CHAPTER 2

LITERATURE SURVEY

2.1 INTRODUCTION

Multimedia databases have emerged as the natural solution to achieve the vast growth of multimedia information. Multimedia databases and the internet not only boost each other, but also have a dramatic impact on developers, network administrators, content providers and users reported by (Johnson 1998). By its very nature, multimedia databases need a large storage area than other conventional databases since they store mainly images, audio and video. An image database contains a large collection of images with similar features that makes the querying mechanism problematic. And currently, many of the research works on image databases rely on low level features, either text or image features, that lead to a number of limitations while trying to get an exact query result. The task of image retrieval is to find and retrieve the most similar figure for a given query. However, image retrieval is a complex process that inherits techniques from different fields like pattern matching, information retrieval and computer graphics.

Typically, an image in multimedia database is searched based on keywords, features and/or concepts. But the problem, specifically in a feature based retrieval, lies on the vast number of attributes to express the image, size, color, shape, texture, location, position, domain, etc. Moreover, the complexity in the nature of two dimensional image data
Given a tremendous amount of image data, the capabilities to support efficient and effective image retrieval have become increasingly important. Generally, there are two general approaches for image retrieval:

1. The text-based approaches apply traditional text retrieval techniques to image annotations or descriptions whereas

2. The content based approaches apply image processing techniques to extract image features and retrieve relevant images.

One obvious approach is to describe the image contents verbally, typically using keywords. Once the verbal descriptions are obtained, text search techniques can be applied to retrieve images in the database allowing query by keyword, but this assumption is seldom met since manual labeling is too expensive. The other approach is to represent images with nonverbal descriptions which can be reliably computed from images. Such descriptions are image features based on color, shape, and texture. Conventional content based image retrieval (CBIR) methods use these image features to define image similarity Shingo et al. (2005).

Generally, in both text based and content based image retrieval systems, the bottleneck to the efficiency of the retrieval is the semantic gap between the high level image interpretations of the users and the low level image features stored in the database for indexing and querying. In other words, there is a difference between what image features can distinguish and what people perceives from the image since human perception is complex and seems to be dependent on context, purpose, emotion, psychological ground and many more individual cases.
It was back in 1970 that the text based image retrieval has got its attention. Early on, keyword based search has become the leading paradigm for querying multimedia databases; while it has lots of limitations. In text based image retrieval systems, the first thing to do is providing textual descriptions/annotations for images, but it is very tiresome, inefficient and expensive and has a problem of undesirable mismatch due to annotation impreciseness. Text annotation is extremely tedious in large image collections. To provide text descriptions or annotations, two approaches can be applied. The first approach acquires descriptions/annotations manually by human annotators. The second approach is to automatically annotate images using machine learning techniques that learn the correlation between image features and textual words from the examples of annotated images proposed by Chen Zhang et al. (2005). In both cases, manual labeling is too expensive while automatic methods are not reliable. Automatic image classification yields limited access, in such a way that only few objects like faces or cars can be recognized reliably from general images reported by (Shingo et al. 2005 and Daniel et al. 1995).

Another problem in the use of keywords is the complexity between words and concepts due to synonymy (different words denote the same concept) or homonymy (same word denotes different concepts) and many search criteria cannot be well described by a few keywords reported by Marin et al. (2008). Current researches proposed by (Chen Zhang et al. 2005, Marin et al. 2008 and Huan et al. 2008) conclude that keyword based searching cannot be successful alone. In finding a target image, the content of the image and the image features play an important role. It is also suggested by Huan et al. (2008) that since images might have no annotations or been incompletely annotated, a joint use of existing text annotations and visual features can provide better retrieval results.
In recent years, indexing and retrieval approaches based on keywords and visual features together are getting attentions. A work by Zhao & Grosky (2008), which can be a good example of such approach, solved to some extent the problem of synonymy and homonymy using Latent Semantic Indexing (LSI). LSI can also be applied to the joint visual and keyword based feature vectors in order to find a hybrid reduced representation that links sets of keywords and images. Unfortunately, to identify meaningful relations between keywords, LSI needs large amounts of data. This requirement can only be met when a relatively large quantity of text rather than just a few keywords is associated to every image. They are the important difficulties that boosted research activities in the field of CBIR in the early 1980s. CBIR came into being to solve those problems inherited by text based systems, and currently many of the researches in image retrieval concentrate on the advancement of CBIR systems in many directions from simple low level features to a combined visual and human interactive system for better achievement.

2.2 CONTENT BASED IMAGE RETRIEVAL

Many of today’s image retrieval systems rely on CBIR with varied techniques ranging from single feature vector to combined visual and conceptual image content descriptions and ontology. In CBIR, images are indexed by their visual content, such as color, texture, and shapes. Moreover, the fundamental difference between content-based and text-based retrieval systems is that the human interaction is an indispensable part of the latter system. Humans tend to use high level features i.e. concepts, such as keywords, text descriptors, etc. to interpret images and measure their similarity. Even though the field of CBIR has been extensively researched in recent years, none of the proposed approaches has achieved satisfactory performance due to the semantic gap; expressing
the discrepancy between the low level features that can be readily extracted from the images and the high level descriptions that are meaningful to the users presented by Abolfazl et al. (2008).

Image content can be described at various levels. It may favor perceptual features like color, texture, shape, structure and spatial relationship, or semantic primitives such as the identification of real world objects and the meaning of the images reported by Renato et al. (2001), and image retrieval using low level visual features is a challenging and important issue in content based image retrieval. However, most of the CBIR systems focused on low level features.

Multimedia contents, in general, can be represented as keyword based, feature based and/or concept based. Most of the CBIR systems are based on the basic low level features (color, texture, shape and spatial relationship). As most researchers agree, due to the difficulty of inferring semantic meaning from low level features, none of CBIR systems are satisfactory. Including yahoo and Google, many well-known search engines are currently limited to textual keywords. Due to the increasing demand of better management and retrieval of multimedia data, the next generation multimedia databases are looking for better performance from CBIR systems. Despite the difficulty of extracting exact images features, the vital step in CBIR system is feature extraction. From the vast amount of research works based on low level features, color is the dominant one due to its robustness and its independence to the size and orientation of images.

It is also said by Remco et al. (2000) that most systems use color and texture features, few systems use shape feature, and still less use layout features. The retrieval on color usually yields images with similar colors. Retrieval on texture does not always yield images that have clearly the same texture, unless the database contains many images with a dominant
texture. Searching on shape gives often surprising results. Apparently, the shape features used for matching are not the most effective ones. Color is a visual feature which is immediately perceived when looking at an image. It is one of mostly used visual features in retrieval and can also be used to find the location in an image and to differentiate a large number of objects proposed by Sridhar et al. (2006).

Image features have been divided into three levels reported by (Eakins 1998) as followings

1) Level 1 - Primitive features such as color, texture, shape or the spatial location of image elements. Typical query example is 'find pictures like this

2) Level 2 - Derived attributes or logical features, involving some degree of inference about the identity of the objects depicted in the image. Typical query example is 'find a picture of a flower

3) Level 3 - Abstract attributes, involving complex reasoning about the significance of the objects or scenes depicted. Typical query example is 'find pictures of a beautiful lady.

The majority of content based image retrieval systems mostly offer level 1 retrieval, a few experimental systems level 2, but none level 3.

Commonly used image features for content based image retrieval were followings
1) Color

Color is one of the visual cues often used for content description, but most medical images are grayscale. Thus, color features are not used for medical image retrieval.

2) Texture

Texture features mean spatial organization of pixel values of an image and used in standard transform domain analysis by tools such as Fourier transform, wavelets, Gabor or Stockwell filters. In the medical images, texture features are useful because they can reflect the details within an image structure.

3) Shape

Shape feature has broad range of visual cues such as contour, curve, surfaces, and so on. Recently, many methods measures similarity between images using shape features has been developed.

The first step to extract color features is to select an appropriate color space. Several color spaces are available, such as RGB, CMYK, and HSV. Most digital images are stored in RGB color space. However, RGB color space is not perceptually uniform, which implies that two colors with larger distance can be perceptually more similar than another two colors with smaller distance, or simply put, the color distance in RGB space does not represent perceptual color distance.

Texture is a distinctive property of virtually all surfaces. It represents the regularity, smoothness and coarseness of the image. Texture gives a direction sense to the spatial arrangement of image intensities reported by (Sridhar et al. 2006). Texture is normally defined purely by
grey levels and as such is orthogonal to color. Texture has qualities such as periodicity and scale; it can be described in terms of direction, coarseness, contrast and so on. It is this that makes texture a particularly interesting facet of images and results in a plethora of ways of extracting texture features.

Color histogram, Color Coherence Vector (CCV) and color moments are the common approaches for color feature representation of an image whereas filter banks and AM-FM models have been used as a common model to represent texture of an image. Those techniques, however, are all global methods in that they extract the visual metadata of the whole image. Global methods are advantageous because they have a compact representation and the extracted visual metadata can be, under certain constraints, efficiently compared and indexed using a Spatial Access Method. The main disadvantage is that in these approaches, there is no information about the spatial distribution of colors inside the images. Thus, images with very different spatial layout may have similar representations (Renato et al. 2001). In order to take spatial distribution into account, several regional methods have been developed and proposed by (Renato et al. 2001, Martin et al. 2004 and Xiangli Xu et al. 2008). The Color Based Clustering (CBC) system is the one encouraging regional systems. In regional methods, an image is segmented into a set of regions according to a predefined visual property and each region is represented individually.

A further recent work of low level feature based CBIR includes users. Incorporating human intuition and emotion into retrieving images is vital for effective result reported by (Sung-Bae Cho & Joo-Young Lee 2002, Ghazanfar & Hasnain 2005). Such kind of human oriented systems facilitate the search in both explicit and implicit queries. Such method,
reported by Ghazanfar & Hasnain (2005) uses Interactive Genetic Algorithm (IGA) together with wavelet transformation. The Interactive Genetic Algorithm adopts the users’ choice as fitness and a user can increase/decrease the effectiveness of a color indefinitely and interactively until he/she gets his/her requirement.

Using a combination of visual features and concept oriented interactive manner yields a more effective search result in image databases in comparison to the traditional CBIR system using single visual feature and simple linear combined low level visual features. To address the challenge of semantic gap reduction for image retrieval, many of recent CBIR systems tried to make full use of image information to extract features.

Many methods have been used in CBIR system, including methods based on color reported by (Martin et al. 2004 and XiangliXu 2008), based on texture reported by Vermaa & Kulkarni (2004), based on shape reported by (George Gagaudakis & Paul Rosin 2002), based on spatial relations reported by Mario et al. (2003) and so on. Most CBIR systems allow query formulation with user setting of relative importance of features to mimic the user's perception of similarity. Similarity matching is crucial in retrieval systems to understand new concepts with existing ones. Similarity matching function is multi-level, until the most primitive distance measures or correlation metrics are used at the leaf level. It is obvious that, as the combined features increase, the retrieval effectiveness also improved; let al. one the time and space complexity.

A recent work on CBIR proposed by XiangliXu et al. (2008) incorporates granularity theory to the combined low level visual features. They propose color information of image (based on histogram) and texture feature (based on gray level co-occurrence matrix/GLCM) as both methods
are considered to be quick and effective classical algorithms. The work contributes much in the application of granularity theory in image retrieval, but it lacks on comparing it to other combined feature approaches. A better result is obviously obtained in comparison to methods adopting only single retrieval algorithms.

To make retrieval more intelligent and interactive, a CBIR system should also incorporate other related complementary disciplines such as the knowledge based system, computer vision, data mining, neural networks and so on. A more recent approach proposed by XiangliXu et al. (2008) focuses on intelligent concept oriented search. As they said, traditional image retrieval based on visual based matching is not effective in multimedia applications. Consequently, the modeling of high-level human sense for image retrieval has been a challenging issue over the past few years. Most of the contemporary CBIR systems explored to look for the primary features to present an image. However, the related efforts are not satisfactory enough due to the gap between low-level features and high-level concepts, and hence problems still exist in traditional classification-based image retrieval.

2.3 SEMANTIC BASED IMAGE RETRIEVAL

The ideal Content Based Image Retrieval System from a user perspective would involve what is referred to as Semantic Retrieval, where the user makes a request like “find pictures of Manmohan Singh”. This type of open ended task is very difficult for computers; because pictures could look different, and Manmohan singh may not always be facing the camera or in the same pose. Current content based image retrieval systems, therefore, generally make use of lower level features like texture, color and shape, although some systems take advantage of very common higher level features.
The fundamental difference between the content based and the text based retrieval systems is that the human interaction is an indispensable part of the latter system. Humans tend to use high-level features (concepts), such as keywords, text descriptors, etc., to interpret images and measure their similarity. While the features automatically extracted using computer vision techniques are mostly low-level features (color, texture, shape, spatial layout, etc.). In general, there is no direct link between the high-level concepts and the low-level features. In Semantic Based Image Retrieval, retrieval from the database is based on a measure of similarity between the sought and the existing visual object process diagram. This measure considers the distances between object-process diagrams while taking into account the fact that they represent visual data. Though many sophisticated algorithms have been designed to describe color, shape, and texture features, these algorithms cannot adequately model image semantics and have many limitations while dealing with broad content image databases.

Semantic based image retrieval is among the recent works of image retrieval systems that focus mainly on the techniques which can reduce the semantic gap of CBIR systems reported by (Abolfazl et al. 2008). An improved Support Vector Machine (SVM) based active relevance feedback frame work together with a hybrid visual and conceptual content representation and retrieval is one approach proposed by Abolfazl et al. (2008). This approach employs the global color, texture and shape where the shape content is described by Hough transform. Computing a conceptual feature vector for every image is the major overhead of the method. From the experimental evaluation reported by Abolfazl et al. (2008), employing both visual and concept based feature vectors visibly improves the quality of the results compared to using visual
features alone, and projecting the keywords on all the key concepts gives better performance than projecting only on their key super concepts.

Both semantic keyword query and query by example are supported in Semantic Based Image Retrieval. According to the submitted semantic keywords (corresponding to semantic categories), the system will first find out those images that contain all of the categories, then rank the documents by sorting the score stored for each image in the database. For those images belonging to the UNKNOWN category, its score will be multiplied by a diminishing factor. When the user submits the query-by-example, if the query image is in the database, its Content will be used as query keywords to perform the retrieval; otherwise it will first go through the above steps to obtain the Content and the Score. When user is interested in retrieving images with, not only same semantics, but also the similar visual features as that of the query. Euclidean distance of saliency value will be calculated between the query image Q and the image in the category.

Semantic retrieval techniques depending on user interaction cover a broad spectrum of intelligence. A relatively low level technique for bridging the semantic gap is content based navigation, the use of generic links specified in terms of text or image features. This can be used to construct a multimedia thesaurus reported by (Lewis et al. 1997) specifying semantic relationships between source items in the link database, whether text, image or sound. This allows system users to build up a database of semantic relationships between text terms and their corresponding images. However, human intelligence is still required to establish linkages between image types and their semantic meanings. The system itself acts purely as a repository for this knowledge; it provides no mechanism for automated
reasoning or learning. More obviously & intelligent' is a family of techniques based on extensions of the relevance feedback principle.

One of the earliest systems to provide this kind of interaction was Four Eyes reported by (Minka 1996), which allowed a user to group arbitrary regions of images (such as particular types of building, or species of plant), and optionally give these regions semantic labels such as grass or sky. Once the user has assigned labels to several examples of the same type from one or more images, the system attempts to induce grouping rules from the positive and negative examples at its disposal.

Image retrieval at the semantic level can be achieved only by reference to some knowledge base of prior experience. In every case, it is possible to identify either a repository of past knowledge about the domain in which they operate, or a mechanism for building up such a store. Such a feature is seldom, if ever, seen in primitive-level systems.

The higher the level of interpretation required to answer a semantic query, the larger and more complex the knowledge base and reasoning mechanisms needed to perform this interpretation.

Successful semantic retrieval involving images of complex objects or scenes requires an adaptive system capable of learning from experience. Models of almost any type of object need to be built up and refined over a period of time in the light of experience. To expect system designers to get every detail of their models right first time is unreasonable. Even if they were to do, another problem would always remain. The visual properties of almost any human artefact are likely to change, perhaps, quite frequently over time. It is simply not feasible for system designers to modify an object model every time a visually different example of that object is encountered. The system itself has to adapt.
If the conjectures outlined above do turn out to reflect the processes underlying semantic image retrieval, one can expect a significant proportion of further advances in this field. As indicated above, a number of specific techniques have already been applied to image retrieval at the semantic level, including rule based reasoning, neural networks, and genetic algorithms.

VisEngine is one of the best Semantic Based Image Retrieval System. It has some characteristics:

1) Main Region Segmentation

Human used to judge an image by its main regions because most semantic information focuses on them. Thus in VisEngine, before retrieval, each image in database is first pre-processed, especially main regions of it are segmented, and stored in computer with middle representation so that they can be favor of further procession.

2) Middle Representation

Middle representation is the description of image in the way that computer can understand. So, proper middle representation can shorten the difference between human and computer, on the understanding of the image. In VisEngine, an object-oriented middle representation is proposed.

3) Semantic-based Interaction and feedback

During interaction, users choose similar images among results according their own understanding as feedback image. In VisEngine, a more natural semantic-based
interaction is provided so that users can express their ideas better and system can learn more from users’ feedback. Hence retrieval effect is improved

2.4 ONTOLOGY BASED IMAGE RETRIEVAL

Ontology is a specification of conceptualization. It consists of concept hierarchy, concept properties and relations between concepts in a topic area. It helps to extract semantic meanings from images, and facilitate retrieval in a convenient way, thus bridging the semantic gap. Users could conveniently formulate a query in various aspects to get the artifacts they need, and the feedback from some human subjects that have used the system is satisfactory. There are two basic types of ontology: upper ontology and domain ontology. The domain dependent ontology defines the fine-grained concepts and allows determining specific relationships between the concepts in a given area.

Ontology is an important discipline that has the huge potential to improve information organization, management and understanding. Accordingly, ontology is the term referring to the shared understanding of some domains of interest which is often conceived as a set of classes (concepts), relations, functions, axioms and instances. Ontology is playing more and more important role in textual analysis, and information exchange between different domains. In fact, image retrieval, taking advantage of concept hierarchy or ontology, is recently proposed, and the semantic concept of images may come from automatic computer vision methods, and manual annotation. This method could establish implicit or explicit relations of different concepts, and make users more naturally obtain images they want.
The ontologies form the core of the system and are used for three purposes:

1) Annotation terminology

The ontological model provides the terminology and concepts by which metadata of the images is expressed.

2) View based search

The ontologies of the model, such as Events, Persons, and Places provide different views into the promotion concepts. They can hence be used by the user to focus the information need and to formulate the queries.

3) Semantic browsing

After finding a focus of interest, an image, the semantic ontology model together with image instance data can be used in finding out relations between the selected image and other images in the repository. Such images are not necessarily included in the answer set of the query. For example, images where the same person occurs but in a different event of the same promotion may be of interest and be recommended to the user, even if such images do not match the query.

Ontology based image retrievals is beyond the two major image retrieval paradigms, text based and content based. In practical applications both have limitations. Ontology based image retrieval has the potential to fully describe the semantic content of an image, allowing the similarity between images and retrieval query to be computed accurately reported by (Huan et al. 2008 and Hyvönen et al. 2003). Semantic technologies like
ontology and the XML markup language provide tools for a promising new approach to image retrieval based on implementing semantic understanding of image content. Ontology based image retrieval has two components namely the semantic image annotation which focuses mainly on the description of image content, and semantic image retrieval, to allow searching and retrieval based on image content. Besides semantic annotation of an image or retrieval query still needs the intervention of a human being. It is an open research area to make it automated and interactive similar to some of the combined visual feature based CBIR systems do.

Another advanced method proposed by Huan et al. (2008) is the Ontology with multi-modality based image retrieval system. Researchers argue that such ontological searches for images are very efficient especially for images with complicated content and ambiguous semantics. Single modality refers to either the text or image features of a domain specific image, whereas multimodality refers both. The goal is to create machine processable queries for such kind of diverse domain specific images, in their example, using both text and image features. The system uses the widely known description logic for representing knowledge in terms of classes, and relationships between classes use a match making algorithm to bind user specified queries with the knowledge base. The system is compared with Google (key word based approach) and other ontology based approaches. Its important privilege lies on its flexibility and the ability to add more domains easily without changing the entire architecture. One major bottle neck in such a system is the extra burden in creation of the ontology, construction of the knowledge base and the classification of the accuracy of the image feature. The tradeoff between annotation work and quality of information retrieval can be balanced by using less detailed ontologies and annotations, if needed.
2.5 MEDICAL IMAGE RETRIEVAL

CBIR plays a key role in medical image retrieval fields such as CT scan images, MRI scan images, X-rays etc. Many systems already have been developed in this domain but each one has some difficulties in some way to provide efficient retrieval. So, efficient content based image retrieval in medical domain is still a challenging problem. CBIR in the medical field also presents a growing trend in publications reported by (Long et al. 2003). Although the number of experimental algorithms comprehending specific problems and databases face a growth its reflection on the number of medical applications and frameworks is still very constrained. Only a few systems exist with relative success. The Cervigram Finder system proposed by Xue et al. (2008) was developed to study the uterine cervix cancer. It is a computer assisted framework where local features from a user-defined region in an image are computed and, using similarity measures, similar images are retrieved from a database.

The works related to Medical Image searching techniques in Content Based Image retrieval include Image Retrieval for Medical applications reported by Lehmann et al. (2004), The Spine Pathology and Image Retrieval System (SPIRS) reported by Antani et al. (2004), The Image Map reported by (Petrakis et al. 2002), The Automatic Search and Selection Engine with Retrieval Tools (ASSERT) reported by Chi-Ren Shyu et al. (1999) and The MIMS reported by Chbeir et al. (2000).

In the medical field, images especially digital images are produced in ever increasing quantities and used for diagnostics and therapy. With Digital Imaging and Communications in Medicine (DICOM) a standard for image communication has been set, and patient information can be stored with the actual image(s), although a few problems still exist with respect to standardization.
Imaging systems and image archives have often been described as an important economic and clinical factor in the hospital environment reported by (Greene, Brinkley 2000 and Kulikowski et al. 2002). Several methods from the computer vision and image processing fields have already been proposed for use in medicine reported by (Sarvazyan 1991). Several radiological teaching files exist reported by (Rosset et al. 2002 and Binet et al. 1995) and radiology reports have also been proposed in a multimedia form in Maloney & Hamlet (1999). Web interfaces to medical databases are described in Frankewitsch & Prokosch (2001).

Medical images have often been used for retrieval systems and the medical domain is often cited as one of the principal application domains for content based access technologies in terms of potential impact. While there is widespread enthusiasm for CBIR in the engineering research community, the incorporation of this technology to solve practical medical problems is a goal yet to be realized. Possible obstacles to the use of CBIR in medicine include:

1) The lack of productive collaborations between medical and engineering experts, an important point related to usability and performance characteristics of CBIR systems,

2) The lack of effective representation of medical content by low-level mathematical features,

3) The lack of thorough evaluation of CBIR system performance and its benefit in health care,

4) The absence of appropriate tools for medical experts to experiment with a CBIR application, another important point again related to usability and performance attributes of CBIR systems.
The early years of medical CBIR have been reviewed by Müller et al. (2003). Content based retrieval has also been proposed several times from the medical community for the inclusion into various applications, Lehmann et al. (2003) often without any implementation. Still, for a real medical application of content based retrieval methods and the integration of these tools into medical practice a very close cooperation between the two fields is necessary for a longer period of time. A simple exchange of data or a list of the necessary functionality will not suffice.

There are several reasons why there is a need for additional, alternative image retrieval methods apart from the steadily growing rate of image production. It is important to explain these needs to discuss possible technical and methodological improvements and the resulting clinical benefits. The goals of medical information systems have often been defined to deliver the needed information at the right time, the right place to the right person in order to improve the quality and efficiency of care process reported by Thies et al. (2005). Such a goal will most likely need more than a query by patient name, or study ID for images. For the clinical decision making process it can be beneficial or even important to find other images of the same modality and the same anatomic region of the same disease. Although part of this information is normally contained in the DICOM headers and many images devices are DICOM compliant at this time, there are still some problems. DICOM headers have proven to contain a fairly high rate of errors, for example error rates of 16% have been reported in Quality of DICOM header information for image categorization. This can hinder the correct retrieval of all wanted images.

Clinical decision support techniques such as case based reasoning or evidence based medicine reported by Le Bozec et al. (1998) can even produce a stronger need to retrieve images that can be valuable
for supporting certain diagnosis. It could even be imagined to have Image Based Reasoning as a new diagnostic aid. Decision support systems in radiology reported by Bui et al. (2002) and computer aided diagnosis for radiological practice as demonstrated at the Radiological Society of North America reported by Boissel et al. (2003) are on the rise and create a need for powerful data and meta data management and retrieval. The general clinical benefit of imaging system has also been already demonstrated by Kaplan et al. (1996). In Horsch et al. (2006) an initiative is described to identify important tasks for medical imaging based on their possible clinical benefits.

It needs to be stated that the purely visual image queries as they are executed in the computer vision domain will most likely not be able to ever replace text based methods as there will always be queries for all images of a certain patient, but they have the potential to be a very good complement to the text based search based on their characteristics. Still, the problems and advantages of the technology have to be stressed to obtain acceptance and use of visual and text based access methods up to their full potential.

Besides diagnostics, teaching and research are expected to improve the use of visual access methods as visually interesting images can be chosen and can actually be found in the existing large repositories. The inclusion of visual features into medical studies is another interesting point for several medical research domains. Visual features do not allow the retrieval of cases with patients having not only similar diagnoses but also visual similarity with different diagnoses. In teaching it can help lecturers as well as students to browse educational image repositories and to visually inspect the results found. It can be used to cross-correlate visual and textual features of the image.
There are a large number of propositions for the use of content-based image retrieval methods in the medical domain in general reported by (Tagare et al. 1997, Lowe 1998 and Bidgood 1999). Other articles describe the use of image retrieval with an image management framework reported by (Chu et al. 1994, Orphanoudekis et al. 1996, Petrakis 2002 and Tsiknakis et al. 1996), sometimes with outstating what has actually been implemented and what is still in the status of ideas. Also the integration into PACS systems reported by (Le Bozec et al. 2000) or other medical image databases reported by (Bucci et al. 1996) has been proposed often, but the implementation details are generally rare.

Most of the general articles such as (Tagare et al. 1997) state that the medical domain is very specialized such that general systems cannot be used. This is true but it is the case for all specialized domains such as trademark retrieval or face recognition, and specialized solutions need to be found. The more specialized the features of a system are the smaller the range of applications and the compromise for each specific application are a need to be found. Domain knowledge needs to be integrated into specialized query engines.

Another proposition of what is needed for an efficient use in the medical domain is given in Lowe (1998), including some implementation details. Clinically relevant indexing and selective retrieval of biomedical images are explained in Bidgood (1999). Some examples are given but no implementation details.

Most of these articles ask for semantic retrieval based on images that are segmented automatically into objects and where diagnoses can be derived easily from the objects’ visual features. This is still a dream, as it has been in the computer vision domain for general segmentation methods for a while. Steps into the direction of solutions have to be taken using
machine learning technique sand by including specific domain knowledge. Implementation of image retrieval systems is a step-by-step process and first systems will definitely not meet all the high requirements that are asked for.

Image retrieval based on visual features is often proposed but unfortunately nothing is said about the visual features used or the performance obtained. Tsiknakis et al. (1996) describes a telemedicine and image management framework and Chu et al. (1994) is another very early article on the architecture of a distributed multimedia database. Chang (1996) describes an active index for medical image data management and in Petrakis (2002) a newer image management environment is described. In the works by Lowe et al. (1995), Wong & Huang (1997), two frameworks for image management and retrieval are described focusing on technical aspects and stating application areas. One of the few frameworks with at least a partial implementation is the Image Retrieval in Medical Applications (IRMA) framework Keysers et al. (2003) that allows for a relatively robust classification of incoming image into anatomical regions, modality and the taken orientation. This project also developed a classification code for medical images based on four axes (modality, body orientations, body region, and biological system) to uniquely classify medical images and allow testing and measuring the performance of classification Lehmann et al. (2003).

2.6 MEDICAL APPLICATIONS OF CONTENT BASED IMAGE RETRIEVAL

There are various techniques which are used or have been proposed for the use in medical image retrieval applications. Many of the techniques are similar to those used for general content based retrieval but some techniques that have not yet been used in medical applications also
identified. A special focus is put on the data sets that are used to evaluate the image retrieval systems and on the measurements used for evaluation. Unfortunately, the performance evaluation of systems is currently neglected.

Content-based image retrieval has frequently been proposed for various applications. This section will discuss three potential applications of medical image retrieval. Content-based image retrieval has been proposed by the medical community for inclusion into Picture Archiving and Communication Systems (PACS) reported by Lehmann et al. (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 reported by 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 reported by (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.
2.7 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 presented by (Hsu & Ho 2004 and Schmidt et al. 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 reported by Chang et al. (2004).

Computer Aided Diagnosis (CAD) has become one of the major research subjects in medical imaging and diagnostic radiology. The basic concept of Computer Aided diagnosis is to provide a computer output as a second opinion to assist radiologist’s image interpretation by improving the accuracy and consistency of radiological diagnosis and also by reducing the image reading time. In this article, a number of Computer Aided Diagnosis schemes are presented, with emphasis on potential clinical applications. These schemes include

1. Detection and classification of lung nodules on digital chest radiographs
2. Detection of nodules in low dose CT
3. Distinction between benign and malignant nodules on high resolution CT
4. Usefulness of similar images for distinction between benign and malignant lesions
(5) Quantitative analysis of diffuse lung diseases on high resolution CT, and
(6) Detection of intracranial aneurysms in magnetic resonance angiography. Because Computer Aided diagnosis can be applied to all imaging modalities, all body parts and all kinds of examinations, it is likely that Computer Aided Diagnosis will have a major impact on medical imaging and diagnostic radiology in the 21st century.

A major difference between the CAD and the computer diagnosis is the way in which the computer output is utilized for the diagnosis. With CAD, radiologists use the computer output as a second opinion, and radiologists make the final decisions. Therefore, for some clinical cases in which radiologists are confident about their judgments, radiologists may agree with the computer output, or disagree and then disregard the computer. However, for cases in which radiologists are less confident, it is expected that the final decision can be improved by use of the computer output. This improvement is possible, of course, only when the computer result is correct. The higher the performance of the computer, better the overall effect on the final diagnosis.

However, the performance level of the computer does not have to be equal to or higher than that of radiologists. With CAD, the potential gain is due to the synergistic effect obtained by combining the radiologist’s competence and the computer’s capability. Because of these multiplicative benefits, the current CAD has become widely used in practical clinical situations.

With automated computer diagnosis, however, the performance level of the computer output is required to be very high. For example, if the
sensitivity for detection of lesions by computer would be lower than the average sensitivity of physicians, it would be difficult to justify the use of automated computer diagnosis.

For development of a successful CAD scheme, it is necessary not only to develop computer algorithms, but also to investigate how useful the computer output would be for radiologists in their diagnoses, how to quantify the benefits of the computer output for radiologists, and how to maximize the effect of the computer output on their diagnoses. Thus, large scale observer performance studies on radiologists using a reliable methodology such as Receiver Operating Characteristic (ROC) analysis are equally as important as the development of computer algorithms in the field of CAD. Therefore, the research and development of CAD has involved a team effort by investigators with different backgrounds such as physicists, radiologists, computer scientists, engineers, psychologists and statisticians.

One of the important events in the history of CAD is that R2 Technology has succeeded in commercialization of the first CAD system for detection of breast lesions in mammography based on licensing of CAD technologies from The University of Chicago, and that it obtained US Food and Drug Administration (FDA) approval for the clinical use of their system in 1998. Subsequently, clinical uses of the mammographic CAD system have begun at many screening sites for breast cancer in the United States, and more than 1500 CAD systems are in current use in assisting radiologists in the early detection of breast cancer at many hospitals, clinics and screening center’s around the world. It has also been reported that CAD has provided a gain of approximately 20% in the early detection of breast cancers on mammograms. In 2001, Deus Technologies developed another CAD system for detection of lung nodules on chest radiographs,
and then received FDA approval for its clinical use. In Japan, Mitsubishi Space Software has developed a CAD system with temporal subtraction of sequential chest radiographs and also for detection of lung nodules on chest images. A number of prototype systems for detection of pulmonary nodules in thoracic CT have been developed by manufacturers and are being evaluated at medical centers around the world. Recently, R2 Technology received FDA approval for their CAD system for detection of pulmonary nodules in CT.

The Development of Multi-Detector CT (MDCT) has produced a large number of CT images that may require additional time and effort in image interpretation by radiologists. It has been expected, therefore, that CAD would assist radiologists in reducing the reading time as well as in improving the diagnostic accuracy. Many investigators have attempted to develop CAD schemes for detection of pulmonary nodules by MDCT. Because the quality of MDCT images has been improved considerably in terms of Three Dimensional (3D) image information over that of conventional CT images with relatively thick slices, the performance of CAD schemes in the detection of nodules on MDCT images has generally been improved.

2.8 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, in which 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.

Machine learning in medical applications also gets increasingly more important and it is essential to research the various possibilities. Unfortunately, most articles that propose content based queries do not explain in detail which visual features have been used or are planned to be used. Sometimes, only a very vague description such as general texture and color or grey level features are given as in Mojsilović & Gomes (2000), Chang (1996) and Chheir et al. (2000). Basically all systems that do give details use color and grey level features, mostly in the form of a histogram reported by (Tang et al. 2000, Kwak et al. 2002 and Brodley et al. 1999). Local and global grey level features are used in Miller et al. (2003). Guld et al. (2001) use statistical distributions of grey levels for the classification of images and Qi & Snyder (1999) proposes a brightness histogram. As many of the images in the medical domain do not contain colors or are taken under controlled conditions, the color properties are not at all in the center of research and the same holds for invariants to lighting conditions. This can change when using such photographs as in dermatology. Pathologic images will need to be normalized in some way as different staining methods can produce different colors reported by Wirflinger et al. (2003).

Within radiology, the normalization of grey levels between different modalities or even for the same modality can cause problems when there is no exact reference point as is for the density of the CT, for example. Toga & Thompson (1998) illustrates the dependency of intensity values of the brain from the used modalities.
As color and grey level features are of less importance in medical images than in stock photography, the texture and shape features gain in importance. Basically all of the standard techniques for texture characterization are used from edge detection using canny operators reported by Veropoulos et al. (1998) to Sobel descriptors reported by Brodley et al. (1999). Brodley et al. (1999) also use Fourier descriptors to characterize shapes, and Antani et al. (2002), Bueno et al. (2002) and Mattie et al. (2000), use invariant moments and Antani et al. (2002) also uses scale-space filtering. Features derived from co-occurrence matrices are also frequently used reported by Brodley et al. (1999), Beretti et al. (2001) and Orphanoudakis et al. (1996), as well as responses of Gabor filters reported by (Tang et al. 2000 and Miller et al. 2003), wavelets reported by Kwak et al. (2002) and Markov texture characteristics reported by Mattie et al. (2000). In mammography, denseness is used for finding small nodules. It would be interesting to have a comparison of several texture descriptors. Many of them model the same information and will most likely deliver very similar results.

Using segments in the images also allows using spatial relationships as visual descriptors of the image. This is often proposed by Chu et al. (1994), Petrakis et al. (2002) and El-Kwae et al. (2000), but rarely any detail is given on how to obtain the objects or segments in the images, and hence, it does not permit to judge whether an implementation is possible. Another article not taking into account the problems of automatic segmentation is Petrakis et al. (2002). Sinha & Kangarloo (2002) and Bucci et al. (1996) propose the use of Eigen images for the retrieval of medical images in analogy to Eigen images for face recognition. These features can be used for classification when a number of images for each class exist. Still, the features are purely statistical and it is hard to actually
explain the similarity of two images based on these features which can
more easily be done for a histogram intersection, for example.

Tissue Time Activity Curve (TTAC) curves for the retrieval of PET image is used in Cai et al. (2000). These are not really image features but rather one dimensional temporal signal that is compared. However, the results seem to be good. Similar to general CBIR, semantic features are proposed for visual similarity queries with medical images reported by (Tang et al. 2000). But again, it comes down to simple textual labels attached to the images and a mapping between the text and the low level features.

Most systems do not give many details on the distance measurements or comparison methods which most likely implies an Euclidian vector space model using either a simple Euclidean distance (L2) or something close such as city block distance or L1. To efficiently work with the Euclidian distances even in large databases, the dimensionality is often reduced. This can be done with methods such as Principal Component Analysis (PCA) or Minimum Description Length (MDL) that tries to reduce the dimensionality while staying as discriminative as possible. In principle, redundant information is removed but this can also remove small but important changes from the feature space. Techniques such as KD trees and R-trees are also used in medicine for efficient access to such a large feature spaces.

On the other hand, statistical methods are used for the comparison of features that can be trained with existing data and that can then be used on new, incoming cases. These can be neural networks for the classification of mammography images or on images extremely reduced in size. Other statistical approaches use Bayesian networks or Hidden Markov Models (HMMs). An associative computing approach is proposed for retrieval assuming that a query is performed with a local part of the images.
An Receiver Operating Characteristic (ROC) curve for the comparison of methods is used. This is well known in the medical domain and easily interpretable.

Already in the general image retrieval domain it is difficult to compare any two retrieval systems. For medical image retrieval systems, the evaluation issue is almost non-existent in most of the research articles. Still, there are several articles on the evaluation of imaging systems in medicine or on general evaluation of clinical systems and the problems with it.

A single example result does not reveal a great deal about the real performance of the system and is not objective as the best possible query result can be chosen arbitrarily by the authors. This problem in retrieval system evaluation is described in detail in Muller et al. (2001). Most other system evaluations show measures with a limited power for comparison. In Brodley et al. (1999), the precision of the four highest ranked images is used which does not reveal much about the number of actually relevant items and gives very limited information about the system. Siha & Kangarloo (2002) measures the number of times a differently scaled or rotated image retrieves the original which is also not very close to medical image retrieval reality.

2.9 CONTENT BASED MEDICAL IMAGE RETRIEVAL SYSTEMS

There are two ways that medical images are retrieved, text based and content based methods. In text based image retrieval, the images are retrieved by the manually annotated text descriptions and traditional database techniques to manage images are used. In content-based image retrieval, the images are retrieved on the basis of features such as color,
So far, a variety of medical image retrieval systems have been developed using either method (text-based or content-based) or combining two methods. A rough classification of medical image retrieval methods is shown in Figure 2.1.

![Figure 2.1 Classification of medical image retrieval methods](image)

The two important medical image retrieval systems are SPIRS and IRMA. The Spine Pathology & Image Retrieval System (SPIRS) was developed at the U. S. National Library of Medicine to retrieve x-ray images from a large dataset of 17,000 digitized radiographs of the spine and associated text records. Users can search these images by providing a sketch of the vertebral outline or selecting an example vertebral image and some relevant text parameters. Pertinent pathology on the image/sketch can
be annotated and weighted to indicate importance. This hybrid text-image query yields images containing similar vertebrae along with relevant fields from associated text records, and hence, it allows users to examine the pathologies of vertebral abnormalities.

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, improved and evaluated over time. This section will introduce several major medical content-based retrieval systems.

2.9.1 Automated Search and Selection Engine with Retrieval Tools

Automatic Search and Selection Engine with Retrieval Tools (ASSERT) reported by Chi-Ren Shyu et al. (1999) is developed by Purdue University, Indiana University, and University of Wisconsin Hospital, USA. The system uses image database formed by High Resolution Computed Tomography (HRCT) of lung where a rich set of textural features derived from the disease bearing regions is important for the characterization of the images. The ASSERT system uses a physician in the loop approach to retrieve 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. Hence ASSERT has extremely high data entry costs, which prohibit its application for clinical routine
In ASSERT, there are two phases namely Image Archiving Phase and Image Retrieval Phase. To archive an image into the database, a physician delineates the Pathology Bearing Regions and any relevant anatomical landmarks. This interaction takes only a few seconds for a trained domain expert (a radiologist). In the meantime, a lung region extraction algorithm is applied to the image to determine the boundary of the lungs. The system then executes a suite of image processing algorithms to create attribute vectors that characterize the PBRs individually and the portion of the image that consists of just the lung regions. These attributes are subject to a sequential forward selection algorithm to reduce the dimensionality of the attribute space while retaining the ability to accurately classify each image as belonging to its associated disease pattern.

2.9.2 CasImage

CasImage is another medical CBIR system. The system was developed in University Hospital of Geneva, Switzerland. The CasImage database consists of 8723 images and represents a mixture of images from clinical routine and drawings from medical education. The image Database contains variety of images from CT, MRI, and radiographs. The CasImage data set also includes color images which demand the additional features to capture their specific properties. The system allows content based image retrieval of medical images, including MR images. CasImage is integrated in a PACS environment. Searching and retrieval in system is made via medGIFT searching system. Three types of features are extracted from the images such as the local and global color characteristics and the texture of the images, so that the searching is enabled. The interesting part of the system is that it uses a combination of textual annotation and visual features to perform the search.
Furthermore, there are images with secondary added contents such as pseudo colourings of segmented ultrasound images or manually placed marker arrows for operation planning, and do not represent original data. The CasImage system, which has been integrated into a PACS environment, contains a teaching and reference database. Combinations of textual labels and visual features are used for medical image retrieval. CasImage is a good idea for complex medical CBIR system, since it combines both, the textual information and visual features of the images. But the problem, here, is that textual information is not added automatically, so the entire system is not fully automated.

2.9.3 MedGIFT

The MedGIFT system is an image retrieval engine reported by (Muller et al. 2003). As medGIFT is a domain specific search tool, the user interface has different requirements from other domains. One important part is the display of not only thumbnail images for the browsing but also the text of the diagnosis. It is based on the open source system GNU Image Finding Tool (GIFT), outcome of the Viper project of the University of Geneva. The MedGIFT project started at the medical faculty of the University of Geneva, Switzerland in 2002 and located in the Institute for Business Information Systems at the HES SO in Sierre (Valais), Switzerland. This system offers components for content-based indexing and retrieval of images such as feature extraction algorithms, feature indexing structures and a communication interface called Multimedia Retrieval Mark-up Language (MRML). MedGIFT retrieval system does not require classification and a priori knowledge for retrieval. It is a retrieval engine and encompassing framework for the retrieval of images by their visual content only. The features themselves are supposed to model the visual similarity of the images. GIFT uses techniques from
text retrieval such as frequency based feature weights, inverted file indexing structures, and relevance feedback mechanisms. Main developments are on the integration of various new components around GIFT to create a domain-specific search and navigation tool. The GIFT framework is given in Figure 2.2.

![GIFT Framework](image)

**Figure 2.2 GIFT framework**

The MedGift system is specializing in applying different approaches to image retrieval in the medical domain. Although the system is purely visual based, it demonstrates how the visual methods are successful in restricted domains. The system requires an example or set of images which could be generated from a previous text based search.

The name stems originally from the use of the GNU Image Finding Tool for medical applications. Over the years the GIFT has been used less frequently and a large set of tools and applications have been developed to advance the field of medical visual information retrieval. 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. To represent images with visual features, four feature groups are chosen, namely

1. Local and global texture features based on responses of Gabor filters, and
2. Color/grey scale characteristics on a global image scale and locally within image regions.

The MedGIFT system retrieves images on the basis of local and global similarities in gray level and texture. The system was evaluated for medical image retrieval in the context of the image-CLEF competition. Some of the evaluation results concern the number of gray levels that delivers the best retrieval results. This number is surprisingly low for optimal image retrieval. The system was used with several gray level quantization’s, and its performance was evaluated against a standard of reference generated by a radiologist.

2.9.4 Yottalook

Yottalook is a free medical imaging search engine that provides decision support at the point of care using proprietary relevance and ranking algorithms by Montage Healthcare Solutions, Inc. Yottalook is designed to provide the practicing radiologists the most important and most relevant information they need at the time of patient care.

Yottalook performs multilingual search in thirty three languages to retrieve images from peer reviewed journal articles on the Web. It uses Google’s indexing technology and a proprietary software called iVirtuoso
for natural query analysis, semantic ontology generation and determination of the relevance. Natural query analysis generates keyword from search queries. Yottalook uses an enhanced version of the RSNA’s RadLex® medical ontology to identify relationships or synonymous terms. This is known as semantic ontology generation. Relevance is automatically derived using a relevance algorithm (part of the iVirtuoso software) and is used to rank the retrieved results.

Yottalook is based on core technologies developed by Montage Healthcare Solutions to achieve optimized search results. First is automated analysis of the search term to understand what the radiologist is trying to look for, and this core technology is called natural query analysis.

Yottalook has also developed a thesaurus of medical terminologies that not only identifies synonyms of terms but also defines relationships between terms. This second core technology is called semantic ontology" and is based on existing medical ontologies.

Third core technology is relevance algorithm for image search that differentiates medical terms from other words in text associated with medical images and uses them to create ranking for Yottalook image search. The fourth core technology is a specialized content delivery system called Yottalinks that provides high yield content based on the search term. This content may also be provided by a third party vendor licensing Yottalook search. Yottalook can be integrated with any web based medical application so that context relevant information is provided to the physician at the point of care.

Yottalook provides intelligent search capabilities to look for peer reviewed radiology content including journals, teaching files, CME, etc. This search engine is optimized to be used as a decision support tool at the
time when you need the information quickly. It uses semantic ontology of medical terminologies that not only identifies synonyms of terms but also defines relationships between terms to expand the search results.

2.9.5 iMedline

iMedline is a multimodal search engine under development at National Library of Medicine with goals to retrieve images from biomedical literature relevant to text and image queries and link evidence automatically extracted from clinical articles to patients’ cases. Along with the traditional elements of search results display, such as titles and author names, iMedline provides captions of the retrieved images and short summaries of the retrieved abstracts. For the document retrieval task, iMedline uses NLM’s Essie search engine. Essie is a phrase-based search engine with UMLS-based query expansion and probabilistic relevance ranking that exploits the document structure. The iMedline user interface provides the Essie search options and displays search results in grid or list views.

In this system, retrieved images can be used as a query input to the relevance of search results. Currently, the image retrieval engine uses low-level visual features, such as color, texture, and shape as the primary building blocks of the visual content in an image. The features are then transformed into visual keywords, which annotate images with a set of labels that indicate the membership of local image regions/patches in various image categories. Similarity between a query image and database images is measured as a weighted linear combination of different features.

The feature weights are updated to reflect the similarity rank of images. Build tools employing a combination of text and image features to enrich traditional bibliographic citations with relevant biomedical images, charts, graphs, diagrams and other illustrations, as well as with patient-
oriented outcomes from the literature. Improve the retrieval of semantically similar images from the literature and from image databases, with the goal of reducing the semantic gap that is a significant hindrance to the use of image retrieval for practical clinical purposes.

2.9.6 Automatic Linguistic Indexing of Pictures in Real Time

Automatic Linguistic Indexing of Pictures in Real Time (ALIPR) is an automatic image annotation system proposed by (Li & Wang 2006, 2008) and recently made public for people to try to have their pictures annotated. As mentioned earlier, presence of reliable tags with pictures is necessary for text-based image retrieval. As part of the ALIPR search engine, an effort to automatically validate computer generated tags with human given annotation is being used in an attempt to build a very large collection of searchable images.

A real time image annotation system ALIPR (automatic linguistic indexing of pictures—real time) has been proposed in Li & Wang (2006). ALIPR inherits its high-level learning architecture from ALIP. However, the modeling approach is simpler, hence leading to real time computations of statistical likelihoods. Being the first real-time image annotation engine, ALIPR has generated considerable interest for real-world applications.

ALIPR is on a mission to assign relevant tags to digital images based on their content, and wants you to help it learn. The system has enabled automatic photo tagging and visual search on the web, thus allowing the users to interpret imaging findings. Much of radiological practice is currently not based on quantitative image analysis, but on heuristics to guide physicians through rules of thumb. The training process in ALIPR system is illustrated in Figure 2.3.
Figure 2.3 Training process - ALIPR

- Feature Extraction
- Region Segmentation
- Statistical Modelling by D2-Clustering

Training DB for Concept 1

Training DB for Concept 2

Training DB for Concept M

Textual description about Concept 1

Textual description about Concept 2

Textual description about Concept M

A trained dictionary of semantic concepts

Model about concept 1

Model about concept 2

Model about concept M
ALIPR annotates images by looking at a large database of images which are already annotated and categorized in groups. ALIPR group images in categories, also called semantic concepts. ALIPR defines models of image features for each concept by making a signature of each image in a concept, and build a generative model from it. A signature consists of color and texture features. ALIPR can compare any image to the category models in their training database, then examines the correlation with the categories, and annotates the image similar to the most matching category. Hence it is able to annotate the query image with the annotation words from the database. This results in a more semantic way of searching, since users can just query for concepts they have in mind. ALIPR is a valuable contribution to the CBIR field. Their real-time annotation is in the first place possible because their algorithm characterizes image features in a statistical distribution, without looking at each individual object within images. Secondly, they have a cumulative approach, only for images of new concepts, the annotation algorithm has to be trained, and previous concepts are stored in profiling models.

2.9.7 Flexible Image Retrieval Engine

FIRE system handles different kinds of medical data as well as nonmedical data like photographic databases. In FIRE, each image is represented by a set of features. To find images similar to a given query image, the features from the images in the database are compared to the features of the query image using an appropriate distance measure.

The main aim of FIRE is to investigate different image descriptors and evaluate their performance. In FIRE, query expansion is implemented as automatic relevance feedback. The user specifies a number of images G that he expects to be relevant after the first query. Then a
query is processed in two steps. First the query is evaluated and the first G images are returned. These G images are automatically used as the set of relevant images Q+ to requery the database and the K best matches of this query are returned.

FIRE is a suite of two programs. A server, a web client which will be running on a web server. The server itself can be running on any computer. If a query is performed from the web-interface the client sends the filename of the image to be queried to the server, the server loads this image and the features extracted for it and then compares these features to all the images from the database using the selected distance measures. Then the server sends the query result to the client, which presents the result to the user. In addition to the images resulting from the query some query performance measures are calculated and also presented to the user.

In FIRE, different features are available to represent images. In this system query by example image is implemented using a large variety of different image features that can be combined and weighted individually and relevance feedback can be used to refine the result. A weighted combination of these features admits very flexible query formulations and helps in processing specific queries.

2.9.8 RadLex - Retrieval of Radiology Information Resource

Retrieval of Radiology Information Resources (RadLex) was proposed by the Radiological Society of North America (RSNA). It provides the radiologist with a unified language to organize and retrieve images, imaging reports and medical records. Recently RadLex has been extended by terms and synonyms for imaging signs. However, the hierarchical relations are partly ambiguous, e.g., within the class
“anatomical entity” all the terms “hand” and “finger”, “arm” and “forearm” can be found at one level, which contradicts the part of relation.

RadLex was developed to create a unifying source for medical imaging terminology and currently contains more than 32,000 standardized terms used in radiology reports. It contains not only domain knowledge but also lexical information such as synonymy. RadLex terminology helps the analysis of radiological information, allows uniform indexing of image databases, and enables structuring medical image information. RadLex also provides a comprehensive and technology friendly replacement for the ACR Index for Radiological Diagnoses. It unifies and supplements other lexicons and standards, such as SNOMED-CT and DICOM. In radiology, RadLex has been created to standardize annotation of images and thus reuse. One example of RadLex use is structured reporting where templates for many domains and types of observations exist and reusing the templates can significantly increase efficiency. Another are of RadLex use is information retrieval, where semantic knowledge of cases can make it much easier to find a specific case with a particular visual observation or a specific disease, again taking into account synonyms and also sub categories of a disease or a specific imaging modality

RadLex (Radiology Lexicon) is a controlled terminology for radiology-a single unified source of radiology terms for radiology practice, education, and research. RadLex enables numerous improvements in the clinical practice of radiology, from the ordering of imaging exams to the use of information in the resulting report. One of the primary goals of this project is to create a technology that can be used to annotate, index and retrieve content from MIRC.

RadLex development is supported both by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) and by the cancer
Biomedical Informatics Grid (caBIG) project, a large NIH sponsored effort to develop unified computing infrastructure for clinical trials.

2.9.9 National Health and Nutrition Examination Survey

The Second National Health and Nutrition Examination Survey (NHANES II) is a system developed by National Library of Medicine, USA. 17,000 cervical and lumbar spine X-ray images from the database. Access to the text and image data from NHANES II is being provided by R&D biomedical multimedia database system, Web based Medical Information Retrieval System. 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.

The National Health and Nutrition Examination Survey (NHANES) is a program conducted regularly by the National Center for Health Statistics (NCHS) with the aim of determining the prevalence of selected diseases and the associated risk factors. The National Library of Medicine (NLM) maintains the data collected by the second survey (NHANES II). This database contains 17,000 digitized cervical and lumbar spine x-ray images (sagittal view) as well as related text metadata: demographic information, anthropometric data, and health and medical history. The NHANES II collection has two kinds of images (cervical and lumbar spine) that are very similar within each group and require localized
descriptors at various levels of detail to distinguish among them. This large data collection is a valuable resource for the study and education of bone morphometry and musculoskeletal diseases. After careful study of the NHANES II image collection medical experts have previously determined that the Anterior Osteophyte (AO) is a pathological feature that can be reliably and consistently detected in the data set, along the anterior superior or inferior edges of the vertebral bodies. It is also a valuable resource for research and development in Content-Based Image Retrieval (CBIR), a technique that addresses the problem of indexing and retrieval of visual data through visual features instead of text.

The goals of NHANES include estimating prevalence of selected diseases, monitoring disease trends, and studying the relationship between nutrition and health. While the intent of the NHANES dataset is not clinical decision support, the combination of health survey and imaging data is useful in retrieving examples of abnormal vertebral images for research and education. For instance, SPIRS can be used to determine what types of features (e.g., protrusion on the anterior edge of the cervical vertebra) are consistently associated with a particular symptom (e.g., neck discomfort).

2.9.10 Spine Pathology and Image Retrieval System

Spine Pathology and Image Retrieval System (SPIRS) is a Web based distributed content based image retrieval framework that supports hybrid visual and text queries, and may be applied to various applications in medicine. It implements novel shape representation and similarity matching embedded with an index tree that allows efficient retrieval. SPIRS is a generalizable framework that consists of four components: a client applet, a gateway, an indexing and retrieval system, and a database of images and associated text data. SPIRS permits exploration of a large
biomedical database of digitized spine X-ray images and data from a national health survey using a combination of visual and textual queries. The prototype system is demonstrated using text and imaging data. SPIRS is built using open standards and is simultaneously developed as a service, which enables its integration with other complementary information retrieval systems.

The SPIRS framework is also being used to query an image database of cervicographic images (cervigrams) created by the National Cancer Institute (NCI) and the National Library of Medicine (NLM) for the study of uterine cervix cancer. The database contains approximately 100,000 cervigrams taken as part of the Guanacaste and ALTS projects that study the natural history of human papillomavirus infection and cervical neoplasia. In addition to cervigrams, correlated clinical, cytologic, and molecular information are also available. Unlike NHANES II, the intent of this dataset is to be used in clinical practice as a method for identifying whether patients have precursors to cervical cancer given the appearance of their cervigram.

SPIRS aims to capture query semantics through support of advanced mechanisms like multiple partial shape matching and iterative querying that provides simple yet effective relevance feedback to the system. SPIRS is built using open standards and is simultaneously developed as a service, which enables its integration with other complementary information retrieval systems. Initial evaluation of the system has shown that the combination of partial shape matching and relevance feedback significantly improves the system’s ability to retrieve similar results. Since its original release in August 2006, SPIRS has undergone several revisions that have added support for additional retrieval
algorithms (i.e., Fourier descriptors) and an improved, cleaner user interface for posing visual and text queries.

SPIRS demonstrates recent developments in shape representation and retrieval from a large data-set of digitized x-ray images of the spine and associated text records from the Second National Health and Nutrition Examination Surveys (NHANES II). Rather than limiting queries to textual keywords or relations (as available in WebMIRS, for example), CBIR enables users to provide a query sketch/image, which is then used to find similar images of the same modality, anatomical region, and disease along with the associated text records

SPIRS provides a Web-based interface for image retrieval using the morphological shape of the vertebral body. A query editor enables users to pose queries either by sketching a unique shape, or by selecting or modifying an existing shape from the database. Additional text fields enable users to supplement visual queries with other relevant data (e.g., anthropometric data, quantitative imaging parameters, patient demographics etc.) These hybrid text-image queries may be annotated with pertinent pathologies by selecting and weighting local features to indicate importance. Query results appear in a customizable window that displays the top matching results and related patient data.

SPIRS is automated, easily accessible, and integratable with other complementary information retrieval systems. The system supports the ability for users to intuitively query large amounts of imaging data by providing visual examples and text keywords and has beneficial implications in the areas of research, education, and patient care.
SPIRS addresses current limitations of CBIR implementations by

1) utilizing open standards to communicate among components, which can be extended to support data encryption to meet privacy regulations.

2) implementing algorithms that automate shape extraction and representation;

3) utilizing feature indexing methods for efficient retrieval

4) combining text and image feature queries for hybrid queries; and

5) performing evaluation of a variety of whole and partial shape retrieval algorithms using a large biomedical dataset.

SPIRS utilizes the query by example paradigm, in which users provide an example shape or image to the system as a method of finding similar images. In addition, it combines visual and text queries to provide users with greater flexibility in retrieving relevant results. SPIRS is a working proof of concept that demonstrates the capability of accommodating large amounts of text and imaging data that is expected in the modern research and healthcare environments. The distributed architecture of SPIRS is illustrated in Figure 2.4. User clients communicate through the Internet to a gateway, which connects each client to the biomedical database and indexing and retrieval algorithms. The gateway acts as a mediator, authenticating users and ensuring that the data is sent to the appropriate user.
2.9.11 FigureSearch

FigureSearch search engine, developed at the University of Wisconsin at Milwaukee, is a component of the ask Hermes system, and is a tool devised to improve the quality of patient care by providing information to physicians at point of care. It uses the Lucene text indexing and search technology to search online medical articles and generates a list view of results. Images are displayed on the left while the title, authors, figure caption and summary are displayed on the right. The search engine separates itself from others with its ability to automatically generate summaries from papers (on the purpose, experimental procedure, outcome and conclusion) using sentences from the main text.

The FigureSearch system uses a supervised machine-learning algorithm for classifying clinical questions, and Lucene for information retrieval over the published medical literature to generate a list view of the results with relevant images, abstracts and summaries.
2.9.12 GoldMiner

GoldMiner Global is a multilingual radiologic image search engine. It incorporates the search technologies and image library of the ARRS GoldMiner system which indexes and retrieves images from peer-reviewed sources by using keyword and concept-based search techniques. GoldMiner has a simple, Web based user interface. The search engine applies two distinct retrieval techniques namely, keyword search and concept search; and returns those images found using either technique. The complementarity of these techniques is a unique aspect of the search engine.

First, GoldMiner searches for the given term as a keyword, that is, as a case insensitive string. For example, the search term gallstone would match any figure with a caption that contained the word gallstone, Gallstone, or GALLSTONE. It would not, however, match text that contained “gall stone” (two words) or “gallstones” (the plural form). The second, more powerful, technique is concept based search. With this technique, GoldMiner uses the knowledge contained in the UMLS Meta thesaurus to search using the meaning of the specified term. The Meta thesaurus contains lexical variants of terms, such as gallstone and gallstones, and also contains synonyms, such as cholelithiasis. The Meta thesaurus also recognizes that gallstones are a subtype of gallbladder disease. Thus, when a user enters gallstone as a search term, GoldMiner understands that images labeled with gallstone, gallstones, and cholelithiasis should be retrieved, too.

The Goldminer searches figure captions to retrieve images from 11000 open-access peer-reviewed journal articles from the websites of American Roentgen Ray Society (ARRS), the American Society of Neuroradiology (ASN), the British Institute of Radiology (BIR), and the
Radiological Society of North America (RSNA). It also allows searches using multiple keywords.

Unlike most internet search engines, ARRS GoldMiner understands medical vocabulary. It uses sophisticated techniques from the U.S. National Library of Medicine (part of NIH) to discover medical concepts in free-text figure captions and uses that information to quickly retrieve relevant images. GoldMiner incorporates standardized vocabularies, such as the Medical Subject Heading (MeSH) terms, which are used to index the medical literature in MEDLINE and PubMed.

ARRS GoldMiner recognizes abbreviations, synonyms and kinds of diseases. Not only does it know that "renal calculi" and "kidney stones" mean the same thing, it also knows that renal calculi are a type of kidney disease. GoldMiner searches by both concepts and keywords. As a result, searches for "renal calculi" and "kidney stones" won't find exactly the same entries. If a figure caption says "stones are seen in the kidney", the words "stones" and "kidney" will be indexed, but not necessarily the concept of kidney stones / renal calculi. As with a conventional search engine, GoldMiner's search results can depend on the presence of specific words in the figure captions. GoldMiner includes the ability to limit, or filter, search results by imaging technique, patient age group, and patient sex. From each figure's caption text, the search engine attempts to identify the imaging technique and the patient's age and sex.

2.9.13 BRISC

The BRISC Really IS Cool (BRISC) project which provides a simple base for future work in pulmonary nodule detection and diagnosis. BRISC provides a framework for texture feature extraction and similarity comparison of Computed Tomography (CT) lung nodule images. The
current design allows for importing, browsing, and retrieval of lung nodule images from the Lung Image Database Consortium (LIDC) database. The Lung Image Database Consortium contains 149 unique pulmonary modules that have been segmented and annotated by up to four different radiologists amounting to a total of 2020 images. These images were taken from a total of 90 Computer Tomography studies of the Chest, each containing between 100 and 400 Digital Imaging and Communication (DICOM) Images. The BRISC System overview is shown in Figure 2.5.

![Figure 2.5 System overview of BRISC](image)

**Figure 2.5 System overview of BRISC**

BRISC system retrieves images of similar nodules from a collection prepared by the Lung Image Database Consortium. The system

1. Extracts images of individual nodules from the LIDC collection based on LIDC expert annotations,

2. Stores the extracted data in a flat XML database,
(3) Calculates a set of quantitative descriptors for each nodule that provide a high-level characterization of its texture, and

(4) Uses various measures to determine the similarity of two nodules and perform queries on a selected query nodule.

Four radiologists marked the contour of nodules and assigned nine semantic terms to each nodule, namely calcification, internal structure, lobulation, malignancy, sphericity, speculation, subtlety, texture and margin. Calcification and internal structural are nominal while the other seven annotations are ordinal. Calcification contains six different categories; internal structure contains four different categories, and the other seven annotations are rated on a scale from one to five each.

Local Gabor and Markov methods of texture characterization perform better than global Haralick co-occurrence methods. The precision of image retrieval can be very high, and so this technique has the potential to be useful as an adjunct to radiologist decision making in the context of pulmonary nodules in CT images. This work was supported by the National Science Foundation.

BRISC will be utilized by radiologists to increase their ability to diagnose pulmonary nodules by retrieving similar cases as it allows the radiologists to directly compare the questionable nodule to other cases.

2.9.14 Yale Image Finder

Yale Image Finder (YIF) developed at Yale University, searches text within biomedical images, captions, abstracts and titles to retrieve images from biomedical journal papers. It uses optical character recognition to recognize text in images in both landscape and portrait
modes and then validates the extracted text against content extracted from corresponding full-text articles.

Yale Image Finder is a publicly accessible search engine featuring a new way of retrieving biomedical images and associated papers based on the text carried inside the images. Image queries can also be issued against the image caption, as well as words in the associated paper abstract and title. A typical search scenario using YIF is as follows: a user provides few search keywords and the most relevant images are returned and presented in the form of thumbnails. Users can click on the image of interest to retrieve the high resolution image. In addition, the search engine will provide two types of related images: those that appear in the same paper, and those from other papers with similar image content. Retrieved images link back to their source papers, allowing users to find related papers starting with an image of interest.

A unique capability of YIF is that users can access related images from the associated papers by directly comparing image content. It uses Lucene technology to index, search, and rank search results.

YIF indexes over 140 000 images from over 34 000 open access papers from PubMed Central. The system is updated on a regular basis. YIF offers more comprehensive research results by searching over text that may not be present in the image caption, and offers the ability to find related images and associated papers by directly comparing image content.

2.9.15 ImageCLEF

ImageCLEF aims to provide an evaluation forum for the cross language annotation and retrieval of images. Motivated by the need to support multilingual users from a global community accessing the ever
growing body of visual information, the main goal of ImageCLEF is to support the advancement of the field of visual media analysis, indexing, classification, and retrieval, by developing the necessary infrastructure for the evaluation of visual information retrieval systems operating in both monolingual, cross language and language-independent contexts. ImageCLEF aims at providing reusable resources for such benchmarking purposes.

The medical retrieval task of ImageCLEF 2011 uses a subset of PubMed Central containing 231,000 images. There will be three types of tasks in 2011.

**Modality Classification**

Previous studies have shown that imaging modality is an important aspect of the image for medical retrieval. In user studies, physicians have indicated that modality is one of the most important filters that they would like to be able to limit their search by. Many image retrieval systems like Goldminer, Yottalook etc allow users to limit the search results to a particular modality. However, this modality is typically extracted from the caption and is often not correct or present. Studies have shown that the modality can be extracted from the image itself using visual features. Additionally, using the modality classification, the search results can be improved significantly. The measure used for this sub task will be classification accuracy. The results of the modality classification can be used to filter the search in the next sub task.

**Ad-hoc image-based retrieval**

This is the classic medical retrieval task, similar to those in organized in 2005-2010. Participants will be given a set of 30 textual
queries with 2-3 sample images for each query. The queries will be classified into textual, mixed and semantic, based on the methods that are expected to yield the best results.

**Case-based retrieval**

Case based retrieval was first introduced in 2009. This is a more complex task. In this task, a case description, with patient demographics, limited symptoms and test results including imaging studies, is provided (but not the final diagnosis). The goal is to retrieve cases including images that might best suit the provided case description. Unlike the ad-hoc task, the unit of retrieval here is a case, not an image.

**2.9.16 IRMA – Image Retrieval for Medical Applications**

Image Retrieval for Medical Applications (IRMA) system, developed at Aachen University of Technology, Germany, aims to integrate text and image-based features for medical image retrieval. In IRMA, images are classified according to anatomy, bio-system, imaging direction and modality of the image (x-ray, CT, MRI, etc.). It applies differential weighting of image features for computer-aided diagnosis. The image features are derived from co-registered training images. IRMA uses semantic layers to describe an image. These layers comprise multi-scale descriptions of the raw image data, extracted features, visual content and its spatial layout within the image. It supports text queries as well as image query by example (QBE) and has been tested on mammograms and bone x-ray images. IRMA splits the retrieval process into seven consecutive steps as shown in Figure 2.6. Each step represents a higher level of image abstraction, reflecting an increasing level of image content understanding.
The IRMA concept is based on a strict logical and algorithmic separation of the following steps to enable complex image content understanding:

1. image-categorization (based on global features)
2. image-registration (in geometry and contrast)
3. feature extraction (based on local features)
4. feature selection (category and query dependent)
5. indexing (multiscale blob-representation)
6. identification (incorporate a-priori knowledge)

7. retrieval (on blob-level)

The categorization step aims at determining for each image entry the imaging modality and its orientation as well as the examined body region and functional system. For that, a detailed hierarchical coding scheme was developed, which exceeds the complexity of existing tags of the Digital Imaging and Communications in Medicine (DICOM) standard, but could be consistently integrated to supplement the standard. Automatic categorization is based on a reference database of 10,000 images selected arbitrarily from clinical routine and manually classified by experienced radiologists. This database integrates medical knowledge at a low level of abstraction. The automatic categorization of query by example images is performed by combining DICOM header information and global image features, i.e. features describing the entire image. However, categorization in IRMA is not exclusive. Subsequent steps of processing are applied for the most likely categories.

Registration in geometry (rotation, translation, scaling) and contrast generates a set of transformation parameters that is stored for the corresponding image in each of its likely categories. In consent with Tagare et al. (1997), registration is based on prototypes which are manually defined for each category, and further incorporate medical expert knowledge into the IRMA system. The transformation is not performed explicitly at this step of processing. Instead, the generated parameters are utilized at higher layers of abstraction.

The feature extraction step derives local image descriptions, i.e. a feature value (or a set of values) is obtained for each pixel. These can be
category free (e.g. resulting from edge detection or regional texture analysis) or category-specific, such as the application of an active shape model that explicitly uses a priori knowledge derived from the respective category.

Decoupling feature selection from feature extraction allows integrating both image category and querying context into the abstraction process. For instance, the same radiograph might be subject to fracture or cancer examination, resulting in a contour-based or texture-based combination of features, the so called feature sets, such as the contour set or texture set, respectively. In order to avoid exhaustive computation during query processing, these feature sets are pre computed for each image in each likely category.

Indexing provides an abstraction of the previously generated and selected image features, resulting in a compact image description. According to the selected feature set, this is done via clustering of similar image parts into regions represented by their second area moment description as ellipses ("blobs"). In contrast to the Blobworld approach, this is done at multiple resolutions yielding a multi scale blob representation of the image. The hierarchical indexing enables the processing of ROIs, which are marked by the user when issuing a query. In contrast to existing approaches to medical image retrieval, the ROI is not determined a priori. In other words, the incorporation of medical mid-level knowledge becomes possible.

According to Tagare et al. (1997), an essential requirement for satisfying medical queries is a high level of image understanding offering object oriented retrieval. The identification step provides linking of medical a priori knowledge to certain blobs generated during the indexing
step. It relies on the prototypes defined for each category, which are labeled locally by medical experts, and the corresponding parameters for geometry and contrast registration. Thus, identification is the fundamental basis to introduce high level image understanding by analyzing regional or temporal relationships between the blobs.

In IRMA, the retrieval itself is processed either on the abstract blob level or referring to identified objects. Only the retrieval step requires online computations while all other steps can be performed automatically in batch mode at entry time of an image into the database. This, of course, requires offline computation of all paths generated by the categorization and the feature selection step.

Currently, the IRMA reference database holds 3,879 images that have been labeled according to the IRMA code. Since the frequency of imaged body regions reflects the clinical situation, chest radiographs occur most often. In addition, the sub-region classes are distributed irregularly. The IRMA system supports modular design of arbitrary retrieval algorithms. Modularity easily enables the verification of isolated processing steps and allows the reuse of programs for various experiments and applications. All kinds of features (global, local, blob trees) are uniformly accessible, which results from the general and flexible feature model provided by IRMA. This includes a straightforward definition of a set of images. The system supports the automatic transfer of new and updated processing components into the pool of retrieval algorithms available to the physicians. In addition, new algorithms can quickly access the image database shortening the cycles between development and testing.
IRMA has developed and implemented high level CBIR methods with application to medico diagnostic tasks on radiological image archives.

Current image data consists of radiographs, with future plans to include medical images from arbitrary modalities. It aims to provide visually rich image management through CBIR techniques applied to medical images using intensity distribution and texture measures taken globally over the entire image. This approach permits queries on a heterogeneous image collection and helps identify images that are similar with respect to global features, e.g., all chest x-rays in the AP (anterior-posterior) view. The IRMA system lacks the ability to find particular pathology that may be localized in particular regions within the image.

The Image Retrieval in Medical Applications (IRMA) project has the following goals

1. automated classification of radiographs based on global features with respect to imaging modality, body orientation with respect to the x-ray beam (e.g., “anterior posterior” or “sagittal”), anatomical body region examined, and the biological system under investigation, and

2. identification of local image features including their constellation within a scene and are relevant for medical diagnosis. These local features are derived from a priori classified and registered images that have been segmented automatically into a multi scale approach.

IRMA analyzes content of medical images using a six layer information model:
(1) raw data,
(2) registered data,
(3) feature,
(4) scheme,
(5) object, and
(6) knowledge.

The IRMA system that is currently available via the Internet retrieves images similar to a query image with respect to a selected set of features. These features can, for example, be based on the visual similarity of certain image structures. Currently, the image data consists of radiographs. It uses a reference database of 10,000 images categorized by image modality, orientation, body region and biological system.

2.10 EVALUATION OF CONTENT BASED MEDICAL IMAGE RETRIEVAL SYSTEMS

Content Based Medical Image Retrieval Systems are compared using the criteria of gaps. Semantic gap refers to the difference in the level of image understanding at the human level versus the computer. Visual features derived from images contain low level pixel information such as color, edge information, textures, etc. Mapping these features to high level concepts such as spatial relationships between organs, identification of anatomical features, disease characterization is the main challenge faced by current image retrieval systems. Four types of gaps were identified in content based medical image retrieval systems are content, feature, usability, and performance gaps. Here we compare the current systems using this framework and is given in the following Table 2.1.
<table>
<thead>
<tr>
<th>Feature/Content</th>
<th>GoldMiner</th>
<th>FigureSearch</th>
<th>BioText</th>
<th>Yottalook</th>
<th>YIF</th>
<th>IRMA</th>
<th>iMedline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic</td>
<td>UMLS concepts</td>
<td>Not addressed</td>
<td>Not addressed</td>
<td>Not addressed</td>
<td>Not addressed</td>
<td>Not addressed</td>
<td>Semantic layers of Blobworld</td>
</tr>
<tr>
<td>Context</td>
<td>Broad</td>
<td>Broad</td>
<td>Broad</td>
<td>Broad</td>
<td>Broad</td>
<td>Narrow</td>
<td>Broad</td>
</tr>
<tr>
<td>Type</td>
<td>Text-only</td>
<td>Text-only</td>
<td>Text-only</td>
<td>Text-only</td>
<td>Text-only</td>
<td>Text and image</td>
<td>Text and image</td>
</tr>
<tr>
<td>Use</td>
<td>Online</td>
<td>Online</td>
<td>Online</td>
<td>Online</td>
<td>Online</td>
<td>Online</td>
<td>Online</td>
</tr>
<tr>
<td>Dataset</td>
<td>ARRS, ASN,</td>
<td>PubMed,</td>
<td>PubMed</td>
<td>750,000 images and figures</td>
<td>PubMed Central</td>
<td>1000 radiographs</td>
<td>PubMed, PubMed Central</td>
</tr>
<tr>
<td>Query</td>
<td>Text</td>
<td>Text</td>
<td>Text</td>
<td>Text</td>
<td>Text</td>
<td>Text and image</td>
<td>Text and image</td>
</tr>
<tr>
<td>Feedback</td>
<td>Not addressed</td>
<td>Not addressed</td>
<td>Not addressed</td>
<td>Not addressed</td>
<td>Not addressed</td>
<td>Image</td>
<td>Image</td>
</tr>
</tbody>
</table>