Chapter IX

A Content-Based Approach to Medical Image Database Retrieval

Chia-Hung Wei, University of Warwick, UK
Chang-Tsun Li, University of Warwick, UK
Roland Wilson, University of Warwick, UK

Abstract

Content-based image retrieval (CBIR) makes use of image features, such as color and texture, to index images with minimal human intervention. Content-based image retrieval can be used to locate medical images in large databases. This chapter introduces a content-based approach to medical image retrieval. Fundamentals of the key components of content-based image retrieval systems are introduced first to give an overview of this area. A case study, which describes the methodology of a CBIR system for retrieving digital mammogram database, is then presented. This chapter is intended to disseminate the knowledge of the CBIR approach to the applications of medical image management and to attract greater interest from various research communities to rapidly advance research in this field.

Introduction

John Doe, a radiologist in a university hospital, takes X-rays and MRI scans for patients producing hundreds of digital images each day. In order to facilitate easy access in the future, he registers each image in a medical image database based on the modality, region, and orientation of the image. One day Alice Smith, a surgeon, comes to discuss a case with John Doe as she suspects there is a tumor on the patient's brain according to the brain MRI. However, she cannot easily judge if it is a benign or malign tumor from the MRI scan, and would like to compare with previous cases to decide if this patient requires a dangerous operation. Understanding Alice's needs, John helps Alice find similar-looking tumors from the previous MRI images. He uses the query-by-example mode of the medical image database, delineates the tumor area in the MRI image, and then requests the database to return the brain MRI images most similar to this one. Alice finds eleven similar images and their accompanying reports after reviewing the search results. Alice compares those cases and verifies the pattern of the tumor. Later on, she tells her patient that it is a benign tumor and the operation is unnecessary unless the tumor grows.

This scenario briefly describes the creation of medical images, categorization of medical images, and a content-based access approach. Although a mature content-based access technology has not appeared yet, this field is developing actively. In the last decade, a large number of digital medical images have been produced in hospitals. Large-scale image databases collect various images, including X-ray, computed tomography (CT), magnetic resonance imaging (MRI), ultrasound (US), nuclear medical imaging, endoscopy, microscopy, and scanning laser ophtalmoscopy (SLO). The most important aspect of image database management is how to effectively retrieve the desired images using a description of image content. This approach of searching images is known as content-based image retrieval (CBIR), which refers to the retrieval of images from a database using information directly derived from the content of images themselves, rather than from accompanying text or annotation (El-Naqa, Yang, Galatsanos, Nishikawa, & Wernick , 2004; Wei & Li, in press).

The main purpose of this chapter is to disseminate the knowledge of the CBIR approach to the applications of medical image retrieval and to attract greater interest from various research communities to rapidly advance research in this field. The rest of the chapter is organized as follows: The second section addresses the problems and challenges of medical image retrieval and describes potential applications of medical CBIR. The third section reviews the

existing medical CBIR systems. The fourth section provides greater details on the key components of content-based image retrieval systems for medical imaging applications. The fifth section presents a case study, which describes the methodology of CBIR systems for digital mammograms. The sixth section discusses potential research issues in the future research agenda. The last section concludes this chapter.

Medical Image Database Retrieval

This section will discuss the problems of image retrieval using the conventional text-based method and addresses the challenges of the CBIR approach. Potential applications of the CBIR approach will also be discussed.

Challenges in Medical Image Retrieval

Before the emergence of content-based retrieval, medical images were annotated with text, allowing the images to be accessed by text-based searching (Feng, Siu, & Zhang, 2003). Through textual description, medical images can be managed based on the classification of imaging modalities, regions, and orientation. This hierarchical structure allows users to easily navigate and browse the database. Searching is mainly carried out through standard Boolean queries.

However, with the emergence of massive image databases, the traditional text-based search suffers from the following limitations (Shah et al., 2004; Wei & Li, 2005):

- Manual annotations require too much time and are expensive to implement. As the number of images in a database grows, the difficulty in finding desired images increases. Muller, Michous, Bandon, and Geissbuhler (2004a) reported that the University Hospital of Geneva produced approximately 12,000 medical images per day. It is not feasible to manually annotate all attributes of the image content for this number of images.
- Manual annotations fail to deal with the discrepancy of subjective perception. The phrase, "an image says more than a thousand words," implies

that the textual description is not sufficient for depicting subjective perception. Typically, a medical image usually contains several objects, which convey specific information. Nevertheless, different interpretations for a pathological area can be made by different radiologists. To capture all knowledge, concepts, thoughts, and feelings for the content of any images is almost impossible.

- The contents of medical images are difficult to be concretely described in words. For example, irregular organic shapes cannot easily be expressed in textual form, but people may expect to search for images with similar contents based on the examples they provide.

These problems limit the feasibility of text-based search for medical image retrieval. In an attempt to overcome these difficulties, content-based retrieval has been proposed to automatically access images with minimal human intervention (Eakins, 2002; Feng et al., 2003). However, due to the nature of medical images, content-based retrieval for medical images is still faced with challenges:

- Low resolution and strong noise are two common characteristics in most medical images (Glatard, Montagnat, & Magnin, 2004). With these characteristics, medical images cannot be precisely segmented and extracted for the visual content of their features. In addition, medical images obtained from different scanning devices may display different features, though some approaches to image correction and normalization have been proposed (Buhler, Just, Will, Kotzerke, & van den Hoff, 2004);
- Medical images are digitally represented in a multitude of formats based on their modality and the scanning device used (Wong & Hoo, 2002). Another characteristic of medical images is that many images are represented in gray level rather than color. Even with the change of intensity, monochrome may fail to clearly display the actual circumstance of lesion area.

Medical Applications of Content-Based Image Retrieval

Content-based image retrieval has frequently been proposed for various applications. This section will discuss three potential applications of medical CBIR.

PACS/Health Database Management

Content-based image retrieval has been proposed by the medical community for inclusion into picture archiving and communication systems (PACS) (Lehmann, Wein, & Greenspan, 2003). The idea of PACS is to integrate imaging modalities and interfaces with hospital and departmental information systems in order to manage the storage and distribution of images to radiologists, physicians, specialists, clinics, and imaging centers (Huang, 2003). A crucial point in PACS is to provide an efficient search function to access desired images. Image search in the digital imaging and communication in medicine (DICOM) protocol is currently carried out according to the alphanumerical order of textual attributes of images. However, the information which users are interested in is the visual content of medical images rather than that residing in alphanumerical format (Lehmann et al., 2003). The content of images is a powerful and direct query which can be used to search for other images containing similar content. Hence, content-based access approaches are expected to have a great impact on PACS and health database management. In addition to PACS, medical imaging databases that are unconnected to the PACS can also obtain benefits from CBIR technology.

Computer-Aided Diagnosis

Computer-aided diagnosis has been proposed to support clinical decision making. One clinical decision-making technique is case-based reasoning, which searches for previous, already- solved problems similar to the current one and tries to apply those solutions to the current problem (Hsu & Ho, 2004; Schmidt, Montani, Bellazzi, Portinale, & Gierl, 2001). This technique has a strong need to search for previous medical images with similar pathological areas, scrutinize the histories of these cases which are valuable for supporting certain diagnoses, and then reason the current case (Chang et al., 2004).

Medical Research, Education, and Training

CBIR technology can benefit any work that requires the finding of images or collections of images with similar contents. In medical research, researchers can use CBIR to find images with similar pathological areas and investigate their association. In medical education, lecturers can easily find images with particular pathological attributes, as those attributes can imply particular diseases. In addition, CBIR can

be used to collect images for medical books, reports, papers, and CD-ROMs based on the educational atlas of medical cells, where typical specimens are collected according to the similarity of their features, and the most typical ones are selected from each group to compose a set of practical calibrators.

Existing Medical CBIR Systems

Although content-based image retrieval has frequently been proposed for use in medical image management, only a few content-based retrieval systems have been developed specifically for medical images. These research-oriented systems are usually constructed in research institutes and continue to be improved, developed, and evaluated over time. This section will introduce several major medical content-based retrieval systems.

ASSERT (Automatic Search and Selection Engine with Retrieval Tools)

Developers: Purdue University, Indiana University, and University of Wisconsin Hospital, USA.

Image Database: High-Resolution Computed Tomography (HRCT) of lung.

Selected References: Shyu, Brodley, Kak, Kosaka, Aisen, and Broderick (1999), and Brodley, Kak, Dy, Shyu, Aisen, and Broderick, (1999).

Web site: http://rvl2.ecn.purdue.edu/~cbirdev/WWW/CBIRmain.html

Main Characteristics:

- The ASSERT system uses a physician-in-the-loop approach to retrieving images of HRCT of the lung. This approach requires users to delineate the pathology-bearing regions and identify certain anatomical landmarks for each image;
- This system extracts 255 features of texture, shape, edges, and gray-scale properties in pathology-bearing regions;
- A multi-dimensional hash table is constructed to index the HRCT images.

CasImage

Developer: University Hospital of Geneva, Switzerland.

Image Database: A variety of images from CT, MRI, and radiographs, to color photos.

Selected References: Muller, Rosset, Vallee, and Geissbuhler (2004b), and Rosset, Ratib, Geissbuhler, and Vallee (2002).

Web site: http://www.casimage.com/

Main Characteristics:

- The CasImage system, which has been integrated into a PACS environment, contains a teaching and reference database, and the medGIFT retrieval system, which is adapted from the open-source GIFT (GNU Image Finding Tool) (Squire, Muller, Muller, Marchand-Maillet, & Pun, 2001);
- The medGIFT retrieval system extracts global and regional color and texture features, including 166 colors in the HSV color space, and Gabor filter responses in four directions each at three different scales;
- Combinations of textual labels and visual features are used for medical image retrieval.

IRMA (Image Retrieval in Medical Applications)

Developer: Aachen University of Technology, Germany.

Image Database: Various imaging modalities.

Selected References: Lehmann, Guld, Keysers, Deselaers, Schubert, Wein, and Spitzer (2004a), and Lehmann, Guld, Thies, Plodowski, Keysers, Ott, and Schubert (2004b).

Web site: http://libra.imib.rwth-aachen.de/irma/

Main Characteristics:

- The IRMA system is implemented as a platform for content-based image retrieval in medical applications;

- This system splits the image retrieval process into seven consecutive steps, including categorization, registration, feature extraction, feature selection, indexing, identification, and retrieval.

NHANES II (The Second National Health And Nutrition Examination Survey)

Developer: National Library of Medicine, USA.

Image Database: 17,000 cervical and lumbar spine X-ray images.

Selected References: Antani, Lee, Long, and Thoma (2004a), and Antani, Xu, Long, and Thoma (2004b).

Web site: http://archive.nlm.nih.gov/proj/dxpnet/nhanes/nhanes.php

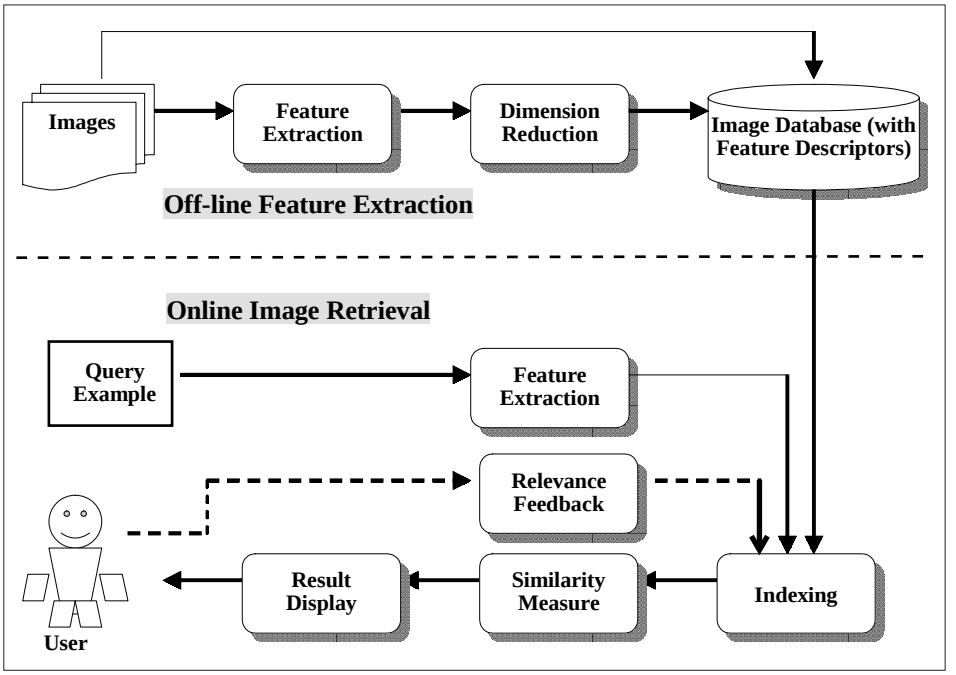
Main Characteristics:

- This system contains the Active Contour Segmentation (ACS) tool, which allows the users to create a template by marking points around the vertebra. If the segmentation of a template is accepted, the ACS tool will estimate the location of the next vertebra, place the template on the image, and then segment it;
- In data representation, a polygon approximation process is applied for eliminating insignificant shape features and reducing the number of data points. The data obtained in the polygon approximation process represent the shape of vertebra. Then, the approximated curve of vertebra is converted to tangent space for similarity measurement.

Content-Based Retrieval Systems

Content-based retrieval uses the contents of images to represent and access the images (Wei & Li, 2005). A typical content-based retrieval system is divided into *off-line feature extraction* and *online image retrieval*. A conceptual framework for content-based image retrieval is illustrated in Figure 1. In off-line feature extraction, the contents of the images in the database are extracted and described with a multi-dimensional feature vector, also called descriptor.

Figure 1. A conceptual framework for content-based image retrieval



The feature vectors of the image constitute a feature dataset stored in the database. In online image retrieval, the user can submit a query example to the retrieval system in search of desired images. The system represents this example with a feature vector. The distances (i.e., similarities) between the feature vectors of the query example and those of the media in the feature dataset are then computed and ranked. Retrieval is conducted by applying an indexing scheme to provide an efficient way of searching the image database. Finally, the system ranks the search results and then returns the results that are most similar to the query examples. If the user is not satisfied with the search results, the user can provide relevance feedback to the retrieval system, which contains a mechanism to learn the user's information needs. The following sections will clearly introduce each component in the system.

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. Feature extraction of the image in the database is typically conducted off-line so computation complexity is not a significant issue. This section will introduce two features — texture and color — which are used most often to extract the features of an image.

Color

Color is a powerful descriptor that simplifies object identification (Gonzalez & Woods, 2002) and is one of the most frequently used visual features for content-based image retrieval. To extract the color features from the content of an image, a proper color space and an effective color descriptor have to be determined.

The purpose of a color space is to facilitate the specification of colors. Each color in the color space is a single point represented in a coordinate system. Several color spaces, such as *RGB, HSV, CIE L*a*b, and CIE L*u*v*, have been developed for different purposes. Although there is no agreement on which color space is the best for CBIR, an appropriate color system is required to ensure perceptual uniformity. Therefore, the *RGB* color space, a widely used system for representing color images, is not suitable for CBIR because it is a perceptually non-uniform and device-dependent system (Gevers, 2001).

The most frequently used technique is to convert color representations from the *RGB* color space to the *HSV*, CIE *L*u*v*, or CIE *L*a*b* color spaces with perceptual uniformity (Li & Yuen, 2000). The *HSV* color space is an intuitive system, which describes a specific color by its hue, saturation and brightness value. This color system is very useful in interactive color selection and manipulation; The CIE *L*u*v* and CIE *L*a*b* color spaces are both perceptually uniform systems, which provide easy use of similar metrics for comparing color (Haeghen, Naeyaert, Lemahieu, & Philips, 2000).

After selecting a color space, an effective color descriptor should be developed in order to represent the color of the global or regional areas. Several color descriptors have been developed from various representation schemes, such as color histograms (Quyang & Tan, 2002), color moments (Yu et al., 2002), color edge (Gevers & Stokman, 2003), color texture (Guan & Wada, 2002), and color correlograms (Moghaddam, Khajoie, & Rouhi, 2003). For example, color histogram, which represents the distribution of the number of pixels for each quantized color bin, is an effective representation of the color content of an image. The color histogram can not only easily characterize the

global and regional distribution of colors in an image, but also be invariant to rotation about the view axis.

For the retrieval of medical images, color allows images to reveal many pathological characteristics (Tamai, 1999). Color also plays an important role in morphological diagnosis (Nishibori, Tsumura, & Miyake, 2004). Color medical images are usually produced in different departments and by various devices. For example, color endoscopic images are taken by a camera that is put into the hollow organs of the body such as stomachs and lungs. A common characteristic in such kind of images is that most colors are made of various stains, though fine variations of natural colors are crucial for diagnosis. Nishibori (2000) pointed out that problems in color medical images include inaccurate color reproduction, rough gradations of color, and insufficient density of pixels. Therefore, effective use of the various color information in images includes absolute color values, ratios of each tristimulus color, differences in colors against adjacent areas, and estimated illumination data. In addition, many medical images are represented in gray level. For this kind of gray level images, CBIR can only regard color as secondary features because gray levels provide limited information about the content of an image. For specific purposes, some gray level images have pseudo-color added to enhance specific areas instead of gray level presentation. Such processing increases difficulties in retrieval.

Texture

Texture in CBIR can be used for at least two purposes (Sebe & Lew, 2002). First, an image can be considered to be a mosaic that consists of different texture regions. These regions can be used as examples to search and retrieve similar areas. Second, texture can be employed for automatically annotating the content of an image. For example, the texture of an infected skin region can be used for annotating regions with the same infection.

Textural representation approaches can be classified into statistical approaches and structural approaches (Li, 1998). Statistical approaches analyze textural characteristics according to the statistical distribution of image intensity. Approaches in this category include gray level co-occurrence matrix, fractal model, Tamura feature, Wold decomposition, and so on (Feng et al., 2003). Structural approaches characterize texture by identifying a set of structural primitives and certain placement rules.

If medical images are represented in gray level, texture becomes a crucial feature, which provides indications about scenic depth, the spatial distribution of tonal variations, and surface orientation (Tourassi, 1999). For example, abnormal symptoms on female breasts include calcification, architectural distortion, asymmetry, masses, and so forth. All of those reveal specific textural patterns on the mammograms. However, selection of texture features for specifying textural structure should take account of the influence from the modulation transfer function on texture (Veenland, Grashuis, Weinans, Ding, & Vrooman, 2002). As the intensifying screens are used to enhance the radiographs, the blurring effect also changes texture features, that is, spatial resolution, contrast, and sharpness are all reduced in the output. Low resolution and contrast result in difficulties in measuring the pattern of tissue and structure of organs (Majumdar, Kothari, Augat, Newitt, Link, Lin, & Lang, 1998).

Dimension Reduction

In an attempt to capture useful contents of an image and to facilitate effective querying of an image database, a CBIR system may extract a large number of features from the content of an image. Feature set of high dimensionality causes the "curse of dimension" problem in which the complexity and computational cost of the query increase exponentially with the number of dimensions (Egecioglu, Ferhatosmanoglu, & Ogras, 2004).

To reduce the dimensionality of a large feature set, the most widely-used technique in image retrieval is principal component analysis (PCA). The goal of principal component analysis is to specify as much variance as possible with the smallest number of variables (Partridge & Calvo, 1998). Principal component analysis involves transforming the original data into a new coordinate system with low dimension, thus creating a new set of data. The new coordinate system removes the redundant data, and the new set of data may better represent the essential information. However, there is a trade-off between the efficiency obtained through dimension reduction and the completeness of the information extracted. As data is represented lower dimensions, the speed of retrieval is increased, but some important information may be lost in the process of data transformation. In the research of medical CBIR, Sinha and Kangarloo (2002) demonstrated the PCA application to the image classification of 100 axial brain images.

Similarity Measure

Selection of similarity metrics has a direct impact on the performance of content-based image retrieval. The kind of feature vectors selected determines the kind of measurement that will be used to compare their similarity (Smeulders, Worring, Santini, Gupta, & Jain, 2000). If the features extracted from the images are presented as multi-dimensional points, the distances between corresponding multi-dimensional points can be calculated. Euclidean distance is the most common metric used to measure the distance between two points in multi-dimensional space (Qian, Sural, Gu, & Pramanik, 2004). For other kinds of features such as color histogram, Euclidean distance may not be an ideal similarity metric or may not be compatible with the human-perceived similarity. Histogram intersection was proposed by Swain and Ballard (1991) to find known objects within images using color histograms. A number of other metrics, such as Mahalanobis Distance, Minkowski-Form Distance, Earth Mover's Distance, and Proportional Transportation Distance, have been proposed for specific purposes. Antani, Long, Thoma, and Lee (2003) used several approaches to code the shape features for different classes of spine X-rays. Each class used a specific similarity metric to compare the distance between two feature vectors.

Multi-Dimensional Indexing

Retrieval of an image is usually based not only on the value of certain features, but also on the location of a feature vector in the multi-dimensional space (Fonseca & Jorge, 2003). A retrieval query on a database of multimedia with multi-dimensional feature vectors usually requires fast execution of search operations. To support such search operations, an appropriate multi-dimensional access method has to be used for indexing the reduced but still high dimensional feature set. Popular multi-dimensional indexing methods include the R-tree (Guttman, 1984) and the R*-tree (Beckmann, Kriegel, Schneider, & Seeger, 1990).

The R-tree, which is a tree-like data structure, is mainly used for indexing multi-dimensional data. Each node of an R-tree has a variable number of entries. Each entry within a non-leaf node can have two pieces of data. The goal of the R-tree is to organize the spatial data in such a way that a search will visit as few spatial objects as possible. The decision on which nodes to visit is made based

on the evaluation of spatial predicates. Hence, the R-tree must be able to hold some sort of spatial data on all nodes. The R*-tree, an improved version of the R-tree, applies more complex criteria for the distribution of minimal- bounding rectangles through the nodes of the tree, such as: The overlap between the minimal- bounding rectangles of the inner nodes should be minimized; the perimeter of a directory rectangles should be minimized; and storage utilization should be maximized. Both techniques perform well in low-dimensional features-space with a limit of up to 20 dimensions. For high-dimensional features-space, it is necessary to reduce the dimensionality using statistic multi-variate analysis techniques such as the aforementioned principal component analysis.

With regard to medical CBIR research, Shyu et al. (1999) successfully applied multi-dimensional indexing in the ASSERT system. In this system, lobular feature sets (LFS) on HRCT images are translated into an index for archiving and retrieval. A multi-dimensional hash table for the LFS classes is constructed for the system. A decision tree algorithm is used to construct a minimum-entropy partition of the feature space where the LFS classes reside. After translating a decision tree to a hash table, the system prunes the set of retrieved LFS classes and candidate images.

Relevance Feedback

Relevance feedback was originally developed for improving the effectiveness of information retrieval systems. The main idea of relevance feedback is for the retrieval system to understand the user's information needs. For a given query, the retrieval system returns initial results based on pre-defined similarity metrics. Then, the user is required to identify the positive examples by labeling those that are relevant to the query. The system subsequently analyzes the user's feedback using a learning algorithm and returns refined results.

A typical relevance feedback mechanism contains a learning component and a dispensing component. The learning component uses the feedback data to estimate the target of the user. The approach taken to learn feedback data is key to the relevance feedback mechanism. In addition to Rocchio's (1971) and Rui and Huang's (2002) learning algorithms, recent work has reported that support vector machine (SVM) is a useful learning approach in relevance feedback (Ferecatu, Crucianu, & Boujemaa, 2004a; Hoi, Chan, Huang, Lyu, & King, 2004; Tao & Tang, 2004). The dispensing component should provide the most appropriate images after obtaining feedback from the user. However,

the dispensing component has two conflicting goals during each feedback round. On the one hand, the dispensing component has to provide as many relevant images as possible. On the other hand, the dispensing component, based on the information needs of the user, has to investigate the images of unknown relevance to the target (Ferecatu, Crucianu, & Boujemaa, 2004b). As the dispensing component returns more relevant images to the user, it has fewer images to mine the needs of the user at each round, and vice versa. A sensible strategy also plays an important role in relevance feedback. Hence, approaches to learning user feedbacks and dispensing strategies for returning the results both determine the performance of relevance feedback mechanisms.

Medical images have a unique characteristic in that their contents always reflect pathological attributes or symptoms for specific diseases. Image classification is often used to group the similar features based on their contents. With this characteristic, relevance feedback is expected to assist in mining the common features of relevant images and finding a specific class where the query example should reside. Those images grouped in the same class have the same semantics and are likely to be target images. El-Naqa, Yang, Galatsanos, and Wernick (2003) proposed a relevance feedback approach based on incremental learning for mammogram retrieval. They adapted support vector machines (SVM) to develop an online learning procedure for similarity learning. The approach they proposed was implemented using clustered micro-calcifications images. They reported that the approach significantly improves the retrieval effectiveness. In addition, El-Naqa et al. (2004) also demonstrated a hierarchical two-stage learning network, which consists of a cascade of a binary classifier and a regression module. Relevance feedback is incorporated into this framework to effectively improve precision based on online interaction with users.

Case Study

This section will propose a general CBIR framework and its application to mammogram retrieval, and demonstrate its method. Breast cancer continues to be a serious disease across the world. Mammography is a reliable method for detection of breast cancer. There are an enormous number of mammograms generated in hospitals. How to effectively retrieve a desired image from mammogram databases is a challenging problem. This study concentrates on

textural analysis based on gray level co-occurrence matrices for the content-based retrieval of mammograms. The objectives of this study are as follows:

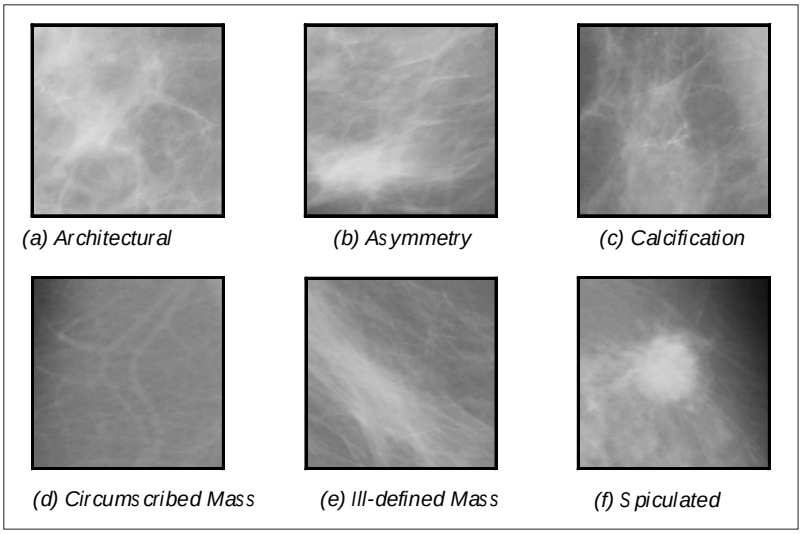
1. To analyze and examine the textural features present in the ROI (Region of Interest) of abnormal breast tissue as compared to the same information presented in normal tissue;
2. To develop the optimal mammographic descriptor generated from gray level co-occurrence matrices; and
3. To evaluate the effectiveness of the CBIR system using descriptors with different unit pixel distances.

The method in this work contains two major stages — *image analysis* and *image retrieval*. The objective of the image analysis stage is to examine the textural features of mammograms, and then test the statistical significance of the differences between normal and abnormal mammograms. These discriminating features are selected to construct a textural descriptor of mammograms. The descriptor constructed in the image analysis stage is embedded into the CBIR system. The feature descriptor is extracted from the query image in order to retrieve the mammograms relevant to the query image. The performance of the CBIR system is then evaluated. The detailed steps and components of the experiment are described in the following sections.

Mammogram Dataset

Mammograms were obtained from the database of the Mammographic Images Analysis Society (MIAS) (Suckling, Parker, Dance, Astley, Hutt, Boggis, Ricketts, Stamatakis, Cerneaz, Kok, Taylor, Betal, & Savage, 1994). The size of each image was 1024×1024 pixels. All of the images have been annotated for class, severity and location of abnormality, character of background tissue, and radius of circle enclosing the abnormality. Abnormalities are classified into calcifications architectural distortions, asymmetries, circumscribed masses, speculated masses, and ill-defined masses. Sub-images of size 200×200 pixels were cropped as ROIs from each mammogram. 122 sample ROIs (including 29 images in calcification class, 19 in architectural distortion class, 15 in asymmetry class, 25 in circumscribed masses class, 19 in speculated masses class, and 15 in other or ill-defined masses class) were selected

Figure 2. *Abnormal mammograms are classified into calcification, architectural distortion, asymmetry, circumscribed masses, speculated masses, and ill-defined masses*



deliberately from abnormal tissues. Another 207 ROIs were obtained arbitrarily from normal tissues. These 329 ROIs were used to analyze their textural features based on gray level co-occurrence matrices.

Feature Analysis

The presence of a breast lesion may cause a disturbance in the homogeneity of tissues, and result in architectural distortions in the surrounding parenchyma (Cheng & Cui, 2004). Therefore, the textures of digital images contain a lot of valuable information for further research and application. This study applies gray level co-occurrence matrices, a statistical textural method, to analyze the textural features of mammograms and develop descriptors for content-based image retrieval. Gray level co-occurrence matrices will be introduced in the following section.

Gray Level Co-occurrence Matrices

Gray level co-occurrence matrix (GLCM) is a statistical method for computing the co-occurrence probability of textural features (Haralick, 1979). Given an

image $f(x, y)$ of size $L_r \times L_c$ with a set of N_g gray levels, define the matrix $p(i, j, d, \theta)$ as

$$P(i, j, d, \theta) = card\{((x_1, y_1), (x_2, y_2)) \in (L_r \times L_c)\ (L_r \times L_c) \mid (x_2, y_2) = (x_1, y_1) + (d\cos\theta, d\sin\theta), f(x_1, y_1) = i, f(x_2, y_2) = j, 0 \le i, j < N_g\} \quad (1)$$

where d denotes the distance between pixels (x_1, y_1) and (x_2, y_2) in the image, θ denotes the orientation aligning (x_1, y_1) and (x_2, y_2), and card $\{\cdot\}$ denotes the number of elements in the set. Texture features that can be extracted from gray level co-occurrence matrices (Haralick, Shanmugan, & Dinstein, 1973) are:

$$\text{Angular Second Moment (ASM)} = \sum_i \sum_j \{p(i, j)\}^2 \quad (2)$$

$$\text{Contrast} = \sum_{n=0}^{Ng-1} n^2 \left\{ \sum_{i=1}^{Ng} \sum_{\substack{j=1 \\ |i-j|=n}}^{Ng} p(i, j) \right\} \quad (3)$$

$$\text{Correlation} = \frac{\sum_i \sum_j (ij)\, p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (4)$$

$$\text{Variance} = \sum_i \sum_j (i - j)^2\, p(i, j) \quad (5)$$

$$\text{Inverse Difference Moment (ID\_Mom)} = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad (6)$$

$$\text{Sum Average (Sum\_Aver)} = \sum_{i=2}^{2Ng} i p_{x+y}(i) \quad (7)$$

$$\text{Sum Variance (Sum\_Var)} = \sum_{i=2}^{2Ng} (i - Sum\_Entro)^2 p_{x+y}(i) \quad (8)$$

$$\text{Sum Entropy (Sum\_Entro)} = -\sum_{i=2}^{2Ng} p_{x+y}(i) \log\{p_{x+y}(i)\} \quad (9)$$

$$\text{Entropy} = -\sum_i \sum_j p(i, j) \log p(i, j) \quad (10)$$

$$\text{Different Variance (Diff\_Vari)} = \text{variance of } p_{x-y} \quad (11)$$

$$\text{Different Entropy (Diff\_Entro)} = -\sum_{i=0}^{Ng-1} p_{x-y}(i) \log\{p_{x-y}(i)\} \quad (12)$$

Mammogram Analysis Using GLCM

In order to develop the tailored descriptors described in the next section, it is necessary to analyze the features of the mammogram. In this study, 12 GLCMs are constructed in order to compute each ROI in the 0°, 45°, 90°, and 135° directions, each with unit pixel distances of 1, 3, and 5, respectively. The 11 features described earlier are computed for the 12 GLCMs, thus resulting in a total of 132 texture features for each ROI.

Feature Selection for Image Retrieval

At the stage of feature analysis, 132 texture features are generated for each ROI. We obtain 5,148 texture feature sample images from 20 normal and 19 abnormal images. To select the most discriminant features, a statistical multivariate *t*-test is used to assess the significance of the difference between the means of two sample set *A* and *B*, which are independent of each other in the obvious sense, that is, the individual measures in set *A* are in no way related to any of the individual measures in set *B*. The value of the *t*-test is obtained as follows (Serdobolskii, 2000):

$$D_a = \sum (A_i - \mu_a)^2 \tag{13}$$

$$D_b = \sum (B_i - \mu_b)^2 \tag{14}$$

$$V = \frac{D_a + D_b}{(n_a - 1) + (n_b - 1)} \tag{15}$$

$$\sigma = \sqrt{\frac{V}{n_a} + \frac{V}{n_b}} \tag{16}$$

$$t = \frac{\mu_a - \mu_b}{\sigma} \tag{17}$$

where A_i and B_i in equations (13) and (14) are the *i*th element of the set *A* and *B*, while μ_a and μ_b are the means of the set *A* and *B*, respectively. D_a and D_b in the equations (13) and (14) are the sum of squared deviates of *the* set A and B. *V* in equation (15) is the estimated variance of the source population. σ in equation (16) is the standard deviation of the sampling distribution of sample-mean differences. *t* in equation (17) is the value of the *t*-test. The degree of freedom (d.f.) is $(n_a - 1) + (n_b - 1)$.

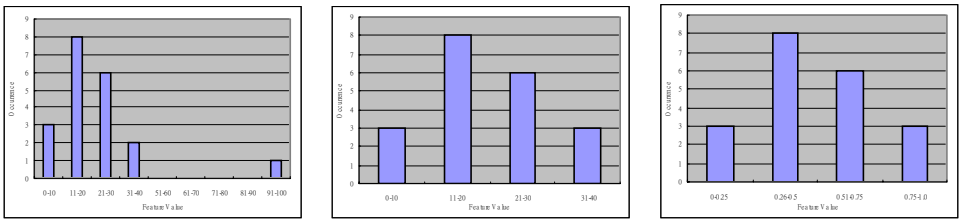
In our case study with 20 normal images (set *A*) and 19 abnormal images (set *B*), the degree of freedom (d.f.) is 37. According to the Table of Critical Values of *t* (Rencher, 1998), the *t* value for 37 degrees of freedom (d.f.) is 1.305. When the value *t* obtained in this experiment is greater than 1.305, it means that there is a significant mean difference between normal and abnormal mammograms with regard to the given feature.

The descriptor is composed of features with significant differences in the *t* statistic. Individual descriptors were developed for three distances ($d = 1, 3,$ and 5) in the gray level co-occurrence matrices.

Data Normalization

The purpose of normalization in this experiment is to assign a weight to all features in order to measure their similarity on the same basis. The technique

Figure 3. Process of normalization



(a) The original distribution *(b) The trimmed distribution* *(c) The normalized distribution*

used was to project each feature onto a unit sphere. However, a potential problem exists — if a few elements in a feature space are extremely large, other elements may be dominated by these large ones after normalization.

To solve this problem, the value located at the point of the top 95% of the distribution is taken as the nominal maximum. All features greater than the nominal minimum in the feature space were clipped to the nominal maximal value, that is, the top 5% of distribution are trimmed. Then all values are divided by the maximal values. An example is given to illustrate the use of this approach to normalization. Figure 3a shows the original distribution of feature values in a feature vector. The results of trimming the top 5% and normalization are illustrated in Figures 3b and 3c.

Similarity Measure

The similarity measure of two images I_a and I_b is the distance between their descriptors f_a and f_a. In this work, L_2 norm was adopted to measure the similarity between the query image and each ROI. L_2 is defined as follows:

$$\| d_{ab} \|_2 = | f_a - f_b |^2 = \sqrt{\sum_{i=1}^{n} | f_{a,i} - f_{b,i} |^2}, \quad (18)$$

where d_{ab} is the similarity distance between descriptors f_a and f_b, $f_{a,i}$ and $f_{b,i}$ are the ith element of f_a and f_a, respectively, and n is the number of elements of the descriptors. The smaller the distance is, the more similar the two images are. After calculating the distance, our CBIR system ranks similarity in descending order and then returns the top five images that are most similar to the query image.

Performance Evaluation

Relevance judgment is a vital part of performance evaluation. The relevance criteria described in Table 1 were developed and used in this work. For example, suppose the query image belongs to the calcification class, the retrieved image would score 0.5 if it belongs to any of the following abnormal classes: ill-defined masses, circumscribed masses, speculated masses, architectural distortion, and asymmetry.

Precision and recall are basic measures used in evaluating the effectiveness of an information retrieval system. Precision is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved (Baeza-Yates, & Ribeiro-Neto, 1999). It indicates the subject score assigned to each of the top five images in this experiment. The formula is expressed as follows:

$$p = \frac{\sum_{i=1}^{n} S_i}{N} \tag{19}$$

where S_i is the score assigned to the ith hit, N is the number of top hits retrieved.

Recall is the ratio of the number of relevant records retrieved to the total number of relevant records in the database (Baeza-Yates & Ribeiro-Neto, 1999). It is defined as follows:

Table 1. Criteria for measurement of performance evaluation of CBIR

Score	Criteria
1.0	The retrieved image belongs to the class of query image.
0.5	The retrieved image belongs to one of the abnormal classes, but not the class of query image.
0	The retrieved image does not belong to any abnormal class.

$$R = \frac{R_n}{T_n} \tag{20}$$

where R_n is the number of retrieved relevant hits, and T_n is the total number of relevant images in the database.

Results of *t*-Test

Table 2 presents the results of the *t* statistic for $d = 1$ of gray level co-occurrence matrices. From the results, it can be seen that the differences between the mean values of ASM, Correlation, Sum_Var (sum variance), and Diff_Entro (difference entropy) of the normal and abnormal ROIs are significant ($t > 1.305$). As a result, these four features are selected to construct the descriptor. Table 3 shows that ASM is the only feature with significant discriminating power for two groups of ROIs when $d = 3$. The descriptor with $d = 3$ contains only ASM.

Table 4 shows that Sum_Var (sum-variance) is the only feature with significant discriminating power for two groups of ROIs when $d = 5$. The descriptor with $d = 5$ contains only Sum_Var.

Table 2. Comparison of mean values obtained by co-occurrence matrices with the distance of 1

	Normal		Abnormal		
Feature	μ_a	D_a	μ_b	D_b	t (37 .d.f.)
ASM	1.7721	0.6127	1.3890	0.5542	1.6042
Contrast	5.8209	2.2742	5.6187	3.7038	0.3755
Correlation	-0.3392	0.0742	-0.6135	0.5345	1.6071
Variance	0.2413	0.0768	1.5446	1.9193	-4.2257
ID_Mom	2.9460	0.0671	3.0654	0.1442	-1.1812
Sum_Aver	2.1943	0.3391	2.6984	1.2137	-1.8444
Sum_Var	3.2636	2.2970	0.7699	1.3261	5.9097
Sum_Entro	-6.7253	2.7305	-0.7922	0.1327	-15.6895
Entropy	-4.0729	0.1476	-3.8611	0.3262	-1.3995
Diff_Vari	5.2007	2.3020	6.7532	1.4977	-3.5953
Diff_Entro	-1.7865	0.0332	-1.9603	0.0141	3.5986

Table 3. Comparison of mean values obtained by co-occurrence matrices with the distance of 3

	Normal		Abnormal		
Feature	μ_a	D_a	μ_b	D_b	t (37 .d.f.)
ASM	1.2846	0.3212	1.0075	0.2905	1.6046
Contrast	1.3822	0.2398	1.6247	0.4604	-1.3191
Correlation	0.8311	0.0402	3.6871	15.2704	-3.3534
Variance	1.5943	3.4598	1.1012	0.9300	1.0584
ID_Mom	1.465	0.0518	1.5281	0.1695	-0.6123
Sum_Aver	1.6135	0.1835	1.9774	0.6539	-1.8161
Sum_Entro	-4.9365	1.4720	-0.5789	0.0708	-15.7084
Entropy	-2.6801	0.1052	-2.4949	0.4599	-1.1262
Diff_Vari	4.1858	1.2969	5.3510	0.8446	-3.5984
Diff_Entro	-1.6118	0.0184	1.7424	0.0088	-91.7379

Table 4. Comparison of mean values obtained by co-occurrence matrices with the distance of 5

	Normal		Abnormal		
Feature	μ_a	D_a	μ_a	D_b	t (37 .d.f.)
ASM	1.2466	0.3019	0.9784	0.2730	0.7924
Contrast	2.1675	0.8227	2.9725	2.0665	-1.5708
Correlation	0.8822	0.0387	4.6453	14.6955	-7.1316
Variance	1.4357	2.8461	1.0654	0.8467	1.0316
ID_Mom	1.2008	0.0497	1.227	0.1625	-0.0742
Sum_Aver	1.5855	0.1775	1.9376	0.6293	-0.8292
Sum_Var	2.3006	2.6245	0.5359	0.6438	4.5870
Sum_Entro	-4.8432	1.4197	-0.5657	0.0677	-8.0337
Entropy	-2.5128	0.1179	-2.3037	0.2179	-0.4203
Diff_Vari	4.1858	1.2969	5.3510	0.8446	-1.6684
Diff_Entro	-1.6118	0.0184	-1.7424	0.0088	0.3150

Results of Performance Evaluation

Precision can be used to describe the accuracy of the proposed CBIR system in finding only relevant images on a search for query images. Table 5 shows that precision rates for the three descriptors (d = 1, 3, and 5) are 47%, 50%, and 51%, respectively. The descriptor with d = 5 obtained the highest value in precision and the smallest values in standard deviation. This indicates that the performance of the descriptor (d = 5) is more stable. On the whole, about half of the results retrieved by these three descriptors are relevant to the query images.

Recall measures how well the CBIR system finds all relevant images in a search for a query image. Table 6 indicates that the descriptor with d = 5 outperforms the other two. However, the recall values are very close. The largest difference is only 1.5%. The three descriptors can retrieve, on average, about 18% of relevant images in the database. In theory, as precision goes up, recall goes down. The relationship explains why the three recall values are low.

The experimental results also show that the descriptor with the largest distance (d = 5) has the best performance in both precision and recall. The descriptor with d = 3 outperforms the descriptor with d = 1 in both measures. Although the larger distance has better performance in this experiment, it is still too early to make any conclusions.

Table 5. Precision for the 3 descriptors

	CALC	CIRC	SPIC	MISC	ARCH	ASYM	Mean	Std
***d*=1**	42%	44%	54%	44%	46%	54%	47%	5.32%
***d*=3**	57%	43%	54%	42%	54%	51%	50%	6.24%
***d*=5**	48%	53%	54%	50%	50%	48%	51%	2.51%

Table 6. Recall for the 3 descriptors

	CALC	CIRC	SPIC	MISC	ARCH	ASYM	Mean	Std
***d*=1**	10.69%	12.80%	20.53%	18.67%	16.84%	26.00%	17.59%	5.51%
***d*=3**	13.45%	11.60%	20.00%	20.00%	18.42%	25.33%	18.13%	4.97%
***d*=5**	12.07%	15.20%	21.05%	24.00%	20.00%	22.00%	19.05%	4.51%

(Notes: CALC = calcification; CIRC = circumscribed masses, SPIC = speculated masses; ARCH = architectural distortion; ASYM = asymmetry; MISC = other or ill-defined masses.)

The main contribution of this work is to present a sound CBIR methodology for mammograms. The methodology is divided into image analysis and image retrieval stages. The purpose of the image analysis is to collect samples from the database, obtain the image signature, and then apply it for the feature extraction in the image retrieval stage. A complete CBIR system based on gray level co-occurrence matrices was implemented. A technique was also proposed to improve the effectiveness of normalization. Three descriptors were evaluated by query images to retrieve the ROIs for the mammogram dataset consisting of 122 images of six sub-classes from abnormal class, and 207 images from normal class. The best precision rate of 51% and recall rate of 19% were achieved with the descriptor using gray level co-occurrence matrices with the pixel distance of 5.

Research Issues

Content-based retrieval for medical images is still in its infancy. There are many challenging research issues. This section identifies and addresses some issues in the future research agenda.

Bridging the Semantic Gap

An ideal medical CBIR system from a user perspective would involve semantic retrieval, in which the user submits a query like "find MRIs of brain with tumor". This kind of open-ended query is very difficult for the current CBIR systems to distinguish brain MRI's from spine MRIs even though the two types of images are visually different. Current medical CBIR systems mainly rely on low-level features like texture, color, and shape.

Systems Integration

Most medical retrieval systems are designed for one particular type of medical image, such as mammogram or MRIs of spine. Specific techniques and modalities are developed based on the characteristics of these highly homogeneous images in their databases. However, medical image databases across different medical institutions have been expected to connect through PACS.

With PACS, a user may make a request to search for medical images among different databases, where images display different characteristics such as degree of resolution, degree of noise, use of color, shape of object, and texture of background. In other words, PACS can be seen as a single retrieval system with distributed image databases, which collect medical images with various modalities. Therefore, systems capable of finding images across heterogeneous image databases are desirable.

Human-Computer Interaction and Usability

Current research on medical CBIR concentrates on the effectiveness of the system, rarely evolving the relationship between CBIR and user interface design. However, innovative retrieval systems alone may not obtain user acceptance as users of medical CBIR systems may include radiologists, surgeons, nurses, or other users without specific knowledge of these systems. The user's experience, or how the user experiences the system, is the key to acceptance. Good user interface design is usually required for the end user to easily learn and use the system. Also, empirical usability testing permits naïve users to provide information about the usability of individual system functions and components.

Performance Evaluation

The National Institute of Standards and Technology (NIST) has developed TREC (Text REtrieval Conference) as the standard test-bed and evaluation paradigm for the information retrieval community (Smeaton, 2003). The image retrieval community still awaits the construction and implementation of a scientifically-valid evaluation framework and standard test bed. To construct a test bed for medical CBIR, imaging modalities, regions, and orientations of images should be taken into account. Due to the complexity of medical images, how to construct a common test bed for medical CBIR is a research issue.

Conclusion

The goal of medical image databases is to provide an effective means for organizing, searching, and indexing large collections of medical images. This

requires intelligent systems that have the ability to recognize, capture, and understand the complex content of medical images. Content-based retrieval is a promising approach to achieve these tasks and has developed a number of techniques used in medical images. Despite recent developments, medical content-based image retrieval still has a long way to go and more efforts are expected to be devoted to this area. Ultimately, a well-organized image database, accompanied by an intelligent retrieval mechanism, can support clinical treatment, and provide a basis for better medical research and education.

References

Antani, S., Lee, D. J., Long, L. R., & Thoma, G. R. (2004a). Evaluation of shape similarity measurement methods for spine X-ray images. *Journal of Visual Communication and Image Representation, 15*(3), 285-302.

Antani, S., Long, L. R., Thoma, G. R., & Lee, D. J. (2003). Evaluation of shape indexing methods for content-based retrieval of X-ray images. In *Proceedings of IS&T/SPIE 15th Annual Symposium on Electronic Imaging, Storage, and Retrieval for Media Databases* (pp. 405-416).

Antani, S., Xu, X., Long, L. R., & Thoma, G. R. (2004b). Partial shape matching for CBIR of spine X-ray images. In *Proceedings of IS&T/SPIE Electronic Imaging — Storage and Retrieval Methods and Applications for Multimedia 2004* (pp. 1-8).

Baeza-Yates, R., & Ribeiro-Neto, B. (Eds.). (1999). *Modern information retrieval*. Boston: Addison-Wesley.

Beckmann, N., Kriegel, H.-P., Schneider, R., & Seeger, B. (1990). The R*-tree: An efficient and robust access method for points and rectangles. In *Proceedings of ACM SIGMOD International Conference on Management of Data* (pp. 322-331).

Brodley, C. E., Kak, A, Dy, J., Shyu, C. R., Aisen, A., & Broderick, L. (1999). Content-based retrieval from medical image database: A synergy of human interaction, machine learning, and computer vision. In *Proceedings of the Sixteenth National Conference on Artificial Intelligence* (pp. 760-767).

Buhler, P., Just, U., Will, E., Kotzerke, J., & van den Hoff, J. (2004). An accurate method for correction of head movement in PET. *IEEE Transactions on Medical Imaging, 23*(9), 1176-1185.

Chang, C.-L., Cheng, B.-W., & Su, J.-L. (2004). Using case-based reasoning to establish a continuing care information system of discharge planning. *Expert Systems with Applications, 26*(4), 601-613.

Cheng, H. D., & Cui, M. (2004). Mass lesion detection with a fuzzy neural network. *Pattern Recognition, 37*(6), 1189-1200.

Eakins, J. P. (2002). Towards in intelligent image retrieval. *Pattern Recognition, 35*(1), 3-14.

Egecioglu, O., Ferhatosmanoglu, H., & Ogras, U. (2004). Dimensionality reduction and similarity computation by inner-product approximations. *IEEE Transactions on Knowledge and Data Engineering, 16*(6), 714-726.

El-Naqa, I., Yang, Y., Galatsanos, N. P., & Wernick, M. N. (2003). Relevance feedback based on incremental learning for mammogram retrieval. In *Proceedings of the International Conference of Image Processing 2003* (pp. 729-732).

El-Naqa, I., Yang, Y., Galatsanos, N. P., Nishikawa, R. M., & Wernick, M. N. (2004). A similarity learning approach to content-based image retrieval: Application to digital mammography. *IEEE Transactions on Medical Imaging, 23*(10), 1233-1244.

Feng, D., Siu, W. C., & Zhang, H. J. (Eds.). (2003). *Multimedia information retrieval and management: Technological fundamentals and applications*. Berlin: Springer.

Ferecatu, M., Crucianu, M., & Boujemaa, N. (2004a). Retrieval of difficult image classes using SVM-based relevance feedback. In *Proceedings of the 6th ACM SIGMM International Workshop on Multimedia Information Retrieval* (pp. 23-30).

Ferecatu, M., Crucianu, M., & Boujemaa, N. (2004b). Sample selection strategies for relevance feedback in region-based image retrieval. In *Pacific-Rim Conference on Multimedia 2004* (pp. 497-504).

Fonseca, M. J., & Jorge, J. A. (2003). Indexing high-dimensional data for content-based retrieval in large database. In *Proceedings of the Eighth International Conference on Database Systems for Advanced Applications* (pp. 267-274).

Gevers, T., (2001). Color-based retrieval. In M. S. Lew (Ed.), *Principles of visual information retrieval* (pp. 11-49). London: Springer.

Gevers, T., & Stokman, H. (2003). Classifying color edges in video into shadow-geometry, highlight, or material transitions. *IEEE Transactions on Multimedia, 5*(2), 237-243.

Glatard, T., Montagnat, J., & Magnin, I. E. (2004). Texture based medical image indexing and retrieval: Application to cardiac imaging. In *Proceedings of the ACM SIGMM International Workshop on Multimedia Information Retrieval* (pp. 135-142).

Gonzalez, R. C., & Woods, R. E. (2002). *Digital image processing*. Upper Saddle River, NJ: Prentice Hall.

Guan, H. & Wada, S. (2002). Flexible color texture retrieval method using multi-resolution mosaic for image classification. In *Proceedings of the 6th International Conference on Signal Processing: Vol. 1* (pp. 612-615).

Guttman, A. (1984). R-trees: A dynamic index structure for spatial searching. In *Proceedings of ACM SIGMOD International Conference on Management of Data* (pp. 47-54).

Haeghen, Y. V., Naeyaert, J. M. A. D., Lemahieu, I., & Philips, W. (2000). An imaging system with calibrated color image acquisition for use in dermatology. *IEEE Transaction on Medical Imaging, 19*(7), 722-730.

Haralick, R. M. (1979). Statistical and structural approaches to texture. In *Proceedings of the IEEE, 67*(5), 786-804.

Haralick, R. M., Shanmugan, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics, 3*(6), 610-621.

Hoi, C.-H., Chan, C.-H., Huang, K., Lyu, M. R., & King, I. (2004). Biased support vector machine for relevance feedback in image retrieval. In *Proceedings of International Joint Conference on Neural Networks* (pp. 3189-3194).

Hsu, C.-C., & Ho, C.-S. (2004). A new hybrid case-based architecture for medical diagnosis. *Information Sciences, 166*(1-4), 231-247.

Huang, H. K. (2003). PACS, image management, and imaging informatics. In D. Feng, W. C. Siu, & H. J. Zhang (Eds.), *Multimedia information retrieval and management: Technological fundamentals and applications* (pp.347-365). New York: Springer.

Lehmann, T. M., Guld, M. O., Keysers, D, Deselaers, T., Schubert, H., Wein B. B., & Spitzer, K. (2004a). Similarity of medical images computed from global feature vectors for content-based retrieval. *Lecture Notes in Artificial Intelligence*, 989-995.

Lehmann, T. M., Guld, M. O., Thies, C., Plodowski, B., Keysers, D., Ott, B., & Schubert, H. (2004b). IRMA — Content-based image retrieval in medical applications. In *Proceedings of the 14th World Congress on Medical Informatics* (pp. 842-848).

Lehmann, T. M., Wein, B. B., & Greenspan, H. (2003). Integration of content-based image retrieval to picture archiving and communication systems. In *Proceedings of Medical Informatics Europe*.

Li, C.-T. (1998). *Unsupervised texture segmentation using multi-resolution Markov random fields*. Doctoral dissertation, University of Warwick, Coventry, UK.

Li, C. H., & Yuen, P. C. (2000). Regularized color clustering in medical image database. *IEEE Transaction on Medical Imaging, 19*(11), 1150-1155.

Majumdar, S., Kothari, M., Augat, P., Newitt, D. C., Link, T. M., Lin, J. C., & Lang, T. (1998). High-resolution magnetic resonance imaging: Three-dimensional trabecular bone architecture and biomechanical properties. *Bone, 22*, 445-454.

Moghaddam, H. A., Khajoie, T. T., & Rouhi, A. H. (2003). A new algorithm for image indexing and retrieval using wavelet correlogram. In *Proceedings of the International Conference on Image Processing 2003: Vol. 3* (pp. 497-500).

Muller, H., Michous, N., Bandon, D., & Geissbuhler, A. (2004a). A review of content-based image retrieval systems in medical applications — Clinical benefits and future directions. *International Journal of Medical Informatics, 73*(1), 1-23.

Muller, H., Rosset, A. Vallee, J.-P., & Geissbuhler, A. (2004b). Comparing features sets for content-based image retrieval in a medical-case database. In *Proceedings of IS&T/SPIE Medical Imaging 2004: PACS and Imaging Informatics* (pp. 99-109).

Nishibori, M. (2000). Problems and solutions in medical color imaging. In *Proceedings of the Second International Symposium on Multi-Spectral Imaging and High Accurate Color Reproduction* (pp. 9-17).

Nishibori, M., Tsumura, N., & Miyake, Y. (2004). Why multi-spectral imaging in medicine? *Journal of Imaging Science and Technology, 48*(2), 125-129.

Partridge, M., & Calvo, R. A. (1998). Fast dimensionality reduction and simple PCA. *Intelligent Data Analysis, 2*(1-4), 203-214.

Qian, G.., Sural, S., Gu, Y., & Pramanik, S. (2004). Similarity between Euclidean and cosine angle distance for nearest neighbor queries. In *Proceedings of 2004 ACM Symposium on Applied Computing* (pp. 1232-1237).

Ouyang, A., & Tan, Y. P (2002). A novel multi-scale spatial-color descriptor for content-based image retrieval. In *Proceedings of the 7th International Conference on Control, Automation, Robotics and Vision: Vol. 3* (pp. 1204-1209).

Rencher, A. C. (1998). *Multi-variate statistical inference and applications*. New York: John Wiley & Sons.

Rocchio, J. J. (1971). Relevance feedback in information retrieval. In G. Salton (Ed.), *The SMART retrieval system-experiments in automatic document processing* (pp.313-323). Englewood Cliffs, NJ: Prentice Hall.

Rosset, A., Ratib, O., Geissbuhler, A., & Vallee, J. P. (2002). Integration of a multimedia teaching and reference database in a PACS environment. *Radiographics, 22*(6), 1567-1577.

Rui, Y., & Huang, T. (2002). Learning based relevance feedback in image retrieval. In A. C. Bovik, C. W. Chen, & D. Goldfof (Eds.), *Advances in image processing and understanding: A festschrift for Thomas S. Huang* (pp. 163-182). New York: World Scientific Publishing.

Schmidt, R., Montani, S., Bellazzi, R., Portinale, L., & Gierl, L. (2001). Cased-based reasoning for medical knowledge-based systems. *International Journal of Medical Informatics, 64*(2-3), 355-367.

Sebe, N., & Lew, M. S. (2002). Texture features for content-based retrieval. In M. S. Lew (Ed.), *Principles of visual information retrieval* (pp.51-85), London: Springer.

Serdobolskii, V. (2000). *Multivariate statistical analysis: A high-dimensional approach*. London: Kluwer Academic Publishers.

Shyu, C., Brodley, C., Kak, A., Kosaka, A., Aisen, A., & Broderick, L. (1999). ASSERT: A physician-in-the-loop content-based image retrieval

system for HRCT image databases. *Computer Vision and Image Understanding, 75*(1/2), 111-132.

Shyu, C.-R., Pavlopoulou, C., Kak, A. C., Brodly, C. E., & Broderick, L. (2002). Using human perceptual categories for content-based retrieval from a medical image database. *Computer Vision and Image Understanding, 88*(3), 119-151.

Smeaton, A. F., & Over, P. (2003). TRECVID: Benchmarking the effectiveness of information retrieval tasks on digital video. In *International Conference on Image and Video Retrieval*, 18-27.

Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(12), 1349-1380.

Sinha, U., & Kangarloo, H. (2002). Principal component analysis for content-based image retrieval. *RadioGraphics, 22*, 1271-1289.

Squire, D. M., Muller, W., Muller, H., & Pun, T. (2000). Content-based query of image databases: inspirations from text retrieval, *Pattern Recognition Letters, 21*(13-14), 1193-1198.

Squire, D. M., Muller, H., Muller, W., Marchand-Maillet, S., & Pun, T. (2001). Design and evaluation of a content-based image retrieval system. In S. M. Rahman (Ed.), *Design and management of multimedia information systems: Opportunities and challenges* (pp. 125-151). Hershey, PA: Idea Group Publishing.

Suckling, J., Parker, J., Dance, D. R., Astley, S., Hutt, I., Boggis, C. R. M., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S.-L., Taylor, P., Betal, D., & Savage, J. (1994). The mammographic image analysis society digital mammogram database. In *Proceedings of the 2nd International Workshop on Digital Mammography*, 375-378.

Swain, M. J., & Ballard, D. H. (1991). Color Indexing. *International Journal of Computer Vision, 7*(1), 11-32.

Tamai, S. (1999). The color of digital imaging in pathology and cytology. *Proceedings of the First Symposium of the "Color" of Digital Imaging in Medicine*.

Tao, D., & Tang, X. (2004). Random sampling based SVM for relevance feedback image retrieval. In *The IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2004*, 647-652.

Tourassi, G. D. (1999). Journey toward computer-aided diagnosis: Role of image texture analysis. *Radiology*, 317-320.

Veenland, J. F., Grashuis, J. L., Weinans, H., Ding, M., & Vrooman, H. A. (2002). Suitability of texture features to assess changes in trabecular bone architecture. *Pattern Recognition Letters, 23*(4), 395-403.

Veltkamp, R. C., & Hagedoorn, M. (2002). State of the art in shape matching. In M. S. Lew (Ed.), *Principles of visual information retrieval* (pp.87-119), London: Springer.

Wei, C.-H, & Li, C.-T. (2005). Design of content-based multimedia retrieval. In M. Pagani (Ed.), *Encyclopedia of multimedia technology and networking*. Hershey, PA: Idea Group Reference.

Wong, S., & Hoo, K. S. (2002). Medical imagery. In V. Castelli, & L. D. Bergman (Eds.), *Image databases: Search and retrieval of digital imagery* (pp. 83-105). New York: John Wiley & Sons, Inc.

Wong, S. T. C. (1998). CBIR in medicine: Still a long way to go. In *Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries*, 114.

Yu, H., Li, M., Zhang, H.-J., & Feng, J. (2002). Color texture moments for content-based image retrieval. In *Proceedings of the International Conference on Image Processing 2002: Vol. 3* (pp. 929-932).