of decision rules. Converting a decision tree to rules has two main advantages: 1) converting to rules removes the distinction between texture features that occur near the root of the tree and those that occur near the leaves, and 2) converting to rules improves readability since rules are often easier for people to understand.

### 25.3.4 Similarity Measures and Performance Evaluation

Similarity metrics definitions: Similarity metrics describe how similar two images (organs in our case) are. There are many similarity measures proposed in the context of CBIR and the choice of a similarity metric is dependent on both the feature space representation and its property to capture the visual human perception of similarity. Rubner et al. defines four categories of similarity measures to calculate the similarity for histogram-based data [41]:

*Heuristic distance metrics:* 1) Minkowski 1-distance,  $d_{L_r}$  (city block distance or  $L_1$  norm) (Equation (25.11)), 2) weighted-mean-variance,  $d_{wmv}$  (uses the means and standard deviations for each of the considered features) (Equation (25.12)):

$$d_{L_r}(H, K) = \left( \sum_i |h_i - k_i|^r \right)^{\frac{1}{r}} \quad (25.11)$$

$$d_{wmv}(H, K) = \sum_i \frac{|\mu_i(H) - \mu_i(K)|}{|\sigma(\mu_i)|} + \frac{|\sigma_i(H) - \sigma_i(K)|}{|\sigma(\sigma_i)|} \quad (25.12)$$

*Non-parametric test statistics:* 1) Cramer-von Mises,  $d_{CvM}$  (similar to the squared Euclidean distance but calculated between the distributions and as the maximal discrepancy between the cumulative distributions) (Equation (25.13)); 2) Kolmogorov-Smirnov distance,  $d_{KS}$  (used for unbinned data distributions and it is invariant to arbitrary monotonic transformations) (Equation (25.14)), and 3) Chi-square statistics,  $d_{\chi^2}$  (used to distinguish whether distributions of the descriptors differ from each other) (Equation (25.15)):

$$d_{CvM}(H, K) = \sum_i \sum_j \left( F^i(j; H) - F^i(j; K) \right)^2 \quad (25.13)$$

$$d_{KS}(H, K) = \sum_i \max_j \left( \left| F^i(j; H) - F^i(j; K) \right| \right) \quad (25.14)$$

$$d_{\chi^2}(H, K) = \sum_i \frac{(h_i - m_i)^2}{m_i}, m_i = \frac{h_i + k_i}{2} \quad (25.15)$$

*Information Theory Divergences:* 1) Jeffrey-Divergence,  $d_{JD}$  (used to compute the distance between class distributions of two values of the same feature) (Equation (25.16)); and 2) Kullback-Leibler (KL) divergence,  $d_{KL}$  (Equation (25.17)):

$$d_{JD}(H, K) = \sum_i \sum_j \left( f(j; H) \log \frac{f(j; H)}{m_j} + f(j; K) \log \frac{f(j; K)}{m_j} \right) \quad (25.16)$$

$$\text{where } m_j = \frac{f(j; H) + f(j; K)}{2}$$

$$d_{KL}(H, K) = \sum_i f(j; H) \log \frac{f(j; H)}{f(j; K)} \quad (25.17)$$

Ground distances: 1) Quadratic Form (QF),  $d_{QF}$ , (Equation (25.18)), and 2) Earth Mover's Distance (EMD),  $d_{EMD}$ : (Equation (25.19))

$$d_{QF}(H, K) = \sqrt{(f_H - f_K)^T A (f_H - f_K)} \quad (25.18)$$

$$d_{EMD}(H, K) = \frac{\sum_{i,j} g_{ij} d_{ij}}{\sum_{i,j} g_{ij}} \quad (25.19)$$

H represents the query image that can be thought as of a point in a 10-dimensional space, where the value along each dimension (each of the 10 Haralick texture features) is given by  $h_i$ ; similarly, K represents a given image from the database and  $k_i$  represents its feature value corresponding to the  $i^{\text{th}}$  dimension in the feature space. Furthermore,  $f_H$  and  $f_K$  are vectors that list all entries in  $f(i;H)$  and  $f(i;K)$ , A denotes the similarity matrix,  $g_{ij}$  is the optimal flow between two distributions and  $d_{ij}$  is the similarity between bin  $i$  and  $j$ .

We evaluate all similarity measures from the first and second category and the Jeffrey-Divergence from the third category within the context of both feature representation requirements for each of the metrics and their retrieval performance. Jeffrey-Divergence ( $d_{JD}$ ) is an information theoretically motivated similarity measure just like Kullback–Leibler [41] and Mutual Information [42]. The latter two are not implemented in this work, but the former is implemented and will serve to represent the performance of similarity measures of this class. Since the two ground distances have high computational complexity and we are interested in evaluating CBIR systems for the medical domain where the retrieval of similar images should be performed very fast and on-the-fly, we do not consider them in the current implementation. In addition to the above metrics, two others are implemented as required by the different texture feature representations: Euclidean distance  $d_A$  (Equation (25.20)) and Hausdorff distance,  $d_{HD}$  (used for texture signature representation) (Equation (25.21)):

$$d_A(H, K) = \sum_i \sqrt{(h_i - k_i)^2} \quad (25.20)$$

$$d_{HD}(H, K) = \max_{h \in H} (\min_{k \in K} (\|h - k\|)) \quad (25.21)$$

For more details on the properties of these similarity metrics we refer the reader to Rubner et al. work [41].

**Performance metrics:** For medical image retrieval systems, the evaluation issue is almost non-existent in most of the papers and those systems that do perform evaluation often use only screenshots of example results to queries. A single example result does not reveal a great deal about the real performance of the system and is not objective as the best possible query can be chosen arbitrarily by the authors.

We evaluate the system's retrieval results using precision and recall as performance metrics, as defined by Equation (25.22) and Equation (25.23):

$$\text{Precision} = \frac{\#of\_relevant\_retrieved\_images}{total\#of\_retrieved\_images} \quad (25.22)$$

$$\text{Recall} = \frac{\#of\_relevant\_retrieved\_images}{total\_\#of\_relevant\_images} \quad (25.23)$$

Each image is considered a query image; therefore, a total of 344 queries are performed and evaluated using the two performance metrics. The  $k$  number of images that are the most similar with the query image is another parameter to be considered in evaluating the two metrics with respect to retrieval.

As a note, the current retrieval system is not required to rank the retrieved organs based on the variability that exists within organs of the same anatomical regions. Therefore, a retrieved image is 'relevant' if belongs to the same anatomical region as the query. However, for the application of this type of system in a more specific discriminatory manner (e.g. to track disease progression, organ distortion, size), it would be useful to rank retrieved images from the same anatomical region.

## 25.4 Experimental Results and their Interpretation

### 25.4.1 Tissue Classification Results

In order to produce and validate the best decision tree, the 344 segmented images were divided into 4 quadrants producing 1360 images (some quadrants did not contain any tissue so they were ignored) that were used for the training set (66% of the data) and the testing set (34%). The training set was used to create the decision trees based on a 10-fold cross-validation technique. Each of the produced decision trees were then validated on the testing data and the decision tree for which the four performance metrics were maximum was chosen as the optimal tree. The optimal tree resulted in a tree with 41 nodes, 8 levels of depth and 21 leaves producing 21 decision rules for the classification process.

For the training set, the overall performance (calculated as the weighted average per organ) for all four metrics was better than 89%. For the testing set, the overall performance for sensitivity and precision was above 80% and the performance for specificity and accuracy was above 90%.

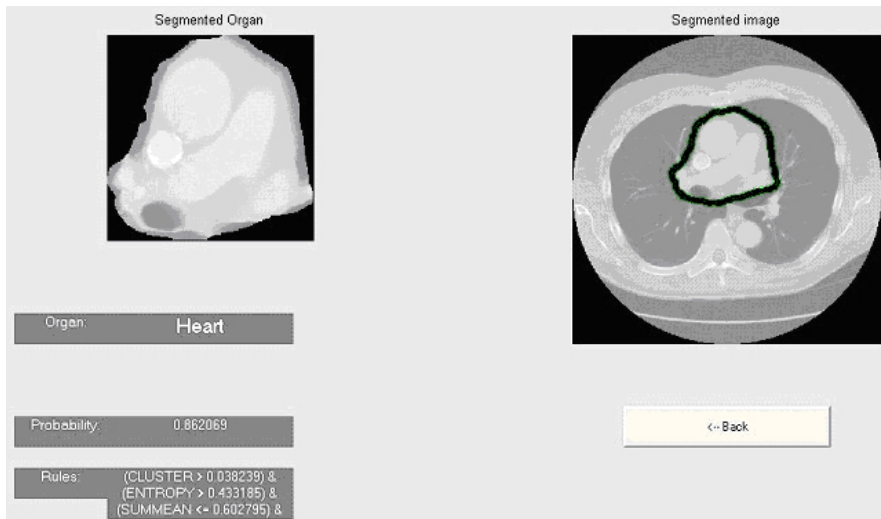
Table 25.1 and Table 25.2 show the four performance metrics for the training and testing data per organ, respectively; Figure 25.2 shows how the rules are applied in order to annotate unlabelled tissues in CT images.

**Table 25.1:** Classification performance on individual tissues of the training set (number of parents = 28, number of children = 5, cross-validation fold =10)

ORGAN	Sensitivity	Specificity	Precision	Accuracy
Backbone	99.7%	99.5%	99.2%	99.6%
Liver	80.0%	96.9%	83.8%	94.1%
Heart	84.6%	98.5%	90.6%	96.5%
Renal	92.7%	97.9%	89.7%	97.1%
Splenic parenchyma	79.5%	96.1%	73.6%	94.1%

**Table 25.2:** Classification performance on individual tissues of the testing set (number of parents = 28, number of children = 5, cross-validation fold =10)

ORGAN	Sensitivity	Specificity	Precision	Accuracy
Backbone	100%	97.6%	96.8%	98.6%
Liver	73.8%	95.9%	76.2%	92.5%
Heart	73.6%	97.2%	84.1%	93.2%
Renal	86.2%	97.8%	87.5%	96.0%
Splenic parenchyma	70.5%	95.1%	62.0%	92.5%



**Fig. 25.2:** An example of heart classification in our proposed approach with a probability of 0.862

The lowest sensitivity and precision values were recorded for spleen which was misclassified as liver most of the time; this indicates that either the used texture



descriptors did not have enough discrimination power or the decision tree classifier was not able to produce the optimal tree with respect to these two types of tissues. Another possible explanation for the misclassification of the liver and spleen comes from the similarity of the gray-levels of these tissues introduced by the linear binning. It is worth mentioning here that low results for the classification of liver and spleen were obtained by Koss et al. [45] when they applied a neural network to segment a CT image into 7 and 8 classes (tissues) of interest using a pixel-based co-occurrence texture model. Therefore, as future work, we plan to investigate other binning strategies, incorporate additional texture models into our texture analysis, and apply other classification techniques in addition to the decision tree one.

The classifier model obtained using decision tree approach generated a set of 21 rules: three rules for the heart, three rules for the kidneys, five rules for the spleen, eight rules for the liver, and two rules for the backbone. The fact that there are multiple rules to describe a single organ suggests that single classes (organs) may have multiple sub-classes (tissues).

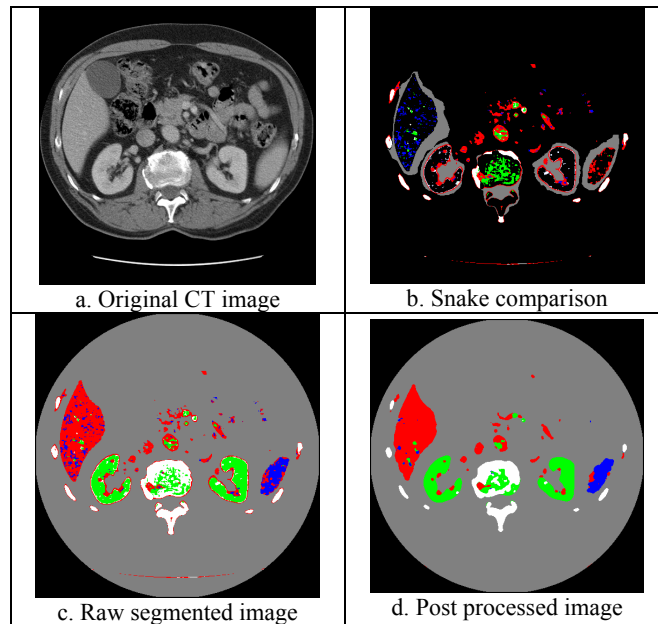
#### **25.4.2 Tissue Segmentation Results**

In order to generate the decision tree for the pixel data, we manually selected patches of pure organ tissues from three consecutive slices. The number of patches and their sizes were chosen such that we have an equal number of pixels for each organ of interest. The pixels received the class label of the patch to which they belonged to. Since we want to use the decision tree for the segmentation of entire CT images, additional patches (not containing the organs of interest) were selected and their pixels were labeled as 'unknown'. We ended up selecting around 1500 pixels for each of the four organs and the 'unknown' class. Furthermore, the training set was used to build the classifier, while the second set was used to estimate the accuracy of the classifier when used for tissue/organ annotation of previously unseen pixels.

In order to select the optimal decision tree (DT) for our data sample, we varied the number of observations (pixels) per node from 25 (number of pixels in a neighborhood) to 1000 and each time we estimated the overall accuracy of the classifier (number of pixels correctly classified divided by the total number of pixels); based on the accuracy of the testing set, the optimal tree was selected. The empirically found optimal parameter for the "observations per parent" was in the range from 274 to 289; any of those values would result in the combined accuracy of the testing set of over 85%. We decided to use a tree with the "observations per parent" of 289 since it resulted in the smallest and most efficient tree.

After we generated and tested the decision tree on the sample data, we applied the tree on several consecutive entire slices in order to segment the organs of interest. We noticed that the unknown class, the kidneys, and the bones were accurately classified while spleen and liver were very often misclassified (liver pixels were classified as spleen and vice-versa). In order to improve the segmentation, either a median filter or spatial information (such as liver is always

in the anatomical left hand side of the abdomen, and the backbone can be used as a point of reference to find the orientation of the CT scan) can be used as a post processing step. For visualization purposes, Figure 25.3 (c-d) shows the pixel-level classification image before and after post processing with a median filter; different colors represent the organs of interest (red – liver, green – kidney, white – backbone, blue – spleen), gray represents the unknown class (organs which were not of interest for the current study) and black is the region outside the body which has not been included in the analysis.



**Fig. 25.3:** Visual representation of classification image, snake comparison, and median filtered image (5 by 5 filter)

### 25.4.3 Tissue Retrieval Results

Since several similarity measures and different feature sets are proposed at both pixel-level and global-level data, we will determine and evaluate *the best combination of the texture feature representation and corresponding similarity measure for retrieval of medical images containing specific anatomical regions*. The best result for pixel-level data approaches would also be compared with the best result from global-level data.

In order to evaluate the significance of the retrieval results, all the 344 images were used as query images. All eleven combinations of feature sets and similarity measures at both levels gave an overall precision over 80% for the number ( $k$ ) of most similar images retrieved equal to 6 (Table 25.3).

**Table 25.3:** Precision at the global and local-level for the entire image database; the overall performance is the weighted average of the retrieval performance by organ (each image was a query image)

GLOBAL LEVEL VECTOR BASED PRECISION						
	Backbone	Heart	Kidney	Liver	Spleen	OVERALL
Euclid Distance	100.0%	90.4%	93.8%	67.8%	62.1%	87.7%
Chi Square Statistics	100.0%	90.7%	93.8%	62.9%	57.5%	86.4%
Minkowski 1 Distance	100.0%	90.1%	92.9%	69.0%	62.5%	87.8%
PIXEL LEVEL VECTOR BASED PRECISION						
	Backbone	Heart	Kidney	Liver	Spleen	OVERALL
Euclid Distance	100.0%	76.0%	85.8%	59.8%	46.7%	81.2%
Chi Square Statistics	100.0%	81.1%	87.7%	60.1%	47.5%	82.4%
Minkowski 1 Distance	100.0%	74.4%	85.2%	59.5%	48.8%	81.0%
Weighted-Mean-Variance	100.0%	87.2%	91.7%	58.9%	53.8%	84.5%
PIXEL LEVEL BINNED HISTOGRAM BASED PRECISION						
	Backbone	Heart	Kidney	Liver	Spleen	OVERALL
Cramer/von Mises	100.0%	88.8%	83.6%	64.1%	51.3%	84.0%
Jeffrey-Divergence	100.0%	91.7%	96.0%	77.9%	75.8%	91.6%
Kolmogorov-Smirnov Distance	100.0%	89.1%	89.8%	69.8%	60.0%	87.0%
PIXEL LEVEL SIGNATURE BASED PRECISION						
	Backbone	Heart	Kidney	Liver	Spleen	OVERALL
Hausdorff 10% v 20% cs	100.0%	81.1%	86.4%	57.8%	42.1%	81.2%

At the global level, there was not much difference in the overall accuracy among the three similarity metrics considered, but the Minkowski and Euclidean distance performed better for liver and spleen than the Chi-square statistics metric. At the pixel level, the retrieval precision was in general higher for the systems that used binned-histogram data together with the Cramer/von Mises, Jeffrey-Divergence and the Kolmogorov-Smirnov metrics. The combination of the binned-histogram feature set and the Jeffrey-divergence metric reached a value of 91.57% making this approach to outperform all the other approaches. Pure comparison of the similarity metrics with respect to the granularity of the feature data can only be made between the metrics that were applied to the data represented in the same feature space. Comparing the Euclidean, Chi Square, and the Minkowski metrics in [5], the global features overall perform better by about 6%. Even though the overall performance is better when using global-level descriptors for these three metrics, one of the systems considered in the pixel level approaches outperforms the global-level approaches by up to as much as 10% to 20% for liver and spleen.

Furthermore, comparing the best feature set and similarity metric combination per organ at the pixel level, we notice that the binned-histogram feature set with

Jeffrey divergence performs the best with respect to each individual organ: backbone (100%), heart (89.7%), kidneys (96%), liver (77.87%) and spleen (75.83%).

Since the best retrieval results were obtained at the pixel level and for the binned histogram feature set and the Jeffrey-Divergence metric, we evaluated further the performance of this system when more than six similar retrieved images were retrieved for this combination of feature representation and similarity metric. By increasing  $k$ , the number of the most similar images retrieved, from 6 to 156 (increments of 10) and calculating the values for precision and recall, we noticed that even for more than six similar images (up to 26 most similar images) the retrieval precision continues to be above 87.79% (the best value obtained for global retrieval and  $k=6$ ). Evaluating the overall recall, its value increased from 8% for  $k=6$  to above 80% for  $k=156$ .

## 25.5 Conclusions

In conclusion, this chapter presents the potential contributions of data mining and image processing to the field of radiology; in particular, we discussed how computer-aided diagnosis (CAD) systems can play a major and important role in early detection, diagnosis, and treatment planning.

While there has been considerable work done for classification of abnormal tissues within different organs (such as liver, lung, heart and brain), to our best knowledge, there is little research in regards to inter-organ classification. Our preliminary results for classification and segmentation show that using only 10 texture descriptors calculated from Hounsfield unit data, it is possible to automatically segment and classify regions of interest representing different organs or tissues in CT images. Furthermore, the results lead us to the conclusion that the incorporation of some other texture models into our proposed approach will increase the performance of the classifier, and will also extend the classification and segmentation functionality to other organs.

The research work on retrieval presents an extensive evaluation of different co-occurrence texture feature representations and similarity metrics for content-based medical image retrieval. Our experimental results show that the presented approaches are promising to offer new possibilities for content-based access to medical images; new intelligent systems could be created that will be able to choose image features relevant to a specific clinical task, analyze the data, and automatically choose the best similarity metric for corresponding data representation.

Even though there are many challenges that need to be addressed before the approaches presented in this chapter will become a true partner to the diagnostic radiologists, the proposed approaches can be considered as an initial step along these efforts and can open other avenues of exploration for other researchers in the field.

## References

1. Hendeer WR and Ritonour ER. *Medical Imaging Physics*. Elsevier Academic Press, 2004.
2. Dreyer KJ. *The Alchemy of Data Mining, Imaging Economics*, 2005
3. Xu D, Lee J, Raicu DS, Furst JD, Channin DS. Texture Classification of Normal Tissues in Computed Tomography, The 2005 Annual Meeting of the Society for Computer Applications in Radiology, 2005.
4. Kalinin M, Raicu DS, Furst JD, Channin DS. A Classification Approach for Anatomical Regions Segmentation, The IEEE International Conference on Image Processing (ICIP), 2005.
5. Corboy A, Tsang W, Raicu DS, Furst J. Texture-Based Image Retrieval for Computerized Tomography Databases, The 18th IEEE International Symposium on Computer-Based Medical Systems(CBMS'05), 2005.
6. Karkanis SA, Magoulas GD, Grigoriadou M, and Schurr M. Detecting abnormalities in colonoscopic images by textural descriptors and neural networks. *Proceedings of the Workshop Machine Learning in Medical Applications*, 1999; 59-62.
7. Chabat F, Yang GZ, and Hansell DM. Obstructive Lung Diseases: Texture Classification for differentiation at CT, RSNA 2003.
8. Sluimer IC, van Waes PF, Viergever MA, van Ginneken B. Computer-aided diagnosis in high resolution CT of the lungs. *Medical Physics*, 2003; 30 (12).
9. Fortson R, Lynch D, and Newell J. Automated Segmentation of Scleroderma in HR CT Imagery. Report LA-UR-95-2401, 1995.
10. Cios KJ, Goodenday LS, Shah KK, and Serpen G. Novel Algorithm for Classification of SPECT Images of a Human Heart. *IEEE Content-Based Medical Systems*, 1996.
11. Albrecht A, Loomes MJ, Steinhöfel K, and Taupitz M. Adaptive Simulated Annealing For CT Image Classification. *Proceedings of IJPRAI*, 2002; 16(5).
12. Wolf M, Ziegenggeist S, Michalik M, Bornholdt F, Michalik S, Meffert B. Classification of brain tumors by CT-image Walsh spectra. *Neuroradiology*, 1990; 32(6).
13. Jain AK and Farrokhia F. Unsupervised texture segmentation using Gabor filters, *Pattern Recognition*, 1991; 24: 1167—1186.
14. Chang T and Kuo CC. Texture segmentation with tree-structured wavelet transform, *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, 1992; 543–546.
15. Unser M. Texture classification and segmentation using wavelet frames, *IEEE Trans. on Im. Proc.*, 1995; 4(11): 1549–1560.
16. Chang KI, Bowyer KW, and Sivagurunath M. Evaluation of texture segmentation algorithms, *IEEE Conference on Computer Vision and Pattern Recognition*, 1999; 294–299.
17. Porter R and Canagarajah N. A robust automatic clustering scheme for image segmentation using wavelets, *IEEE Transactions on Image Processing*, 1996; 5(4): 662–665.

18. Beveridge JR, Gri J, Kohler RR, Hanson AR, and Riseman EM. Segmenting images using localized histograms and region merging. *International Journal Computer Vision*, 1989; 2: 311-347.
19. Adams R and Bischof L. Seeded region growing. *IEEE Transactions Pattern Analysis and Machine Intelligence.*, 1994; 16(6): 641-647.
20. Kass M, Witkin A and Terzopoulos D. Snakes: Active contour models, *International Journal of Computer Vision*, 1988; 1: 321-331.
21. Tagare DH, Jaffe CC, & Duncan J. Medical Image Databases: A Content-based retrieval approach. *Journal of American Medical Informatics Association*, 1997; 4 (3): 184 - 198.
22. Baker JA, Kornguth PJ, Soo MS, Walsh R, Mengoni P. Sonography of solid breast lesions: observer variability of lesion description and assessment. *AJR* 1999; 172: 1621-1625.
23. Müller H, Rosset A, Vallée JP, Geissbuhler A. Comparing feature sets for content-based medical information retrieval. *SPIE Medical Imaging*, 2004.
24. Orphanoudakis SC, Chronaki CE, Kostomanolakis S. I2Cnet: A system for the indexing, storage and retrieval of medical images by content, *Medical Informatics*, 1994; 4(3): 109-122.
25. El-Kwae EA, Xu H, Kabuka MR. Content-based retrieval in picture archiving and communication systems. *Journal of Digital Imaging*, 2000; 13 (2): 70-81.
26. Guld MO, Wein BB, Keysers D, Thies C, Kohnen M, Schubert H, Lehmann TM. A distributed architecture for content-based image retrieval in medical applications, in: *Proceedings of the International Conference on Enterprise Information Systems (ICEIS2001)*, 2001; 299-314.
27. Lehmann T, Wein B, Dahmen J, Bredno J, Vogelsang F, and Kohnen M. Content-based Image Retrieval in Medical Applications: A Novel Multi-step Approach. In *Procs. Int. Society for Optical Engineering (SPIE)*, 2000; 3972(32): 312-331.
28. Chu WW, Cardenas AF, Taira RK. KMED: A knowledge-based multimedia distributed database system, *Information Systems* 19 (4) (1994) 33-54.
29. Müller H, Fabry P, Geissbuhler A. MedGIFT - Retrieving medical images by there visual content. *World Summit of the Information Society, Forum Science and Society*, 2003.
30. Image Engine: <http://www.bristol.ac.uk/radiology/IntLink/ImageEngine.html>
31. Glatard T, Montagnat J, and Magnin IE. Texture based medical image indexing and retrieval: application to cardiac imaging. *ACM SIGMM international workshop on Multimedia Information Retrieval (MIR'04)*, *Proceedings of ACM Multimedia 2004*.
32. Rubner Y and Tomasi C. Texture Metrics. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, 1998; 4601-4607.
33. Brodley C, Kak A, Shyu C, Dy J, Broderick L, and Aisen AM. Content-Based Retrieval from Medical Image Databases: A Synergy of Human Interaction, Machine Learning and Computer Vision. In *Proc. of the Sixteenth National Conference on Artificial Intelligence (AAAI'99)*, 1999.
34. Zheng L, Wetzel AW, Gilbertson J, Becich MJ. Design and Analysis of Content-based Pathology Image Retrieval System. *IEEE Transactions on Information Technology in Biomedicine*, 2003; 7 (4).
35. Wei CH, Li CT and Wilson R. A General Framework for Content-Based Medical Image Retrieval with its Application to Mammograms. in *Proc. SPIE Int'l Symposium on Medical Imaging*, 2005.

36. Haralick RM, Shanmugam K, and Dinstein I. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 1973; Smc-3 (6): 610-621.
37. Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. *International Journal of Computer Vision*, 1988; 1(4).
38. Tourassi GD. Journey toward Computer-Aided diagnosis role of image texture analysis. *Radiology*, 1999; 213: 317-320.
39. Raicu DS, Furst JD, Channin DS, Xu DH, & Kurani A. A Texture Dictionary for Human Organs Tissues' Classification. *Proceed. of the 8th World Multiconf. on Syst., Cyber. and Inform.*, 2004.
40. Bentley JL. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 1975; 18: 509-517.
41. Rubner Y, Puzicha J, Tomasi C, and Buhmann JM. Empirical Evaluation of Dissimilarity Measures for Color and Texture. *International Conference on Computer Vision*, 1999; 2: 1165.
42. Pluim JPW, Maintz JBA, and Viergever MA. Mutual-information-based registration of medical images: A survey. *IEEE Transactions on Medical Imaging*, 2003; 22(8): 986-1004
43. Rubner Y, Tomasi C, and Guibas L. The Earth Mover's Distance as a metric for image retrieval. *Technical Report STAN-CS-TN-98-86*, Computer Science Department, Stanford University, 1998.
44. Wei G, Li D and Sethi IK. Detection of Side-view Faces in Color Images. *Proceedings Fifth IEEE Workshop on Applications of Computer Vision*, 2000; 79-84.
45. Koss JE, Newman FD, Johnson TK, Kirch DL. Abdominal organ segmentation using texture transforms and a Hopfield neural network, *IEEE Transactions on Medical Imaging*. 1999; 18.