Development of an Image Retrieval Model for Biomedical Image Databases

Achimugu Philip[1], Babajide Afolabi[2],
Adeniran Oluwaranti[2] and Oluwagbemi Oluwatolani[1]
[1]Department of Computer Science, Lead City University, Ibadan
[2]Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife
Nigeria

1. Introduction

Image processing is the field of signal processing where both the input and output signals are images. Images can be thought of as two-dimensional signals via a matrix representation. Image processing a is very important subject, and finds itself in such fields as photography, satellite imaging, medical imaging, and image compression, to name but a few. In the past, image processing was largely done using analog devices (Cheng et al., 2006). However, as computers became more powerful, processing shifted toward the digital domain. Like one-dimensional digital signal processing, digital image processing overcomes traditional analog "problems" such as noise, distortion during processing, inflexibility of system to change, and difficulty of implementation.

Image retrieval, popularly referred to as content-based image retrieval is an emerging technology that allows a user to retrieve relevant images in an effective and efficient manner. Digital imaging has extensive applications in our daily lives and it is being used for several applications. Examples of imaging applications are in museums for archiving important images and manuscripts from art gallery and museum management. Many useful applications of imaging are found in security for tracking an intruder, crime prevention, law enforcement and object recognition in digital forensic.

In particular, image retrieval is potentially useful in discovering brain activation patterns, in classifications and in diagnoses by comparing observed patterns with those of known diseases, leading to clinical applications. In biomedicine, content-based image retrieval is critically important in patient digital libraries, clinical diagnosis, clinical trials, searching for 2-D electrophoresis gels, and pathological slides. Most existing content-based image retrieval systems (Flickner et al., 1995; Gupta and Jain, 1997; Ma and Manjunath, 1997; Rubner, 1999 and Wang et al., 1998) are designed for general purpose picture libraries such as photos and graph.

The storage, manipulation and analysis of the contents of digital images are essential requirements for the next generation of healthcare information infrastructure. The aim of this infrastructure is to bring timely health information to support communication among healthcare decision makers and communities at large. Among several healthcare services that can be provided with the aid of the emerging grid technology for ubiquitous access, image classification and diagnosis services are important. The ubiquitous access and

retrieval services enable the storage, retrieval, analysis, management, manipulation and sharing of all kinds of healthcare specific digital images. The healthcare community is currently exploring collaborative approaches for managing image data and exchanging useful knowledge. It also eases distributed management of clinical data and scenarios for integration of image retrieval access methods into picture archiving and communication systems as well as healthcare information infrastructure. However, retrieval results of the existing image retrieval systems are generally not satisfactory due to the weak connection between low level image features and high level image semantics. Moreover, traditional text-based description requires images to be manually annotated. This can be a very time-consuming task that is cumbersome, error prone and prohibitively expensive. Additionally, images can have contents that texts alone cannot adequately convey. Thus, due to the rich content of images, traditional text-based retrieval methods can be complemented with efficient and effective image retrieval algorithms and techniques to enhance medical diagnosis and therapy planning.

2. What is digital image processing?

An *image* as defined in the "real world" is considered to be a function of two real variables, for example, a(x, y) with **a** as the amplitude (e.g. brightness) of the image at the real coordinate position (x, y). Furthermore, an image may be considered to contain sub-images sometimes referred to as regions-of-interest, ROIs, or simply regions. This concept reflects the fact that images frequently contain collections of objects each of which can be the basis for a region.

Therefore, *Digital image processing* is the use of computer algorithms to perform image processing on digital images. As a subfield of digital signal processing, digital image processing has many advantages over analog image processing; it allows a much wider range of algorithms to be applied to the input data, and can avoid problems such as the build-up of noise and signal distortion during processing.

2.1 What can be done by image processing?
a. Geometric transformations such as enlargement, reduction, and rotation.
b. Color corrections such as brightness and contrast adjustments, quantization, or conversion to a different color space.
c. Registration (or alignment) of two or more images.
d. Combination of two or more images, e.g. into an average, blend, difference, or image composite.
e. Interpolation and recovery of a full image from a RAW image format.
f. Segmentation of the image into regions.
g. Image editing and Digital retouching.
h. Extending dynamic range by combining differently exposed images.

2.2 Applications of image processing
Image Processing finds applications in the following areas:
a. Photography and Printing
b. Satellite Image Processing
c. Medical Image Processing

    d.   Face detection, Feature detection, Face identification
    e.   Microscope image processing

3. Digital image basics

When using digital equipment to capture, store, modify and view photographic images, they must first be converted to a set of numbers in a process called digitization or scanning. Computers are very good at storing and manipulating numbers, so once your image has been digitized you can use your computer to archive, examine, alter, display, transmit, or print your photographs in an incredible variety of ways.

Digital images are composed of *pixels* (short for picture elements). Each pixel represents the color (or gray level for black and white photos) at a single point in the image, so a pixel is like a tiny dot of a particular color. By measuring the color of an image at a large number of points, we can create a digital approximation of the image from which a copy of the original can be reconstructed. Pixels are a little like grain particles in a conventional photographic image, but arranged in a regular pattern of rows and columns and store information somewhat differently. A digital image is a rectangular array of pixels sometimes called a *bitmap*.

For photographic purposes, there are two important types of digital images; color and black and white. Color images are made up of colored pixels while black and white images are made of pixels in different shades of gray.

A black and white image is made up of pixels each of which holds a single number corresponding to the gray level of the image at a particular location. These gray levels span the full range from black to white in a series of very fine steps, normally 256 different grays. While a color image is made up of pixels each of which holds three numbers corresponding to the red, green, and blue levels of the image at a particular location. Red, green, and blue (sometimes referred to as RGB) are the primary colors for mixing light; these so-called additive primary colors are different from the subtractive primary colors used for mixing paints (cyan, magenta, and yellow). Any color can be created by mixing the correct amounts of red, green, and blue light. Assuming 256 levels for each primary, each color pixel can be stored in three bytes (24 bits) of memory. This corresponds to roughly 16.7 million different possible colors. Note that for images of the same size, a black and white version will use three times less memory than a color version.

3.1 Binary or Bi-level Images
Binary images use only a single bit to represent each pixel. Since a bit can only exist in two states; on or off, every pixel in a binary image must be one of two colors, usually black or white. This inability to represent intermediate shades of gray is what limits their usefulness in dealing with photographic images.

3.2 Indexed color images
Some color images are created using a limited palette of colors, typically 256 different colors. These images are referred to as indexed color images because the data for each pixel consists of a palette index indicating which of the colors in the palette applies to that pixel. There are several problems with using indexed color to represent photographic images. First, if the image contains more different colors than are in the palette, techniques such as dithering must be applied to represent the missing colors and this degrades the image. Second,

combining two indexed color images that use different palettes or even retouching part of a single indexed color image creates problems because of the limited number of available colors.

3.3 Resolution

The more points at which we sample the image by measuring its color, the more detail we can capture. The density of pixels in an image is referred to as its *resolution*. The higher the resolution, the more information the image contains. If we keep the image size the same and increase the resolution, the image gets sharper and more detailed. Alternatively, with a higher resolution image, we can produce a larger image with the same amount of detail and the resolution of an image is reduced no matter the size; the pixels get larger and larger and there is less and less detail in the image.

4. Image digitalization

The process of converting an image to pixels is called *digitizing* or *scanning* and this function can be performed in many different ways as described below:

4.1 Film scanners

This type of scanner is sometimes called a slide or transparency scanner. They are specifically designed for scanning film, usually 35mm slides or negatives, but some of the more expensive ones can also scan medium and large format film. These scanners work by passing a narrowly focused beam of light through the film and reading the intensity and color of the light that emerges.

4.2 Flatbed scanners

This type of scanner is sometimes called a reflective scanner. They are designed for scanning prints or other flat, opaque materials. These scanners work by shining white light onto the object and reading the intensity and color of the light that is reflected from it, usually a line at a time. Some flatbed scanners have available transparency adapters, but for a number of reasons, in most cases these are not very well suited to scanning film. On the other hand, flatbed scanners can be used as a sort of lensless camera to directly digitize flat objects like leaves.

4.3 Digital cameras

One of the most direct ways to capture an image is a digital camera which uses a special semiconductor chip called a CCD (charge coupled device) to convert light to electrical signals right at the image plane. The quality of the images created in this manner is closely related to the number of pixels the CCD can capture. Affordable digital cameras suffer from relatively low resolution, limited dynamic range, and low ISO film speed equivalent, and consequently do not always produce high quality digital images. To get images with quality comparable to film photography currently requires very expensive digital cameras.

4.4 Video frame grabbers

This type of scanner uses a video camera to capture a scene or object and then converts the video signal that comes out of the camera to a digital image in your computer memory. A

video camera can be used to digitize scenes containing 3- dimensional objects, but they usually have much lower image quality than film or flatbed scanners.

4.5 Scanning services

Photo CD is a service started by Kodak a number of years ago whereby your film can be scanned using a high quality scanner and written to a compact disk your computer can access. Using Photo CD service is an inexpensive way to get high quality scans of your images without purchasing a scanner. Many other scanning services are available which can scan prints or film to floppy disks or removable disk cartridges. These vary from low resolution snapshot quality images to professional drum scans at very high resolution.

5. Image compression

Image compression is the application of data compression on digital images. In effect, the objective is to reduce redundancy of the image data in order to be able to store or transmit data in an efficient form. Image compression can be lossy or lossless. Lossless compression is sometimes preferred for artificial images such as technical drawings, icons or comics. This is because lossy compression methods, especially when used at low bit rates, introduce compression artifacts. Lossless compression methods may also be preferred for high value content, such as medical imagery or image scans made for archival purposes. Lossy methods are especially suitable for natural images such as photos in applications where minor (sometimes imperceptible) loss of fidelity is acceptable to achieve a substantial reduction in bit rate. Compressing an image is significantly different than compressing raw binary data. Of course, general purpose compression programs can be used to compress images, but the result is less than optimal. This is because images have certain statistical properties which can be exploited by encoders specifically designed for them. Also, some of the finer details in the image can be sacrificed for the sake of saving a little more bandwidth or storage space. This also means that lossy compression techniques can be used in this area. The image compression technique that is most often used is transformed coding. A typical image's energy often varies significantly throughout the image, which makes compressing it in the spatial domain difficult; however, images tend to have a compact representation in the frequency domain packed around the low frequencies, which makes compression in the frequency domain more efficient and effective. Transform coding is an image compression technique that first switches to the frequency domain, then does it's compressing. The transform coefficients should be de-correlated to reduce redundancy and to have a maximum amount of information stored in the smallest space. These coefficients are then coded as accurately as possible to not lose information.

6. Statement of problem

The retrieval results of the existing image retrieval systems are generally not satisfactory due to the weak connection between low-level features and the semantics of images. This problem is generally referred to as the semantic gap problem. Image retrieval approaches based on low level concepts naturally have some inherited problems for human perceptual recognition. Humans recognize images based on high-level concepts such as texts and they typically query images by their semantics. Alternatively, high level concepts alone are not sufficient for image retrieval, because images can contain important information that texts

cannot reveal. Another very important problem in image retrieval is the need for suitable similarity measures and image representations that can lead to an improved retrieval result. In particular, variations can occur among semantically similar objects in medical images such as MRI, x-ray, ultrasound, digital radiography, etc; which have direct reliance on medical diagnosis and intervention. These variations can cause serious problems for an image representation method, making it difficult to conceive a measure for similarity in image retrieval (Olugbara, 2008). Therefore, it would be pertinent to develop a database processing models or schemes for improving retrieval results in healthcare applications; hence this study.

7. Related works

A lot of opinions and methods have been developed in solving problems that relates image processing and computer graphics. This has led to extensive research that resulted, in less than a decade, into numerous commercial and research based system, including QBIC (Niblack, et al., 1993), Virage (Bach, et al., 1996), Photobook (Pentland, et al., 1996), Excalibur (Feder, 1996), Chabot (OVE and Chabot, 1995), and Visual SEEK (Smith, 1997). These systems allow users to formulate queries using combinations of low-level image features such as colour, texture and shape. The queries are specified explicitly by providing the desired feature values or implicitly by specifying an example image. Some system also uses the spatial organization of the image features, so that similarity is determined not only by the existence of certain features but also by their absolute or relative location in the image (Hou, et al., 1992; Bach, et al., 1996; Stricker and Dimai, 1996; Smith, 1997; Das, et al., 1997; Cohen, 1999).

Early systems focused on the search engine (given a query, find the best matches) and did not use precious queries to understand better what the user was looking for. Recent system allows the user to redefine the search by indicating the relevance (or irrelevance) of images in the returned set. This is known as relevance feedback (Rui, et al., 1997; Minka and Picard, 1997; Smith, 1997).

In (Idris and Panchanathan, 1997), the authors provide an in-depth review of content-based image retrieval systems. They also identify a number of unanswered key research questions, including the development of more robust and compact image content feature and dissimilarity measures that model perceptual similarity more accurately.

(Michalski, 1972; 1973) examined how symbolic AQ rule learning could be used for discrimination between textures or between simple structures. These seminal papers presented the Multi-Level Logical Template (MLT) methodology in which windowing operators scanned an image and extracted local features. These features were used to learn rules describing textures (or simple structures); the rules were then used for texture (or simple structure) recognition.

(Shepherd, 1983), encoded examples as feature vectors, learned decision trees for an industrial inspection task especially, classification of the shapes of chocolates. Comparisons of classification accuracy were made between decision tree, k-nearest neighbor (k-nn), and minimum distance classifiers. Experimental results for these classifiers were similar, with the minimum distance classifier producing the highest accuracy, 82%.

(Channic, 1989) extended the MLT methodology in (Michalski, 1972; 1973) by using convolution operators in conjunction with the original set of windowing operators for feature extraction. Using the AQ learning system, Channic investigated incremental

learning and iterative learning from sequences of images using ultrasound images of laminated objects.

Instead of representing examples using feature vectors, (Connell and Brady, 1987) learned generalized semantic networks from images of classes of hammers and of overhead views of commercial aircraft. Training examples were generated by a vision system that took gray scale images as input and produced semantic networks for the objects; a learning system, which was a modified version of (Winston, 1984) ANALOGY program, learned by generalizing the training examples. The learning system was extended to learn disjunctive concepts and to learn from only positive examples. These generalized representations were used to classify unknown objects.

(Cromwell and Kak, 1991) proceeded as Shepherd did, using feature vectors to characterize shapes. Electrical component shapes were learned using a symbolic induction methodology based on that developed by Michalski. They reported that their method achieved 72% on testing data, but no comparisons were made with other learning methods.

(Pachowicz and Bala, 1991) also used the MLT methodology, following (Michalski, 1972; 1973) and (Channic, 1989), but added a modified set of Laws' masks for texture feature extraction. They also applied techniques for handling noise in symbolic data. These techniques included optimizing learned symbolic descriptions by truncating rules, as well as removing training examples covered by weak rules and relearning. The PRAX method for learning a large number of classes was introduced by Bala, Michalski, and Wnek in 1992 and 1993.

(Segen, 1994) used a hybrid shape representation consisting of a hierarchical graph that takes into account local features of high curvature, and the angles and distances between these local features. This representation is invariant to both planar rotation and translation. Shapes were silhouettes of hand gestures. Segen's system runs in real time and has been applied to airplane simulator control as well as to control of a graphics editor program. Error rates were between 5% and 10%, but most errors were unknowns rather than misclassifications.

(Cho and Dunn, 1994) described a new learning algorithm for learning shape. This algorithm memorizes property lists and updates associated weights as training proceeds. Forgetting mechanisms remove useless property lists. Shapes are modeled by a series of line segments. Using the orientations of these segments, local spatial measures are computed and form a property list for a shape. The system was used to classify tools and hand gestures and achieved predictive accuracies of 92% and 96% on these problems.

(Dutta and Bhanu, 1994) presented a 3D CAD-based recognition system in which genetic algorithms are used to optimize segmentation parameters. Qualitative experimental results were presented for indoor and outdoor motion sequences in which the system recognized images of wedges (traffic cones) and cans from gray scale and depth map images. Sung and (Poggio, 1994) worked on automatic human face detection. An example-based learning approach was tested for locating un-occluded frontal views of human faces in complex scenes. The space of human faces was represented by a few \face" and \non-face" pattern prototypes. At each image location, a two-valued distance measure was computed between the local image pattern and each prototype. A trained classifier was used to determine whether a human face is present. The authors showed that their distance metric is critical for the success of their system.

(Zheng and Bhanu, 1996) examined how Hebbian learning mechanisms could be used to improve the performance of an image thresholding algorithm for automatic target detection

and recognition. Qualitative results were presented in which the adaptive thresholding algorithm was shown to be superior to the classical thresholding algorithm for both SAR and FLIR images.

(Rowley, et al., *1996*) built a neural network-based face detection system by using a retinally connected neural network to examine small windows of an image and decide on the existence of a face. A bootstrap algorithm was implemented during training so as to add false detection into the training set and as a consequence, eliminate the difficult task of manually selecting non-face training examples. Experimental results showed better performance in terms of detection and false-positive rates.

(Romano, et al., 1996) built a real-time system for face verification. Experiments showed that simple correlation strategies on template-based models are sufficient for many applications in which the identity of a face in a novel image must be verified quickly and reliably from a single reference image. The authors suggested that this automatic real-time face verification technique could be put to use in such human-machine interface applications as automated security systems. The technique has been integrated into a screen locking application which permits access to workstations by performing face verification in lieu of password authentication.

The MLT methodology developed by Michalski in 1973 and 1973 has recently been extended into the Multi-Level Image Sampling and Transformation (MIST) methodology. MIST has been applied to a variety of problems including natural scene segmentation and identification of blasting caps in x-ray images. For classifying natural scenes, three learning techniques were compared: AQ15c, a back propagation neural network, and AQ-NN.

AQ-NN is a multi-strategy learning technique in that it uses two different representations and two different learning strategies. Specifically, the AQ learning algorithm is used to learn attributional decision rules from training examples. These decision rules are then used to structure a neural network architecture. A back propagation algorithm is then used as a learning step to further optimize the AQ induced descriptions. In such a system, learning times and recognition rates are often significantly decreased, while predictive accuracy is improved, with respect to conventional neural network learning. To learn classes such as ground, grass, trees and sky, hue, intensity, and convolution operators are used to extract features from a user-designated training area. These examples are then presented to the learning system, which induces a class description.

AQ15c, used alone, achieved a predictive accuracy of 94%, while AQ-NN and a standard neural network achieved predictive accuracies near 100%. The training time of AQ-NN was approximately two orders of magnitude shorter than the training time of the standard NN.

8. Our approach

An environment has been developed and implemented which support the interactive processing, classification and browsing of medical images. Such an environment can be used over an existing Image DataBase (IDB) system for the purpose of assisting the interaction between the user and the IDB. An "Image DataBase" (IDB) is a "system in which a large amount of image data and their related information are integratedly stored". Important considerations in the design and implementation of IDB systems are: image feature extraction, image content representation, organization of stored information, search and retrieval strategies, as well as user interface design. In particular, the following were taking into consideration when designing the system:

8.1 Processing and segmentation of medical images

Before any required image descriptions are extracted and used (e.g., stored or matched with others stored in the IDB), images must first be segmented into disjoint regions or objects (Petrakis and Orphanoudakis, 1992). Segmentation refers to the process of partitioning a digital image into multiple segments (sets of pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (edge detection). Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s). Furthermore, image segmentation help compute a variety of image features specific to a particular image representation. Therefore, the active contours method was used for segmenting images in this research because of its protrusious ability and specific topological effect.

8.2 Classification of medical images

In the classification procedure, each test image is represented using the local features selected. To approximate the posterior probability that a certain local features image belongs to an image in a given class, the *k-nearest* neighbour algorithm was used. This posterior probability is then used to obtain a combined decision for the set of local features. Therefore, images were pre-classified into high level semantic categories like graph or photograph, texture or non-texture, which are relatively simple to classify. After this classification, the retrieval system returns only the images belonging to the same semantic categories.

8.3 Efficient Image Database (IDB) organization

A fast, simple and low complexity technique for image database organization is what is required. The basic idea is to reveal the connectivity relations of the database in order to obtain information of the database structure and facilitate the clustering processes. This will mean partitioning the image database into segments that reduces the search space thereby enhancing faster retrieval since a query addresses only a specific database segment as depicted in Figure 1. It was achieved by randomly selecting a certain number of prototype data and using appropriately the membership values of the rest data points to the selected prototypes. The clustering was easily performed in the final step by using graph theory methodology.

8.4 Retrieval of medical images

Content-based image retrieval (CBIR), a technique that uses visual contents to search images from large scale image databases according to users' interests, has been an active and fast advancing research area since the 1990s. During the past decade, remarkable progress has been made in both theoretical research and system development. Figure 2 shows the general model used in CBIR.
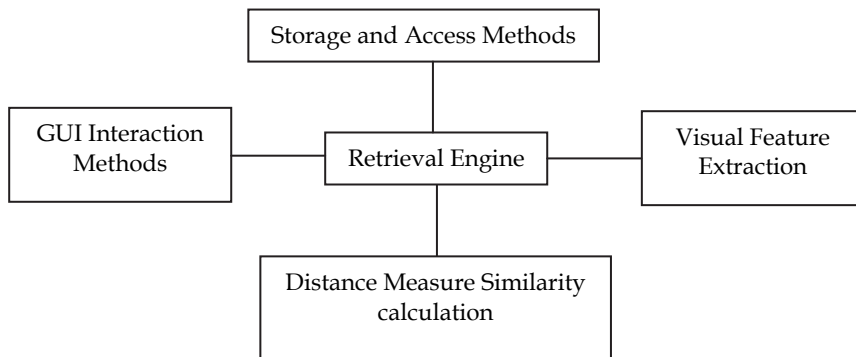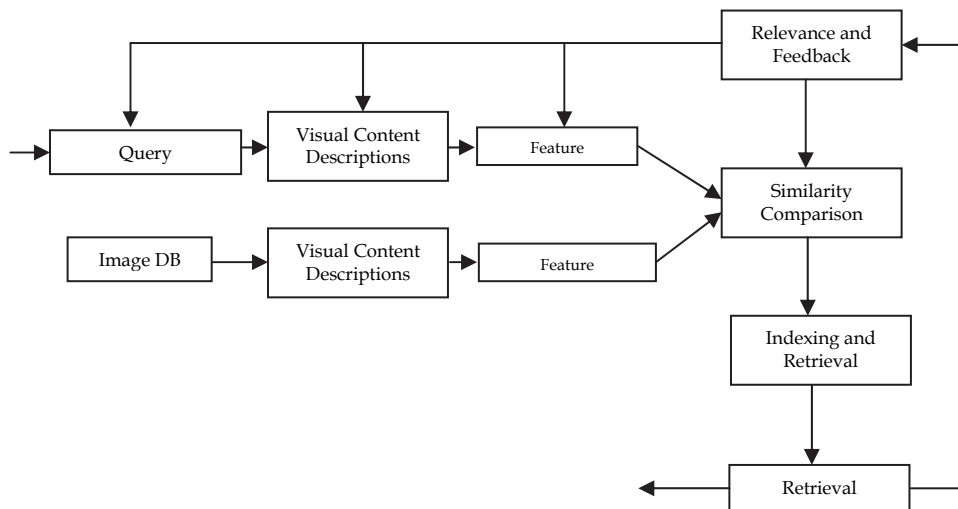
Fig. 1. Principal components of a CBIR system

Fig. 2. Diagram of content based image retrieval system

Before the emergence of content-based retrieval, medical images were annotated with text, allowing the images to be accessed by text-based searching. 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:

a.   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.

b.   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.

c.   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.

d.   Low resolution and strong noise are two common characteristics in most medical images. With these characteristics, medical images cannot be precisely segmented and extracted for the visual content of their features.

For the purpose of this research however, queries can be specified either by a conditional statement indicating various attributes or by specifying an example image (grey level image or sketch). Therefore, a user may point or specify the actual database segment to be searched for and all images satisfying the query criteria will be retrieved. A user is then allowed to make a final selection by browsing the retrieved images in order to determine the correctness (precision) of the retrieved image.

8.5 Browsing the IDB

The main goal of healthcare information infrastructures is to provide the needed information on time, at the right place and to the right persons so as to improve the quality and efficiency of care processes. (Olugbara, 2008). Therefore, in order to achieve this goal, the user interface is loaded with friendly graphical tools and mechanisms, called the "image browser" for both displaying and selecting image classes as well as class their respective properties and schemas.

9. Image representation, indexing, storage and retrieval

Quantized images are commonly represented as sets of pixels encoding color or brightness information in matrix form. An alternative model is based on contour lines. A contour representation enhances easy retrieval of images in a bitmap form. It is primarily used for data compression of an image. The idea is to encode, for each level, the boundaries of connected regions of pixels and to reconstruct an original image from those boundaries. One problem is how to store such representations in a compact manner. In practice, one rarely needs the entire contour representation. An image subset is considered to be the basic entity in the proposed indexing scheme. On the other hand, images are indexed based on representations of the set of all the derived subsets. The objects contained in each group are first ordered. Ordering must be based on criteria that clearly differentiate the objects among them. Position is such a criterion, since objects are usually scattered in the image. Size or shape does not necessarily provide good ordering criteria, since images may contain similar objects. Each of these ordered subsets is then represented by a set of attribute strings corresponding to the set of properties involved in a particular image representation. An individual image subset is indexed by treating each of its corresponding attribute strings as a separate key.

Furthermore, representing image contents is indeed the first problem to face when defining an automatic annotation system. There are different strategies to extract features from images depending on which is considered the most relevant information to capture. As the x–ray images do not contain any color information, edge, shape and global texture features play an important role in this task and were used by several groups (Bo et al., 2005; Deselaers et al., 2005; Liu et al., 2006). Various methods used the pixel values directly and accounted for possible deformation of the images (Image Distortion Model, IDM) (G¨uld et al., 2005; Deselaers et al., 2005). Approaches coming from the object recognition field mostly followed the currently widely adopted assumption that an object in images consists of parts that can be modeled independently. Thus these methods considered local features extracted around interest points and used a wide variety of bag-of-features approaches (Mar´ee et al., 2005; Liu et al., 2006; Tommasi et al., 2007; Avni et al., 2008). Generally the ordering of the visual words is not taken into account and only the frequency of the individual visual word is used to form the feature vectors. However, some groups added the spatial information to patches extracted from images (Deselaers et al., 2006; Avni et al., 2008) after observing that radiographs of a certain body part are typically taken in the same spatial arrangement. Another widely adopted strategy consists of combining different local and global descriptors into a unique feature representation (Bo et al., 2005; Liu et al., 2006; Tommasi et al., 2008a).

9.1 Indexing biomedical images

The purpose of content-based indexing and searching for biomedical image database is now an important subject in the field of image processing. This is because; users of a biomedical image database are often interested in images with similar objects at the finest scale. Therefore, various indexing techniques were reviewed and analyzed for suitable use. This has to do with selecting the rules that form the basis of the annotation process. Many different indexing strategies were applied and while in the earlier years nearest neighbour-based approaches were most common and most successful; for example, (Deselaers et al., 2005; G¨uld et al., 2005), in 2006 and later, discriminative approaches such as log–linear models (Deselaers et al, 2006), and decision trees (Setia et al, 2008), as well as Support Vector Machines (Setia et al, 2008; Tommasi et al, 2008a; Avni et al, 2008) became more and more common and outperformed the nearest neighbor–based approaches. The training images are used as the image database and the test images are used to query it. For each query, the training images are ranked according to their similarity and the nearest neighbor decision rule is applied, that is, the class of the most similar training image is chosen for every test image.

In this research however, our aim is to propose a system that will be able to link an image or image region with a semantic label that combines the Unified Medical Language Systems (UMLS) concepts and visual percepts using statistical learning. In this way, we have a common language for both images and the associated text. Two complementary indexing schemes were used within this learning framework; that is, a global indexing to access image modality and a local indexing to access semantic local features that are related to modality, anatomy, and pathology concepts.

In our model, a set of disjoint UMLS-based concepts that do -not share common visual instances with visual appearance in medical images is first selected to define a Visual and Medical vocabulary. Secondly, low-level features are extracted from image region instances

z to represent each Visual and Medical terms using color, texture and shape. Thirdly, these low-level features are used as training examples to build hierarchical semantic classifiers according to the Visual Medical vocabulary. The classifier for the Visual Medical vocabulary is designed using Support Vector Machine (SVM) classifiers. The conditional probability that an example *z* belong to a class *c* given that the class belongs to its super class *C* is computed using the softmax function as follows:

$$P(c \mid z, C) = \frac{\exp D_c(z)}{\sum_{j \in C} \exp D_j(z)} \qquad\qquad \text{i}$$

Where, *Dc* is the signed distance to the SVM hyper plane that separates class *c* from the other classes in *C*. The probability of a Visual Medical term VMT*i* for *z* is:

$$P(VMT_i \mid z) = \prod_{l=1}^{L} P(C^l(VMT_i) \mid z, C^{l-1}(VMT_i)) \qquad\qquad \text{ii}$$

Where, *L* is the number of levels, $C^{l-1}(VMTi)$ is the super class of $C^l(VMTi)$, $C^0(VMTi)$ is the class for all Visual Medical terms, and $P(C^l(VMTi) \mid z, C^{l-1}(VMTi))$ is given by Equation (Lim et al., 2007).
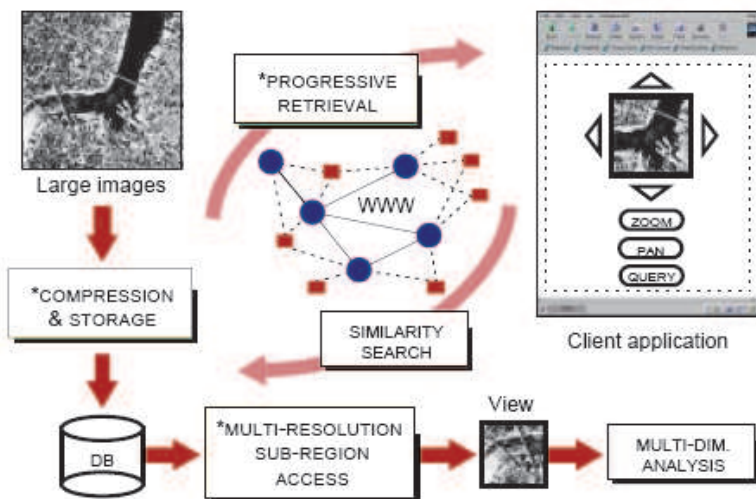
9.2 Image storage

Biomedical images are crucial asset for biomedical research and medical practice. There are multitude of devices for biomedical image acquisition that range from simple; for example, a digital camera coupled with a conventional optical microscope, to complex; for example, specialized equipment for Positron Emission Tomography (PET). These devices are routinely used in the daily medical practice and biomedical research, generating a continuous stream of images. The great majority of these images are digital and a good amount of them are permanently stored in digital image repositories. These image collections are a potential source of information and knowledge. However, the realization of this potential requires effective mechanisms for image storage and retrieval, image analysis, image-collection analysis and image-collection exploration.

Image data can be distinguished into "physical" and "logical" (Chang, 2006). Original (grey level) images, segmented images, and image miniatures are physical images. On the other hand, image related data (information extracted from images, attributes, text etc.) are logical images. Physical and logical images are stored separately in a physical and a logical database respectively (Petrakis and Orphanoudakis, 1992). Pointers are implemented from the logical to the physical database. To reduce storage requirements, physical images are compressed prior to storage and conversely decompressed upon retrievals. Therefore, images are stored in clusters based on the likelihood of being retrieved together in response to a particular query (for example, the set of all images corresponding to a patient's examination may be stored close together on disc). The logical database consists of a set (H_2 …, H_{Kmax}) of relational tables, where K_{max} is the maximum size of image subsets under consideration. The specification of value K_{max} depends on the application. The purpose of each H_k table is to resolve queries which specify *k* objects. Therefore, K_{max} can be set equal to the maximum number of objects allowed in queries, if such a value can be specified in advance. In general, K_{max} may take any value greater than or equal to 2. Typically, the number of objects specified in image queries is not greater than 6. Therefore, we set $K_{max} = 6$.

The image subsets of size k, $2 \leq k \leq K_{max}$, together with their representations are all stored in table H_k. Each image subset is represented by a tuple of one dimensional strings of the form: $(p, r, w, \ldots)$ where p is the ordered sequence of object indices, r is its corresponding rank string representing the "left/right" and "below/above" relationships between objects, w is the inclusion string, and the remaining strings correspond to properties of individual objects (Petrakis and Orphanoudakis, 1992). Two such strings have been used, s to encode the size property and c to encode the roundness property of those objects whose indices are in p. Therefore, the representation of an image subset is given a tuple of the form (r, w, s, c). Indexing is performed by creating a secondary index for each attribute string. In general, the logical database may be considered as consisting of a set $(H, H_1, H_2 \ldots H_{Kmax})$ of relational tables. Table H stores information about each image as a whole and has attributes such as: image file name, dates, names, text descriptions, image header information etc. Table H_1 on the other hand, can be used to store representations of the shape of objects such as those proposed in (Tsai and Yu, 1985; Stien and Madioni, 1990).

Figure 3 demonstrate how medical images are stored, compressed and retrieved; which means that, the application has to provide access to multi-resolution of sub-region views of images. This requires that the system be able to efficiently extract the views from the stored compressed data. The system typically provides query facilities such as content based retrieval methods that allow the users to search for images in the database.

Source: (Smith et al., 1999)

Fig. 3. Large Image Storage and Retrieval Environment

9.3 Image retrieval

Image Database Management Systems (IDBMS) aim to store large collections of image data, and to support efficient content-based retrieval of these data. Like any other Database Management Systems (DBMS), an IDBMS has a storage facility and the data are arranged according to some type of data model. It also has a query facility where images are retrieved according to the queries. All queries address the logical database rather than the raw

image data stored in the physical database. Generally the selection queries are used. For example, selecting a few images from the database, based on their content. Imagine a query facility that loads all the images from the database and checks each image to see if it fits the selection criterion. In all but extreme cases this query facility would be inefficient. The only other option is to have more information about the images stored in the database so it can be used for evaluating the selection criterion. This is known as indexing, discussed from the previous section. When an image data is stored in the database, an index is constructed and stored as well for that data. Queries are executed using only the information in the indexes. Images are retrieved only after confirming that the index conforms to the selection criterion. This means that any content information queried must be in the index. Therefore, the type of information in the index limits the expressiveness of any IDBMS query language. Efficient processing of queries in a database requires sophisticated indexing techniques. Content-based querying of medical image databases requires development of indexing techniques that are significantly different than traditional DBMSs. In general, two types of indexes occur: text indexes to facilitate queries of the type "Find all the images that are labeled as 'ear' and content similarity indexes that facilitate queries of the type "Find all the images that are similar, in color, to this sample image". When an IDMBS is to be developed, some extra data types (arrays, images, etc.) are required that are not in the standard relational data model. Therefore, what is needed is an object relational data model with the capacity to store large uninterrupted images.

10. Results

Evaluations have been carried out based on a number of test queries addressing a prototype IDB storing 30 medical images. Queries are distinguished based on the number m of objects they specify. To obtain average performance measures, for each value of m ranging from 2 to 6, 5 image queries have been used and the average performance to queries specifying an equal number of objects has been computed. Measurements of both the answer set (percentage of images returned with respect to the total number of images stored) and of the retrieval response times have been taken and shown in Figure 4. Queries become more specific and the size of an answer set decreases as the number of objects contained in queries increases. Therefore, response time decreases with the number of query objects since search space and thus the amount of information to be processed decreases too.
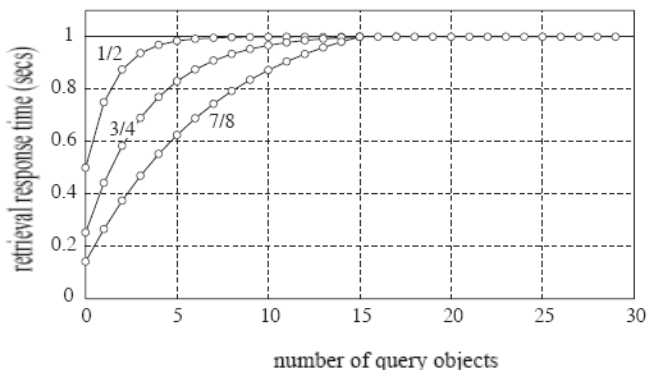

Fig. 4. Number of query objects Vs response time

11. Discussion

An IDB system has been described which supports the efficient processing, archiving, and retrieval of medical images by content. The system consists of an interactive IDB environment, which supports the communication between the user and the various system components, and a methodology which supports the automated indexing and retrieval of images. Such tools and methodologies will be integrated into a prototype IDB system which is currently under development. The system can be easily extended with additional features and mechanisms facilitating the processing and accessing of image data. For instance, the user interface may be extended with additional tools and mechanisms for image registration and image processing, as well as a powerful query language supporting various types of image queries, in addition to queries by example (e.g., queries by specifying an identifier, a class, a hierarchy, range queries etc.), such as those proposed in (Petros and Stelios, 1992). Furthermore, the proposed methodology for image indexing can be easily extended to include the indexing of image sequences in three or four dimensions, where the fourth dimension is time. In developing an IDB system which supports the efficient management of various kinds of image data, is extensible and easy to use, we have to choose an appropriate data model and a database management system (DBMS) which provides persistent storage of both the model and the data, as well as mechanisms for defining, creating, modifying and accessing the model and the data. Furthermore, it must provide a query language, transaction management, concurrency control and authorization for a multiuser environment, as well as performance features such as secondary indexing and clustering.

Such features and mechanisms are inherent within a relational DBMS. Besides, a relational DBMS offers the most direct way of implementing the logical database proposed in this work. However, in order for a DBMS to be appropriate for IDB work it must be augmented with semantic data modeling concepts (e.g., class definition and hierarchies) to assist application modeling. In particular, in developing an IDB system which satisfies the need for a hierarchical database organization, takes advantage of the property of inheritance, and is extensible, the object-oriented approach seems to be more appropriate. All database entities (e.g., various types of data) will be defined as either primitive or complex objects, while system functions (e.g., image processing and retrieval functions) will be defined as methods encapsulated within the same representation with the above database entities. Extensions of relational data models and DBMS's with object-oriented characteristics do exist and can be used to develop an IDB system which satisfies our needs.

However, the main challenge of this system is scalability. Therefore, further research work can focus on developing a system that can accommodate more images.
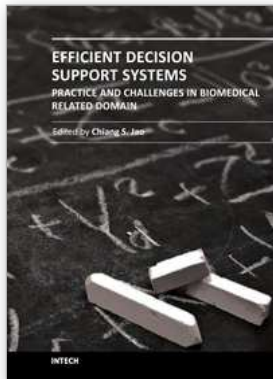
12. Conclusion

In this research, we have described a method for storage, retrieval and manipulation of digital medical images by content. The system consists of an interactive IDB environment, which supports the communication between the user and the various system components which supports the automated indexing and retrieval of images. Our analytical and empirical results have shown the effectiveness of the image retrieval system. The work also presented a possible direction for further future research.

13. References

Avni, U., Goldberger, J., Greenspan, H., (2008) TAU MIPLAB at ImageClef 2008. In: Working Notesof CLEF 2008, Aarhus, Denmark

Bach, J., Fuller, C., Gupta, A., Hampapur, B., Horowitz, B., Humphrey, R., Jain, R., and Shu, C., (1996). Virage Images Search Engine: An Open Framework for Image Management. *In SPIE Conference on Storage and Retrieval Management for Image Database IV,* vol 2670, pp76-87.

Bo, Q., Wei, X., Qi, T., Chang, S., (2005) Report for annotation task in ImageCLEFmed 2005. In: Working Notes of CLEF 2005, Vienna, Austria

Cohen, S., (1999). Finding Colours and Shape Pattern in Images. PhD, Stanford University.

Cheng, H., Zhicong, H., & Mark, K., (2006). Final Project Proposal-Image Processing Techniques. UCLA Electrical Engineering, EE 113D, Spring.

Dana, H., Ballard, B., and Christopher, M., (1982). *Computer Vision*, chapter 4, pages 119–149. Prentice Hall.

Das, M., Riseman, E., and Draper, B., (1997). FOCUS: Searching for Multi-Coloured Object in a Diverse Image Database. *In; Processing of IEEE Conference on Computer Vision and Pattern Recognition.* Pp756-761.

Deselaers, T., Weyand, T., Keysers, D., Macherey, W., Ney, H., (2005) FIRE in Image CLEF: Combining content–based image retrieval with textual information retrieval. In: CLEF 2005 Proceedings, Springer, Lecture Notes in Computer Science (LNCS), vol 4022, pp 652–661

Deselaers, T.,Weyand, T., and Ney, H., (2006) Image retrieval and annotation using maximum entropy. In: CLEF 2006 Proceedings, Springer, Lecture Notes in Computer Science (LNCS), vol 4730, pp 725–734

Feder, J., (1996). Image Recognition and Content-Based Image Retrieval for the World Wide Web. *Advance Imaging* 11(1):26-28.

Flicker, M.., Sawhney, H., Nibleck, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D. & Yanker, P. (1995). Query by Image and Video Content: The QBIC system. *IEEE Computer.* 28(9), 23-32.

Fridtjof, S., and Gerard, M., (1990). Efficient Two Dimensional Object Recognition. *In Proceedings of 10th International Conference on Pattern Recognition*, pages 13–17, Atlantic City, New Jersey.

Gianni, L., Vernazza, S., Serpico, B., and Silvana G., (1987). A Knowledge-Based System for Biomedical Image Processing and Recognition. *IEEE Transactions on Circuits and Systems*, Cas-34(11):1399–1416.

G¨uld, M., Christian, B., Lehmann, T., (2005) Combining global features for content–based retrieval of medical images. In: Working Notes of CLEF 2005, Vienna, Austria

Gupta, A., and Jain, R., (1997). Visual Information Retrieval. *Comm. Assoc. Comp. Mach.,* vol. 40, no. 5, pp. 70-79.

Hou, T., Hsu, A., Liu, P., and Chiu, M., (1992). A Content-based Indexing Technique Using Relative Geometric Feature. *SPIE Conference on Storage and Retrieval Management for Image Database IV,* vol 1662, pp59-68.

Idris, H., and Panchanathan, R., (1997). Review of Image and Video Indexing Techniques. *Journal of Visual Communication and Image Representation.* 8(2):146-166.

Ioannis, K., (1990). Segmentation and Feature Extraction for Magnetic Resonance Brain Image Analysis. In *Proceedings of 10th International Conference on Pattern Recognition*, pages 583–590, Atlantic City, New Jersey, June 1990.

Lim, J., Chevallet, J., Thi, D., Le, H., and Goh, H., (2007). Bi-modal Conceptual Indexing forMedical Image Retrieval. Satoh, S., Nack, F., and Etoh, M., (Eds.): MMM 2008, LNCS 4903, pp. 456–465,c_Springer-Verlag Berlin Heidelberg.

John, S., Vittorio, C., and Chung-Sheng L., (1999). Adaptive Storage and Retrieval of Large Compressed Images. IS&T/SPIE Symposium on Electronic Imaging: Science and Technology-Storage and Retrieval of Images Database VII. Pp 2.

Ma, W., and Manjunath, B., (1997). NaTra: A Toolbox for Navigating Large Image Databases. *Proc. IEEE Int. Conf.   Image Processing,* pp. 568-71.

Minka, T., and Picard, R., (1997). Interactive Learning with a Society of Models. Pattern Recognition 30(4):565-581.

Muller, H., Michoux, N., Bandon, D. & Geissbuhler, A. (2004). A Review of Content-based Image Retrieval Systems in Medical Application-clinical Benefits and Future Directions. *International Journal of Medical Informatics.* 73(1), 1-23.

Niblack, W., Barber, R., Equitz, W., Flickner, M., Glassman, E., Petkovic, D., Yanker, P., Faloutsos, C., Taubin, G., and Height, W., (1993). Querying Image by Content using Color, Texture and Shape. *SPIE   Conference on Storage and Retrieval Management for Image Database IV,* vol 1908, pp173-187.

Lehmann, T., Guld, M., Thies, C., Fischer, B., Spitzer, K., Keysers, D., Ney, H., Kohnen, M., Schubert, H., & Wein, B. (2003). The IRMA Project: A State of the Art Report on Content-based Image Retrieval in Medical Applications. *Proceedings of Korea-Germany Joint Workshop on Advanced Medical Image Processing.* 161-171.

Liu, J., Hu, Y., Li, M., Ma, S., Ying, M., (2006). Medical image annotation and retrieval using visual features. In: CLEF 2006 Proceedings, Springer, Lecture Notes in Computer Science (LNCS), vol 4730, pp 678–685

Olugbara, O. (2008). Content-based image retrieval in Healthcare information infrastructure for Classification and diagnosis. PhD Thesis, University of Zululand, South Africa.

OVE, and Chabot, S., (1995). Retrieval from a Relational Database of Images. *Computer.* Pp 40-48.

Petros, K., and Stelios, C., (1991). Graphical Tools and Retrieval Strategies for Image Indexing by Content. In *Proceedings of Computer Assisted Radiology, CAR91,* pages 519–524, Berlin, July 1991.

Pentland, A., Picard, R., and Sclaroff (1996). Photobook: Content-based Manipulation Image Databases. *International Journal of Computer Vision,* 18(3):233-254.

Rubner, Y., (1999). Perceptual Metrics for Image Database Navigation, PhD Dissertation, Computer Science Department, Stanford University.

Setia, L., Teynor, A., Halawani, A., Burkhardt, H., (2008) Grayscale medical image   annotation using local relational features. Pattern Recognition Letters 29(15):2039–2045

Smeulders, A., 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.

Smith, J., (1997). Integrated Spatial and Feature Image Systems: Retrieval, Analysis and Compression. PhD Thesis, University of Columbia.

Shi-Kuo, C., (1989). *Principles of Pictorial Information Systems Design,* chapter 8, pages 172–211. Prentice Hall International Editions, Englewood Cliffs, New Jersey 07632.

Tommasi, T., Orabona, F., Caputo, B., (2008a). CLEF2008 Image Annotation Task: an SVM Confidence–Based Approach. In: Working Notes of CLEF 2008, Aarhus, Denmark

Tsai, W., and Shiaw-Shian., Y., (1985). Attributed String Matching with Merging for Shape Recognition. *IEEE Trans. on Patt. Anal. and Mach. Intel.,* 7(4):453–462.

Efficient Decision Support Systems - Practice and Challenges in Biomedical Related Domain

Edited by Prof. Chiang Jao

ISBN 978-953-307-258-6

Hard cover, 328 pages

Publisher InTeh

Published online 06, September, 2011

Published in print edition September, 2011

This series is directed to diverse managerial professionals who are leading the transformation of individual domains by using expert information and domain knowledge to drive decision support systems (DSSs). The series offers a broad range of subjects addressed in specific areas such as health care, business management, banking, agriculture, environmental improvement, natural resource and spatial management, aviation administration, and hybrid applications of information technology aimed to interdisciplinary issues. This book series is composed of three volumes: Volume 1 consists of general concepts and methodology of DSSs; Volume 2 consists of applications of DSSs in the biomedical domain; Volume 3 consists of hybrid applications of DSSs in multidisciplinary domains. The book is shaped decision support strategies in the new infrastructure that assists the readers in full use of the creative technology to manipulate input data and to transform information into useful decisions for decision makers.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Achimugu Philip, Babajide Afolabi, Adeniran Oluwaranti and Oluwagbemi Oluwatolani (2011). Development of an Image Retrieval Model for Biomedical Image Databases, Efficient Decision Support Systems - Practice and Challenges in Biomedical Related Domain, Prof. Chiang Jao (Ed.), ISBN: 978-953-307-258-6, InTech, Available from: http://www.intechopen.com/books/efficient-decision-support-systems-practice-and-challenges-in-biomedical-related-domain/development-of-an-image-retrieval-model-for-biomedical-image-databases

INTECH
open science | open minds