Salient-Object-Based Image Query by Visual Content

By

DAWIT BULCHA AMENU

Thesis

Submitted to the School of Graduate Studies of Addis Ababa University in

partial fulfillment of the requirements for the Degree of Master of Science

in Computer Science

July 2004

Addis Ababa University School of Graduate Studies

Department of Computer Science

Salient-Object-Based Image Query by Visual Content

By

DAWIT BULCHA AMENU

Name and Signature of Members of the Examining Board

____

____

____

____

To the memory of my grandfather

Amenu Gudina Waji

for his great inspirations in my early ages

Acknowledgements

First of all, I would like to forward my deepest appreciation and thanks to my advisor Dr. Solomon Atnafu for his motivative and constructive guidance throughout the work. Many thanks and appreciations go to him for the discussions with him always made me think that things are possible. His enthusiasm and encouragement has always inspired me to accelerate to the completion of the work.

I would also like to thank all my instructors at the Department of Computer Science, for their personal commitment and contribution to the success of the graduate program.

My thanks go to my friend and mate Mitiku Mamuye for he was tireless for the many times printing of the material and standing by me in all the difficult moments.

I also would like to extend my thanks to my peers Fekade Getahun and Seifu Geleta working on related thesis topics. The many times discussions and sharing of ideas and resources with them had a significant contribution to the success of this work.

Finally, I am grateful to my father, my first teacher; my grandmother, who took the role of a mother and a grandmother to raise me; and my uncles and aunts and all the rest of my families, friends, and peers who, one or the other way brought me up to a success in my academic endeavor.

Table of Contents

List of Tables

List of Figures

Abstract

Salient-Object-Based Image Query By Visual Content

Dawit Bulcha

Advisor: Solomon Atnafu (PhD)

July 2004

The rise in the intense utilization of images in our daily life resulted in a high volume of images produced from different sectors of human endeavor. This resulted in the need for an efficient management of image data. Recently, Content-based image retrieval has attracted much attention from the research community. As exact matching is not possible with image retrieval, the approach is to use similarity-based matching. Much of the works on similarity-based image retrieval use the global features (color, shape, texture, etc) of the entire image to compute similarity score between two images.

Equally important to using the entire image is the use of salient-objects; objects in an image that are of particular interest to the user, as the basis of similarity-based computation. The current works on content-based image retrieval do not address very well the issues related to salient-objects based image retrieval.

In this work, we have proposed an extension to a previous work on image database modeling and query processing. To support salient object based image retrieval, we have proposed an extension of the data repository model so that spatial features of contained salient objects are captured. Moreover, we proposed an extension to the similarity-based selection operator defined earlier so that salient object based selection operation be part of image database systems for similarity-based image retrieval. We have also proposed spatial operators that can be used to compute spatial relation between an image and contained salient objects. We have reviewed and presented refined formulations of previous works on spatial relations between objects in 2D space to compute spatial relation between salient objects.

To demonstrate the viability of salient-objects-based image retrieval, we have extended a previous work named EMIMS, to develop a system named EMIMS-S (Extended Medical Image Management System to support Salient objects). We have also used this prototype to experimentally show the retrieval effectiveness of salient-objects-based image queries.

Keywords: Salient-object-based image retrieval, similarity of salient-objects, image database, image data model, similarity-based algebra, spatial relation of salient-objects.

Chapter 1

Introduction

1.1. Image retrieval

Image retrieval has been a topic of active research since the 1970s. The research communities that are mainly involved in this area are from the fields of database management and computer vision [4]. The development of interest of researchers is derived by the rise in the intense utilization of images in our daily life which resulted in a high volume of images produced from different sectors of human endeavor.

Images have long been in use in the history of mankind. Expressing a real-world phenomenon with paintings and drawings is the practice of mankind since the old times. With the growth of imaging technologies, the storage and processing capabilities of computing devices and communication technologies, the use of images has grown in every sector of life. Some of the most important sectors, where images become part of information systems, as described in many literatures, include: medicine, crime prevention, architecture, fashion, Geographic Information Systems, Art galleries, Art history, and the like. A study at the University of California at Berkeley on the size of information worldwide in the year 2000 indicated that there are 410 petabytes (4.10x10^11 MB) of images from photography, 0.016 petabytes in motion pictures production and 17.2 petabytes of X-Rays are produced annually[25]. The study further emphasizes the desperate need for better understanding and better methods of image management to take full advantage of the ever-increasing supply of information. Searching for an image of particular interest in such a large collection manually is a daunting task. This growth in size of image data production and utilization indicates an increasing need for an efficient management of the images for its better utilization.

1.2. Content-based image retrieval

Traditional database management systems mainly deal with the storage and processing of alphanumeric data. These database systems were geared mainly towards business applications where data are mainly simple types. These database systems effectively address the common database issues of data integrity, transaction processing, concurrency and the like [15]. Relational database management systems are well matured and developed technology to effectively address the requirements of storage and processing of alphanumeric types of data [14].

The traditional approach in frequent use for image retrieval is to annotate the image with keywords and then use keyword-based DBMSs to perform the retrieval [4]. This involves describing the images with textual information such as date, producer of the image, device used, etc. and some semantic information on the image depending on the domain of application. An example of such semantic information in a medical application is the diagnostic description of X-Ray, CT, MRI, etc images. There are two basic problems in this approach. The first is that manual annotation is infeasible for large collection of images. The other is that as images are rich in information, a lot of subjectivity will be introduced in the process of annotation as a result of difference in human-perception. The report in [8] describes the richness of images as follows:

> ”*… unlike books, images make no attempt to tell us what they are about and that often may be used for purposes not anticipated by their originators. Images are rich in information and can be used by researchers from a broad range of disciplines …*”

The traditional approach is a heavy burden on the users and still inefficient as it is impossible to completely describe the content such as its color, shape, texture, and regions in the image.

As a result, retrieval of images from an image database requires techniques for processing image query based on these low level image features – a technique known in the literature as Content-Based Image Retrieval (CBIR). There are a lot of ongoing researches on CBIR but it is still not in its stage of maturity and its contemporary scale of commercial use is not significant [4]. The richness in content of image poses a new challenge in its management not addressed in the traditional database systems. A typical CBIR involves two processes: the extraction of the low level image features (color, texture, shape, …etc) and the management and processing of these features for use with retrieval.

A major distinction between content-based image retrieval and alphanumeric information retrieval is that most of the alphanumeric information retrieval is based on exact matching. In content-based image retrieval, due to the complex nature of images, exact-matching is not possible. An approach used is *similarity*. Matching images based on similarity is performed by computing the closeness of the low level features of the images.

1.3. Salient Object-based-Retrieval

In the current state-of-the-art, similarity matching is performed by considering the whole image. In this approach, global features of the whole image are used for similarity comparison between two images. A comparison that considers part(s) of images for similarity is a more natural approach to image retrieval. The approach is more effective in application domains where only part of the image is of interest. In the real world, humans usually compare parts of an object (for example, it is common to say that a child has similar eyes to that of his father). In this case if one has a database of faces, it is more meaningful to compare the images using the constituent regions of the faces than to compare the entire face. These regions of image that are of particular interest are termed as *salient objects* of the image. A tumor in a brain image and cancer in an X-Ray or CT image from the medical image domain, the image of a particular actor in a frame of a segmented video, can be considered as examples of salient objects. Image retrieval based on salient objects is the particular focus of this work.

1.4. Problem statement

The general objective of this research is to develop a data model for the management of salient objects of images and techniques for the processing of queries that involve content-based image retrieval that utilizes the salient objects, and in a way, contribute to the general theme: Content-Based Image Retrieval.

Specifically, this work addresses modeling salient objects, assessment of spatial relations of salient objects, and specification and integration of query algebra involving salient objects of images.

This thesis is outlined as follows: Chapter 2 introduces some motivations on why salient-object-based image retrieval is of interest with illustrative scenarios from the real-world. Chapter 3 discusses related works in image retrieval in general and salient-objects. Chapter 4 discusses an extended data repository model for salient objects, in chapter 5, the image query algebra and spatial operators supporting salient-objects-based retrieval are presented. Chapter 6 discusses EMIMS-S, a prototype extended from EMIMS [13] that demonstrates the use of salient-objects-based queries. Chapter 7 presents the conclusions and prospectives.

Chapter 2

Motivation and Problem Definition

2.1 Motivation

As mentioned in chapter 1, similarity-based retrieval of images is possible either using the entire image or the salient-objects in the image. The importance of salient-objects in image query is given high importance both in the database and computer vision communities [9, 13, 18, 19, 20, 21, 22, 23, 24]. The work in [19] states that matching images solely on the basis of global similarities is often a too crude approach to produce satisfactory results. It further describes that clustering of the images into perceptually salient regions-of-interest that should be assigned higher weights in similarity computations can serve as an intermediate level processing between the lower pixel level processing and higher semantic level processing. Clustering also helps in eliminating an unnecessary effect of retrieval caused by the background of the image [21, 24].

The various progresses made in the development of algorithms for salient feature extraction of images clearly indicate that salient-object-based image query is an important issue to be addressed [18, 19, 20, 21, 22, 24]. These works mainly focused on the extraction of the salient features from the image. Though there are promising works in the database community, salient-object-based modeling and processing of image data was not given considerable treatment; no work has given sufficient consideration for the modeling and query algebra development that utilize salient objects. The work by S. Atnafu in [13] has laid a profound foundation to the modeling and processing of similarity-based image retrieval but did not treat the issue of salient-objects-based retrieval in depth.

In summary, most of the contemporary development of CBIR systems concentrated on the extraction of the low-level image features and similarity based retrieval based on the entire image. Though some developments were made in the modeling and processing of image data, much attention was not given to the modeling and query processing of images that make use of salient objects. This thesis focuses on modeling and processing of salient-object-based image query by visual content.

2.2 Example Scenarios

In the real world, there are many scenarios in different problem domains where retrieval of images is more important and meaningful when based on salient objects. In the following sections, we will see real-world problems that show the necessity of image retrieval using salient-objects.

1 In a medical image database, it is of interest for physicians to study malfunctioning human organs based on certain infected parts. Following is an example:

> *Given a brain image of a patient with a tumor, a physician might be interested to search for other brain images with similar anomalies in the past.*

This would enable the physician to get feedback from the medical history of past patients with similar problems.

2 In crime prevention, a police officer investigating a crime case might be interested in the following:

> *Search in a face database of criminals, for images having a certain facial feature (salient object), similar to that of a suspect under investigation.*

The facial feature here could be some special mark on the face of the suspect or a common feature such as the geometry of the nose or the shape of the mouth.

3 In Art history, the study of works of art such as paintings, sculpture and architecture is of interest to researchers, students, and the public in general. Their history, construction and meaning as cultural products are important. Image databases are used as visual substitutes that approximate the art works as closely as possible. Management of such database is of prime importance like any other image database. In such a database a researcher might for example be interested in a query of the form:

> *Retrieve images/paintings with similar constituent features to a given sample image.*

Such retrievals can be more useful when the requester has only part of the historical image due to a damage of the painting for several reasons.

As mentioned above, the application of image retrieval by using salient objects is diverse and provides the end user with systems that are more natural and intuitive to use. Thus, research in this area will result in applications of important practical implications.

As indicated in Figure 2-1 below, the image data can be represented constituting the image and its salient objects. The figure shows the RGB color distribution (color histogram of the salient objects). A database that supports salient-object-based retrieval should capture and store the features of the salient object in addition to the main image and its features. It is not important to store the salient object separately, since in the real world, the salient object is part of the image, and usually is not needed as a separate entity.

In addition to the feature vector, textual description for both the main images and the salient objects is also important. This is because of the fact that content-based and keyword/text-based retrieval can be used in a complementary way to develop a more efficient multi-criteria query.

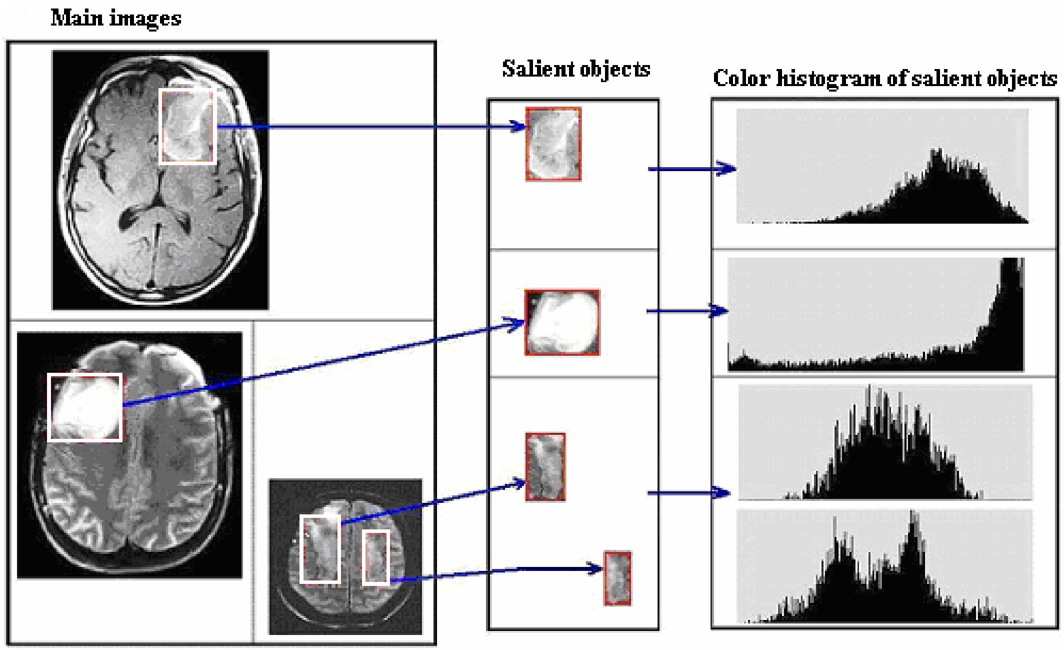


Figure 2-1 Salient objects extraction from images and color histogram representation.

In addition to the feature vector and textual representation, the spatial relation of salient-objects is important. This is needed as some retrievals require taking the spatial position of the salient object into consideration. Therefore, a data repository incorporating salient-objects should be able to capture this information.

In sections that follow, we will see examples of retrievals using salient objects that are applicable in the domain of medical applications. In all the examples, we assume that there is a historical database with a collection of brain images of patients.

Query 1:

Find all brain images that contain a tumor similar to a tumor in a given brain image.

In this scenario, the user provides an image to the system in a similar way to the one given in Figure 2-2 and indicates the region of interest. In the case of medical image, the salient object of interest is usually the anomalous part. Then, the system performs similarity computation using the feature of the query salient object and the features of the salient objects of the images in the image database. The result of the query will be images having similar salient object.

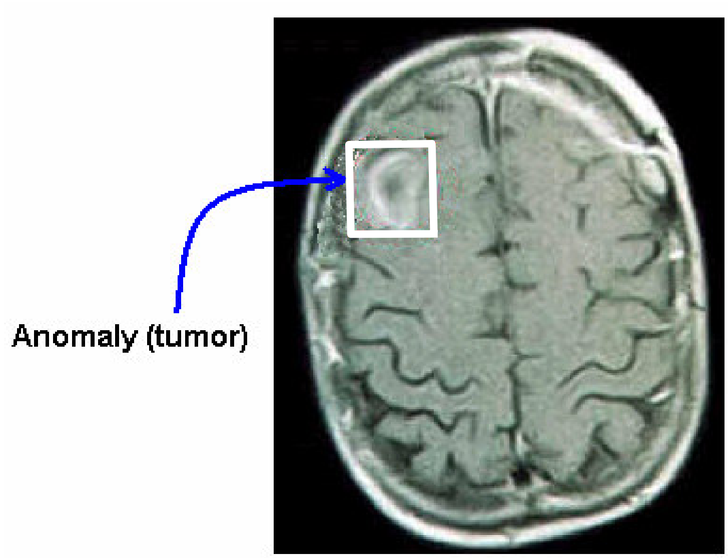


Figure 2-2 Example query mage with one salient object

Query 2:

Find all brain images that contain a similar tumor, located at the same position as that of a sample image.

Considering the image in Figure 2-2, in this scenario, the request is to find images with an anomaly (salient object) located at the *top left* part and similar to the given anomaly. Therefore, here, in addition to the similarity of the salient object, the spatial position is also important.

Query 3:

Find all images with two anomalies (salient objects) as in the query image, where one is located to the left of the other.

As indicated in the example query image in Figure 2-3 below, this query involves both the existence of salient objects and the directional relation between the two salient objects. Therefore, this requires the retrieval to consider similarity of the salient objects as well as the relative spatial position of the salient objects.

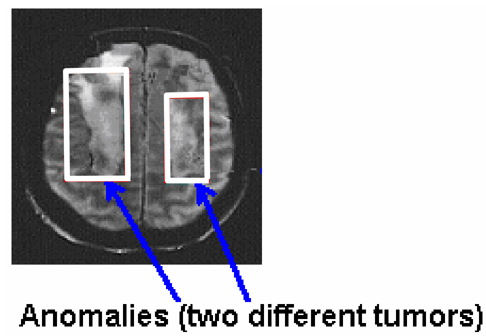


Figure 2-3 Example query mage with two salient objects

Query 4:

Find brain mages of patients between 25-30 years of age, diagnosed in the last six months with a tumor at the top left position, similar to that of a sample image.

This query requires all the three types of information in the retrieval: Salient object similarity (*tumor*), alphanumeric information (*between 25-30 years of age*, and *last six months*), and spatial position of the salient object, (*top left)*.

Query 5:

Find all brain images that have a tumor with the same size as that of a tumor in a query image.

In this query, the important consideration is the size of the salient object. In such types of queries, the comparison may not be exact, specially if the salient object is manually specified by the user. Therefore, it is worth considering the closeness of the sizes by having some mechanism of specifying some threshold.

2.3 Summary

As described in the query scenarios and the previous sections, image retrieval with salient-objects is more intuitive and relates to real world similarity-based comparison. In addition, the scenarios discussed above show that not only the content but also the traditional keyword-based/textual description of images/salient objects are also important, indicating that the two are complementary. Another important characterization of salient objects is their spatial position, which is also important in most real-world applications. As in the examples, in medical applications, the location where a tumor, or a cancer appears is so important for the physician to perform comparative analysis of the anomaly with past patient history of similar problems.

Most of the existing image data management systems focus on retrievals that utilize the global features of an image: content and textual. They do not give due consideration to the characteristics of salient-objects and their implications on retrieval.

The proceeding chapters 4 and 5 focus on the data repository modeling and query algebra that integrate salient objects in such a way that they can be used as an additional intermediate level image retrieval processing.

Chapter 3

Related Work

3.1 Digital Image representation

As mentioned in chapter 1, an image is a complex object rich in content. As a result, its representation is also complex unlike traditional data. The output of most sensors is a continuous voltage waveform whose amplitude and spatial behavior are related to the physical phenomenon being sensed. To create a digital image, we need to convert continuous sensed data into digital form. This involves two processes: sampling and quantization.

An image captured by a sensor is expressed as a continuous function f(x,y) of two co-ordinates in the plane, and also in amplitude [37]. To convert such image to digital form, we have to sample the continuous image in both coordinates and the amplitude. Digitizing the coordinate values involves just the pixel coordinates and is called sampling. Digitizing the amplitude, which is the gray level, is called quantization. Quantization involves the conversion of continuous gray-level (amplitude) into discrete quantities.

An image can be represented by an MxN matrix as shown below. Each point in the matrix is a sample point.

$$f(x,y) = \begin{pmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & & & \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{pmatrix}$$

f is a function that assigns a gray-level value to each distinct coordinate (quantization)

The number of bits required to store a digital image is: b = **M**x**N**x**k.** Where k is an integer such that 2^k is the number of gray-levels. Such an image is called a “k-bit” image. Therefore, an image with 256 possible gray level values is called an 8-bit image [37]. From the representation indicated, it is clear that, image data has huge storage requirement.

3.2 Content-Based Image Retrieval techniques and systems

There have been a lot of research works conducted on image retrieval in the past few decades, especially in the 1990s and later [4, 8]. Content-based retrieval using the visual content of images (color, shape, texture) have been studied by the computer vision community to alleviate the problems of manual image annotations. Related issues such as multidimensional-indexing, image data modeling, image query processing has been studied by the database community [13, 8, 16, 4].

Research in the image feature extraction focused mainly on how to extract the low level image features (color, shape, texture) for efficient content-based retrieval. This includes models for the representation of color (color spaces). Examples are the RGB and HSV color spaces. Each of these analysis techniques determines how color features are extracted from the image and represented mathematically for use in CBIR. Different techniques were developed for shape and texture representation in the literature [4, 8, 9].

Multidimensional indexing techniques have been studied since the middle of the 1970s. As image data have complex feature that can not be described with the traditional single dimensional data structures such as B-trees, such indexing technique are important. As a result, data structures such as R-trees, R*-trees, X-trees, SS-trees, TV-trees, SR-trees and others were developed. Some of these are variants of others optimized for efficiency of storage, query types supported, simpler data structures, etc. Detailed review of the different multidimensional indexing structures can be found in [16].

Currently, there are several CBIR systems that have been developed and in use. Most of these are research prototypes while few were converted to commercial systems. Most of these systems use low-level features such as color, texture, and shape. Systems such as QBIC (Query By Image Content) of IBM [26], Photobook of MIT [27], the VIR image search engine of Virage Inc, MARS (Multimedia Analysis and Retrieval System) of the Dept. of Computer Science of the university of Illinois at Urbana-Champaign, Surfimage of the research group at INRIA Rocquencourt, France, CBVQ (Content-Based Visual Query) of Image and Advanced Television Lab., are some [13,17]. As stated in the study by S. Atnafu [13], many efforts are being made to realize effective CBIR techniques, and each has made some contribution, but most of these works concentrated on retrieval using the entire image.

As mentioned in chapter 1, image retrieval involves similarity-based matching. Given a query image, it is possible to search its similar images from a set of images using the techniques of image analysis and processing developed in the field of computer vision [13]. Such retrieval technique has been a topic of research by many researchers [2, 3, 7, 13]. Two approaches are used in this regard. The first one is retrieval by similarity threshold, where all images within a predetermined similarity value (say ε) are retrieved, a technique known as *range query*. The other is the retrieval of the k most similar images (k Nearest Neighbors: *k-NN*) to a given query image. Many promising developments were made in these areas [4, 7, 8, 13, 16]. As the traditional DBMSs do not address the issue of similarity, a new technique is needed to deal with the problem.

Most of the researches in image retrieval mainly concentrated on image feature extraction, multidimensional indexing, and similarity matching using the low level features. A significant and pioneering work which is used as a framework in this thesis is the work by S. Atnafu [13]. This work proposed a generic and practical framework for image data management that can be effectively implemented in an object-relational environment.

In summary, most of the contemporary development of CBIR systems concentrated on the extraction of the low level image features and similarity-based retrieval using the features of the entire image. Though some works were done in the modeling and processing of image data as mentioned above, yet much attention was not given to the modeling and query processing of images that integrate salient objects. In fact, salient object-based query of images is more natural and closer to the human characterization of similarity of images.

3.3 Salient Object based image queries

The works in computer vision deal with the segmentation/clustering of an image into semantically meaningful categories that are perceptually closer to segmentation by human. These works concentrated on developing reliable clustering algorithms. Some of the clustering algorithms referred in the literature include the *K-means*, *Hierarchical clustering*, *parametric density estimation*, and *Non-parametric density estimation* [19]. These algorithms make use of some mathematical and statistical techniques to partition an image into visually meaningful parts.

As a way to segmentation, some researchers have developed algorithms for the elimination of the background so that the Figures or objects of interest are left out [24]. This is important since in most queries that are based on salient objects, the background of the images is of no use. This facilitates an image query whose purpose is searching for images containing a specific object of interest by avoiding irrelevant results that might be obtained due to the inadvertent similarity contribution of the background.

The various progresses made in the development of algorithms for salient feature extraction of images clearly indicate that salient-objects based image query is an important issue to be addressed [18, 19, 20, 21, 22, 24].

The works in the community of computer vision discussed earlier do not address the problem of the modeling and representation of the features for efficient query processing in a database context. There are some works on developing a data model for salient-objects of images and their usage in image retrieval in a database context. The DISIMA project is one that uses object-oriented approach for modeling of images and their salient objects. The model is based on the MOQL (Multimedia Object Query Language) which is an extension of the OQL (Object Query Language) proposed by the ODMG (Object Data Management Group). The DISIMA approach models an image using two blocks: the image block and the salient-object block. It views the content of an image as a set of salient objects with certain spatial relationship to each other [9]. The DISIMA approach requires a priori type definition and classification according to the application domain.

The work in [23] proposes a four level architecture for a system named Content-based Retrieval Engine (CORE) for a multimedia information system. The image level, which is the lowest level, the segmented image level which is the second level, the description and measures level, and the highest level which is the interpretation level. In this model, the segmented image level is the layer of salient-objects. This work has made a significant concept development on how to approach image data modeling but does not give particular focus to the physical, spatial, and semantic modeling of salient objects.

The work by S. Atnafu [13] intensified the importance of salient-object-based image retrieval and proposed further development. This work has proposed a possible extension of their data repository model for capturing salient objects. In addition to the image data repository model, S. Atnafu [13] have developed similarity-based algebra and related query optimization techniques. This is a major work that has formalized image data modeling and query processing in the context of a database system. The model is suitable for an implementation in the context of the evolving Object-Relational Database Management System. The model can also be extended to other types of multimedia data such as audio and video. Though this work laid a foundation for salient-object data repository, it does not treat the spatial relation of salient objects in the model and does not integrate salient objects in similarity retrieval. In this thesis we use the framework developed in [13] and propose a mechanism of integrating salient-object-based image retrieval under an ORDBMS paradigm.

3.4 Spatial Relationship of Salient Objects

The work in [30] classifies queries related to spatiotemporal relationships of salient objects into four. These are: *salient object existence*, *temporal relationships*, *spatial relationships*, and *spatiotemporal relationships*. The temporal and spatiotemporal types of relationships are important for video data as they involve timing in their retrieval. We consider only salient object existence and spatial relationships as these are of interest for salient-object based image queries.

1. Salient Object existence
   In these types of queries, users are only interested in the appearance of an object.

2. Spatial relationships
   In these queries, users express simple directional or topological relationships among salient objects. Directional relations are generally determined on the basis of the order in space between objects such as: right, left, north, south, etc. Topological relations describe the neighborhood and incidence between objects such as: disjoint, touch, overlap, etc. An example in a medical application is where a physician requests to retrieve lung x-rays in which a tumor is visible at the *top of the left lung*. Here, in

addition to the existence of the salient object (a tumor), the spatial position (*top of the left lung*) is also important.

The first types of queries do not require consideration of spatial relationships, it suffices to check the existence of the salient object of the type requested, whereas in the second types, we need a detailed analysis of the spatial relationship between the salient objects and/or the salient object and the image.

A detailed analysis of spatial relationships is important for the purpose of modeling the representation of spatial behavior of salient objects. Transitively, a data model determines the ways queries on a database can be performed.

Directional and topological relationships are the most extensively studied relations between objects [30]. In the sections that follow, we will make detailed analysis of the use of these relations in image retrievals involving salient objects.

3.4.1 Topological relations

Topological relations between contiguous objects without holes are defined by the nine-intersection model [31, 32]. According to this model, each object **p** is represented in 2D space as a point set which has an interior, a boundary, and an exterior. The topological relation between any two objects **p** and **q** is then described by the nine intersections of **p**'s interior, boundary, and exterior, with the interior, boundary, and exterior of **q**. Out of the 512 (=2^9) different relations that can be distinguished, only eight are meaningful for region objects. These are: *disjoint*, *meet*, *equal*, *overlap*, *contains*, *inside*, *covers,* and *covered_by*, these are shown in Figure 3-1 below.

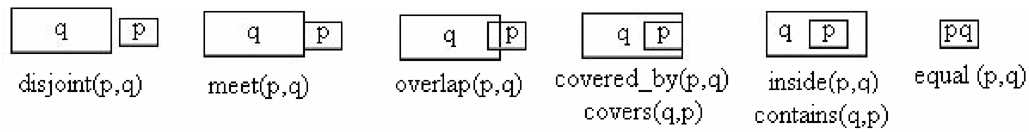


Figure 3-1 Topological Relations [31]

Tests have demonstrated that this model is able to define cognitively meaningful relations. Due to this, it has been implemented in Geographic Information Systems, and some commercial systems like Intergraph and Oracle MD [28].

3.4.1.1 Object Approximations and Topological Relations

Objects of the real world are usually irregular in shape. As a result, they are approximated with some regular geometric objects in order to facilitate query processing and approximate their spatial relations. Several approximations are proposed in the literature to represent these complex real world objects. Such approximations include: Minimum bounding rectangle(MBR) also called Minimum bounding box (MBB), Rotated minimum bounding box (RMBB), Minimum bounding circle (MBC), Minimum bounding ellipse (MBE), Convex hull (CH), and Minimum bounding n-corner (n-C) [31, 33]. A common problem with most of the approximation mechanisms is that the relationship between object approximations does not always result in the same relationship between the actual objects. The result is that there are always false hits in retrievals [31, 33]. Nevertheless, these approximations are used as filters for further analysis of the relationship between the query object and the candidate object, which is called a refinement step. This refinement step involves the use of complex algorithms from the field of computational geometry.

Though approximations can be performed using several geometries, there are some trade-offs in selecting one, such as the storage space required, simplicity of the approximation, and number of false hits in refining the candidate objects. In this thesis we have selected to use the

MBR approximation due to its simplicity, lower storage requirement, and popularity in usage. The work in [31], describes that MBRs have been used extensively to approximate objects in spatial data structures and reasoning, because they need only two points for their representation.

An object **q** can be represented as an ordered pair (**q**'$_l$, **q**'$_u$) of points corresponding to the lower left and upper right corner of the MBR **q**' that covers **q** (q'$_l$ stands for the lower and q'$_u$ for the upper point of the MBR) [31]. The topological relations we consider, therefore, are between the MBRs, and are used to approximate the relations between the actual objects.

We refer the object to be located as the *primary object* and the object in relation to which the primary object is to be located as the *reference object*. The reference object is fixed in position in the 1D space and we analyze the relationship by varying the position of the primary object. In table 1 below, the MBRs of the reference object are identified as gray and that of the primary object as white.

In [31], it is indicated that the number of pairwise disjoint relations between objects in 1D space is 13 as shown in Figure 3-2. The symbols **q**'$_l$ and **q**'$_u$ denote the edge points (lower and upper) for the reference object and the characters l and u the lower and upper points of the primary object.

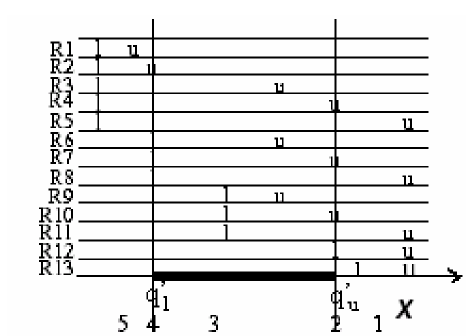


Figure 3-2 The 13 possible relations in 1D space [31]

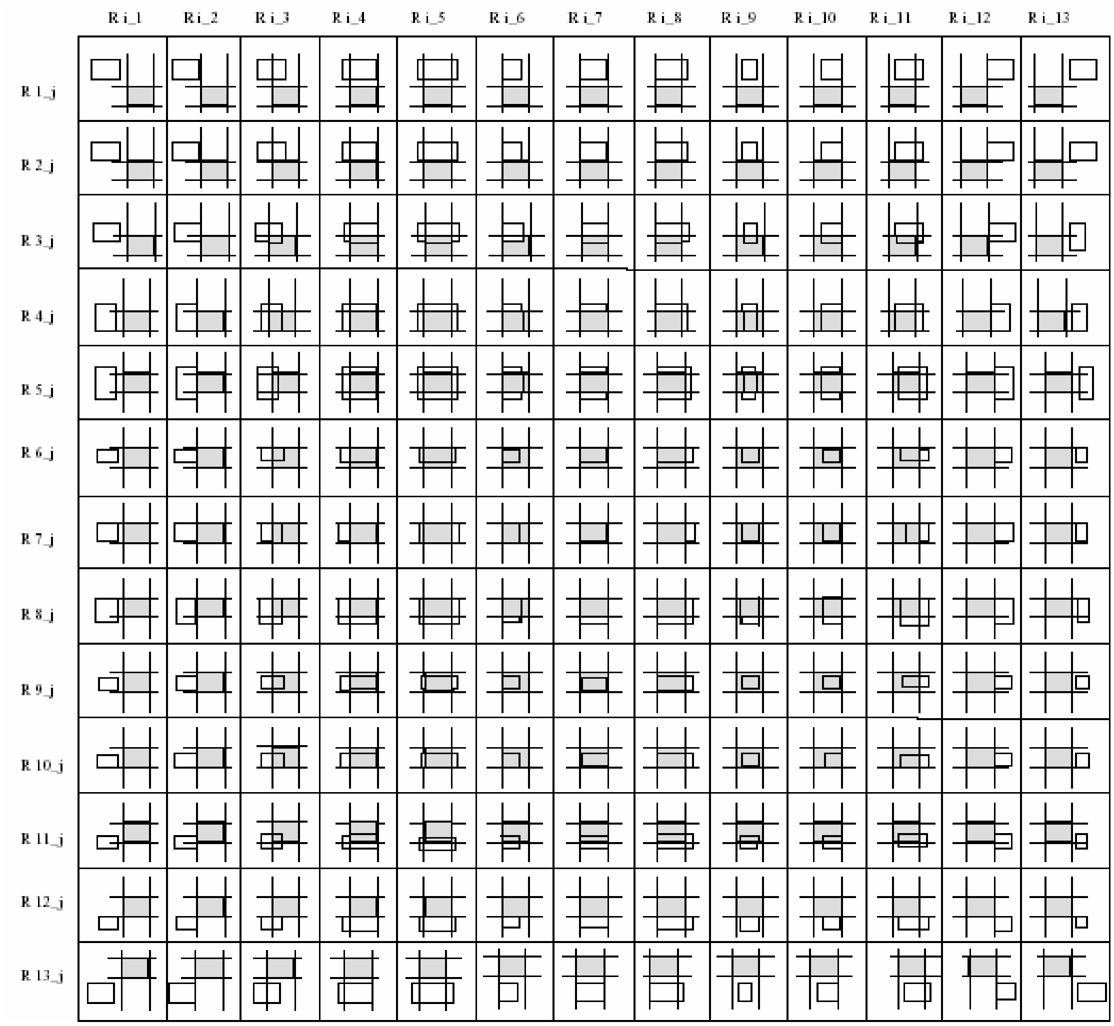


Table 1 Possible relations between MBRs [31]

The 13 relations in Figure 3-2 correspond to the time interval relations introduced by Allen [31]. The number of pair wise disjoint relations in a 2D space is 169. This is because in a 2D space, what we have is the 13 1D relations squared, resulting in 169 possible relations. The 169 possible relations are indicated in Table 1 above.

When summarized, these 169 possible relations correspond to one or more of the eight topological relations indicated in Table 2 below. As indicated in the table, the frequencies of the relations differ significantly, indicating that the chances of occurrence of some of the relations are lower and of the others are relatively higher. Therefore, an algorithm that computes a topological relation between two MBRs can consider the frequency of the relations to optimize the computation. In this regard, Clementi et al. [32] studied algorithms for minimizing these computations by exploiting the semantics of the spatial relations.

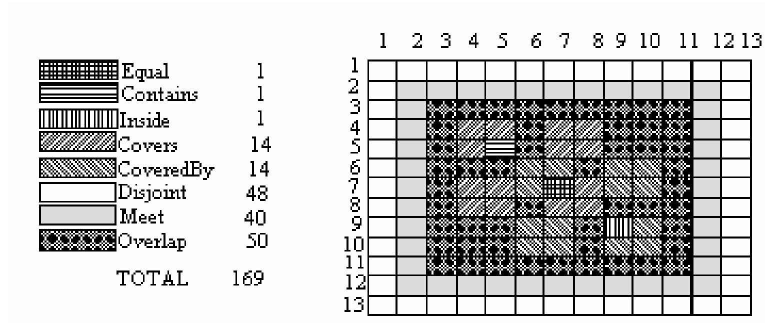
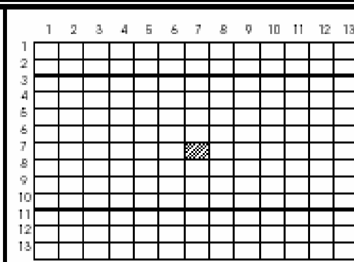
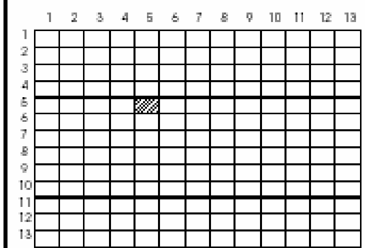
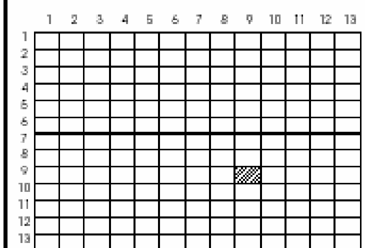
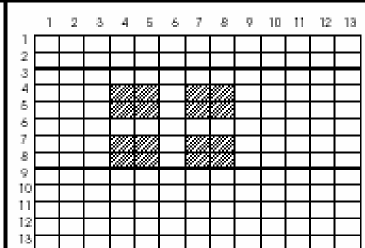


Table 2 Topological relations between MBRs [31]

3.4.1.2 Topological relations conveyed by MBRs about the actual objects

As noted earlier, topological relations between the MBRs may not necessarily convey the topological relations between the actual objects. An example is the query “find all objects p equal to q”, in this case, we need to retrieve all MBRs that are equal to the MBR of the reference (query) object. But the relation between the actual query object and the objects in the retrieved MBRs could be any of: *equal*, *overlap*, *covered_by*, or *covers* [32]. As a result, a refinement step is needed to further analyze the relation between the actual objects using Computational Geometry techniques [31]. In this thesis, we will deal only with the relations approximated by the MBRs.

Implementation of the topological relations is shown in Table 3 below. Most of the relations require a refinement step except in some cases of *disjoint* and *overlap* relations as indicated in Table 4. In these cases, it is certain that the approximated relations are the same as the relations between the actual objects.

Relation between actual objects p and q	Projection relations that the MBRs p' to be retrieved should satisfy with respect to MBR q'	Illustration of the corresponding projections
equal(p,q)	R_{7_7}(p',q')	
contains(p,q)	R_{5_5}(p',q')	
inside(p,q)	R_{9_9}(p',q')	
covers(p,q)	R_{i_j}(p',q') where i,j in {4,5,7,8}	

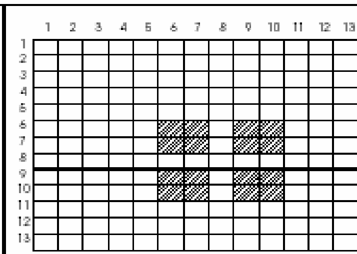
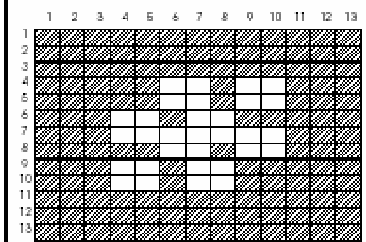
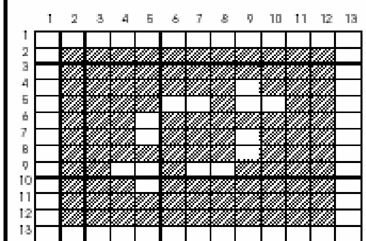
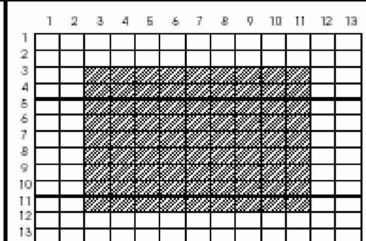
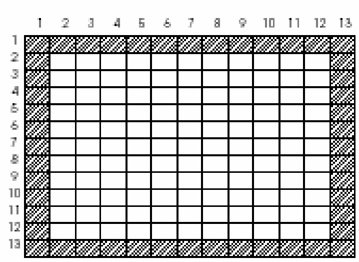
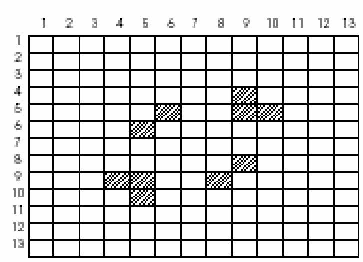
covered_by(p,q)	$R_{i,j}(p',q')$ where i,j in {6,7,9,10}	
disjoint(p,q)	any relation except $R_{i,j}(p',q')$ where i in {4,5,7,8}, j in {6,7,9,10} or i in {6,7,9,10}, j in {4,5,7,8}	
meet(p,q)	any relation except $R_{i,j}(p',q')$ where i in {1,13} or j in {1,13} or (i,j) in {(4,9), (5,6), (5,7), (5,9), (5,10), (6,5), (7,5), (7,9), (8,9), (9,4), (9,5), (9,7), (9,8), (10,5)}	
overlap(p,q)	any relation except $R_{i,j}(p',q')$ where i in {1,2,12,13} or j in {1,2,12,13}	

Table 3 Topological relations implemented [31]



disjoint



overlap

Table 4 Configurations for which a refinement step is not needed [31]

3.4.2 Directional Relations

Directional relations between two spatial objects describe such relations as *north*, *south*, *above below* and etc. Li et al [29] classified directional relations into three categories, a total of twelve relations as follows:

- Strict directional relations: *north*, *south*, *west*, and *east*
- Mixed directional relations: *northeast*, *northwest*, *southeast*, *southwest*
- Positional directional relations: *above*, *below*, *left*, and *right*

Relation	Meaning	Relation Definition
A ST B	South	A_x {d, di, s, si, f, fi, e} $B_x \wedge$ Ay {b,m} B_y
A NT B	North	A_x {d, di, s, si, f, fi, e} $B_x \wedge$ Ay {bi, mi} B_y
A WT B	West	A_x {b,m} $B_x \wedge A_y$ {d, di, s, si, f, fi, e} B_y
A ET B	East	A_x {bi, mi} $B_x \wedge A_y$ {d, di, s, si, f, fi, e} B_y
A NW B	Northwest	(A_x {b,m} $B_x \wedge A_y$ {bi, mi, oi} B_y) $\vee$ (A_x {o} $B_x \wedge A_y$ {bi, mi} B_y)
A NE B	Northeast	(A_x {bi, mi} $B_x \wedge A_y$ {bi, mi, oi} B_y) $\vee$ (A_x {oi} $B_x \wedge A_y$ {bi, mi} B_y)
A SW B	Southwest	(A_x {b,m} $B_x \wedge A_y$ {b,m, o} B_y) $\vee$ (A_x {o} $B_x \wedge A_y$ {b,m} B_y)
A SE B	Southeast	(A_x {b,m} $B_x \wedge A_y$ {b,m, o} B_y) $\vee$ (A_x {oi} $B_x \wedge A_y$ {b,m} B_y)
A LT B	Left	A_x {b,m} B_x
A RT B	Right	A_x {bi, mi} B_x
A BL B	Below	A_y {b,m} B_y
A AB B	Above	A_y {bi, mi} B_y
A EQ B	Equal	A_x {e} $B_x \wedge A_y$ {e} B_y
A IN B	Inside	A_x {d} $B_x \wedge A_y$ {d} B_y
A CV B	Cover	(A_x {di} $B_x \wedge$ Ay {fi, si, e} B_y) $\vee$ (A_x {e} $B_x \wedge A_y$ {di, fi, si} B_y) $\vee$ (A_x {fi, si} $B_x \wedge A_y$ {di, fi, si, e} A_y)
A OL B	Overlap	A_x {d, di, s, si, f, fi, o, oi, e} $B_x \wedge A_y$ {d, di, s, si, f, fi, o, oi, e} B_y
A EC B	Externally Connected	(A_x {m, mi} $B_x \wedge A_y$ {d, di, s, si, f, fi, o, oi, m, mi, e} B_y) $\vee$ (A_x {d, di, s, si, f, fi, o, oi,m, mi, e} $B_x \wedge A_y$ {m, mi} B_y)
A DJ B	Disjoint	A_x {b, bi} $B_x \vee A_y$ {b, bi} B_y

Table 5 Directional and Topological Relation Definitions [29]

The interpretations of the basic temporal interval relations from which the Directional and Topological Relations are derived are indicated in Table 6 below. The concept of temporal relations here is used in application to the relation between static objects in space at a specific point in time. Therefore, the relations between the objects define a fixed relation with no regard to change in time.

Relation	Symbol	Inverse	Meaning
A *before* B	b	bi	AAA BBB
A *meets* B	m	mi	AAABBB
A *overlaps* B	o	oi	AAA BBB
A *during* B	d	di	AAA BBBBB
A *starts* B	s	si	AAA BBBBB
A *finishes* B	f	fi	AAA BBBBB
A *equal* B	e	e	AAA BBB

Table 6 Interpretations of the basic temporal interval relations [29]

Li et al. [29], as indicated in Table 5, specified a complete definition of the combined topological and directional relations between spatial objects in terms of Allen's temporal interval algebra. A and B in the Table 5 above represent arbitrary spatial objects and their projected intervals on the x and y axes are denoted as A_x, A_y , and B_x, B_y respectively. $\wedge$ and $\vee$ are the logical AND and OR operators respectively. The notation { } is used to substitute the $\vee$ operator over relations. An example is shown below.

A_x {b,m,o} B_x is equivalent to A_x b B_x $\vee$ A_x m B_x $\vee$ A_x o B_x.

As a result, we have twelve directional relations and six topological relations, a total of eighteen spatial relations. The topological relations are reduced from eight to six as two of them are the inverse each other.

- *Covered_by* is the inverse of *covers*
- *Contains* is the inverse of *inside*.

In addition to similarity comparison between salient objects of images, the spatial (topological and directional) relationships are also of interest depending on the application domain. In the medical image domain for example, it is of interest to the physician to retrieve brain images according to some location of an anomaly in the image.

In this section, we have discussed the minimum bounding rectangle approximation of objects and their usage in the evaluation of spatial relationships between objects. We will use the definitions discussed here in the modeling of salient object representation in a manner that enable us to retrieve topological relations of salient objects belonging to a given image. We will present a refined mathematical formulation of the 18 topological and directional relations in chapter 5.

Another important spatial relation is the relation between an image and contained salient objects. An example is when a user wants to know whether a salient object is at the top left of the image. The topological and directional relations are not sufficient to describe such relations. In chapter 5, we will define important relations that can be used to describe such relations.

3.5 Image Segmentation

One of the major problems and challenging area in content-based image retrieval is the semantic gap between the lower level image content such as color, texture, shape, etc. and the higher level semantic perception of humans. Humans perceive high level semantics such as “water”, “sky”, “mountains”, “sunset”, etc. The extraction and correlation of the low level features to the higher level semantic perception of humans is crucial and challenging [38].

Humans can visually perceive and identify parts of an image that stand-out from the rest of the image such as the background. The problem with this manual type of identification is the difficulty of accurately locating the object of interest. Therefore, an automatic or semi-automatic segmentation of an image into perceptually meaningful regions is crucial in salient-object based image retrieval.

Segmentation subdivides an image into its constituent regions or objects, called segments. These segments are regions of the image that are homogenous with respect to some homogeneity predicate such as color. The level to which the subdivision is carried out depends on the problem being solved. That is, segmentation should stop when the objects of interest in an application have been isolated. The accuracy of segmentation determines the eventual success or failure of computerized analysis procedures.

Manual segmentation can be performed by human specialists of the domain of application, such as a radiologist in a medical image domain. Automatic segmentation requires some software tool/engine capable of fragmenting an image into visually identifiable parts, discriminating the background. In the case of automated segmentation, the resulting categorization should be meaningful to humans. Chen et al [38] stressed this fact:

> "*Since humans are the ultimate users of most CBIR systems, it is important to obtain segmentations that can be used to organize image contents according to categories that are meaningful to humans*."

Image segmentation algorithms are generally based on one of two basic properties of intensity values: *discontinuity* and *similarity* [37]. Several techniques of segmentation use algorithms to detect three basic types of gray-level discontinuities in a digital image: points, lines, and edges.

The segmentation problem is approached by finding boundaries between regions based on discontinuities in gray levels or via the utilization of threshold values based on the distribution of pixel properties, such as color, intensity, or hue. Other techniques are based on finding regions directly. Texture segmentation is performed with similar technique as color segmentation.

A major problem in the current state is that there is no standard algorithm or tool developed that can be utilized for automatic segmentation of an image even though there are many promising researches and experiments going on that demonstrate the viability and use of image segmentation and its uses in content-based image retrieval [18, 19, 20, 21,22, 24, 28, 38]. The MPEG-7 standard [36] does not standardize such area of technical analysis, for the reason of allowing good use of the expected improvements.

3.6 Image Data Models

Data models define the structure and content of information to be stored about an entity in an abstract manner. As image data is a complex data rich in content, we need a model that serves as a framework for capturing complete and meaningful information about an image. Various developments have been made in defining a generic model that can be used to capture image data.

The work by J.K. Wu et. al. in CORE [23] emphasizes that a multimedia information system is more than a database as it requires considerations such as processing the dataset, feature measures and extraction, and assignment of meaning to the dataset. The model proposed for an image data, referred in the paper as Multimedia Object (MOB) is as follows:

$\mathbf{O_{mob} = \{U, F, M, A, O^p, S\}}$

Where:

- **U** A multimedia object (image, video, etc)

- **F** = $\{F^1, F^2, \ldots\}$ set of features derived from data.

- $\mathbf{M^i} = \{M^i_1, M^i_2, \ldots\}$ Represents the interpretation of feature F^i. These are for example in a facial image, characterizations of facial features such as eyes, nose, mouth, etc.

- **A** stands for a set of attributes or particulars of O_{mob}. As an example, trademark number, owner, and date of registration are attributes of a trademark.

- $\mathbf{O^p}$ is a set of pointer or links to super objects, sub objects, and other objects respectively, which forms object hierarchies.

- **S** represents a set of states of O_{mob} (persistent, nonpersistent, completely defined, and incomplete)

This model is used to represent complex objects consisting of sub objects and the links among them. This work further states the necessity of segmentation so that regions of interest can be identified, extracted, and analyzed.

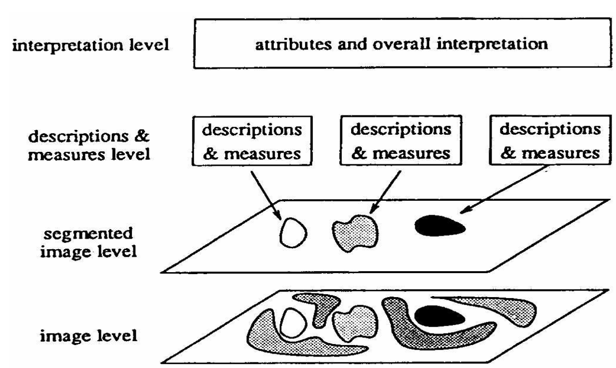


Figure 3-3 Image representation scheme of CORE [23]

As indicated in Figure 3-3 above, the CORE representation scheme provides a basic framework that can be used for the abstraction of digital images. In addition to the global image feature, this representation scheme incorporates the segmented image level where what we call salient objects naturally fit. Nevertheless, this model does not treat salient objects in a more detailed manner.

The Image Data model proposed by R. Chbeir et. al [39] provides a global view of an image. The model supports both metadata and low level descriptions of images in such a way that a multi-criteria query involving both metadata and the low-level content can be used in combination resulting in efficient image data retrieval. The model has two main spaces: the external space and the content space (Figure 3-4).

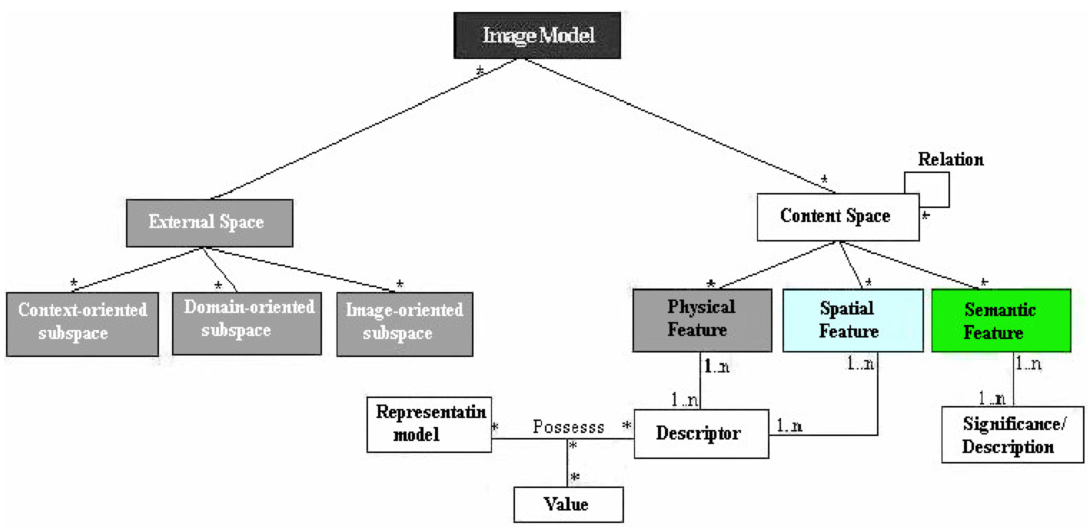


Figure 3-4 An image data model in UML by R. Chbeir et. al. [39]

The external space

The external space captures alphanumeric information associated to the image that are not related to its content. This component has three subspaces.

The context-oriented subspace: contains application-oriented data that are completely independent of the image content and have no impact on the image description. In a medical application, such information could include the hospital name, the physician identity, the patient name, patient's age, etc.

The domain-oriented subspace: consists of data that are directly or indirectly related to the image. This subspace allows one to highlight several associated issues. For example, in medical image domain, it contains information like, the medical doctor's general observations, previous associated diseases, etc. The domain-oriented subspace can also assist in identifying associated medical anomalies.

The image-oriented subspace: this subspace describes the information that is directly associated to the image creation, storage, and type. As an example, in medical domain, we need to distinguish the image compression type, the image format, creation (radiography, scanner, MRI, etc.), the incidence (sagittal, coronal, axial, etc.), the scene, the study (thoracic traumatism due to a cyclist accident), the series, image acquisition date, etc. These data can help in describing the content of the image.

The Content Space

The content space describes the content of the image and the contained salient objects. In addition to the content, it also enables description using metadata. It consists of: the physical, the spatial and the semantic features. The spatial subspace maintains relations between the salient objects, and the salient objects and the image.

The Physical Feature: describes the image (or the salient object) using its low-level features such as color, texture, etc. The color feature, for instance, can be described via several descriptors such as color Distribution, histograms, dominant color, etc. The use of such physical features allows responding to non-traditional queries. In a medical system for example, it allows to respond to queries such as: Find lung x-rays where they contain objects that are similar (by color) to a salient object S_2.

The Spatial Feature: is an intermediate (middle-level) feature that concerns geometric aspects of images (or salient objects) such as shape and position. Each spatial feature can have several representation forms such as: MBR (Minimum Bounding Rectangle), bounding circle, surface, volume, etc. The use of spatial features allows to respond to queries in medical systems such as: Find lung x-rays where an object S_1 is above object S_2 and their surfaces are disjoint.

The Semantic Feature: integrates high-level descriptions of image (or salient-objects) with the use of an application domain oriented keywords. In the medical domain, for example, terms such as name (lungs, trachea, tumor, etc.), states (inflated, exhausted, dangerous, etc.), and semantic relations (invade, attack, compress, etc.) are used to describe medical image content. Use of semantic features is important to respond to traditional queries. In medical systems, queries could be such as: Find lung x-rays where hypervascularized tumor is invading the left lung..

This model provides all the necessary descriptions of an image data, both content and metadata. The model provides a generic and complete view of an image data and can be used in defining image data repositories independent of application domains.

S. Atnafu [13] proposed an image data repository model, termed as a meta model as it is a generic model independent of any specific implementation. The model can be used to describe both alphanumeric (textual) and content information of an image. This model is developed by considering important issues on the storage and retrieval requirements of image data. It also complies with and implements the abstract image model of R. Chbier et. al. [39] described earlier.

This work proposed the following image data repository model.

*An **image data repository model** is a schema of five components*
M(id, O, F, A, P), *under an object relational model, where,*

id is a unique identifier of an instance of M

O is a reference to the image object itself that can be stored as a BLOB internally in the table or can be referenced as an external BFILE (binary file).

F is a feature vector representation of the object O. It stores the feature vectors representing all or part of the color, texture, shape, and layout contents extracted from the image.

A is an attribute component that may be used to describe the object O using textual data or keyword like annotations. This can be declared as an object, a set of objects, a table, or set of attributes linked to other relational tables allowing flexibility.

P is a data structure that is used to capture pointer links to instances of other image tables as a result of a binary similarity operation. This component holds a structure composed of the referenced table, id of the instance whose image is found to be similar, and the corresponding similarity score for each image found to be similar from the referenced table in the binary operation.

This work emphasized the importance of salient objects and the need to represent them in the model and proposed a salient object repository model as a schema of three components as follows:

S (id_s, F_s, A_s)

- **id_s** an identifier of a salient object
- **F_s** feature vector extracted to represent the low-level feature of the salient object.
- **A_s** is an attribute component that is used to capture semantic description of the salient object using textual data or keyword like annotations.

This repository model describes that the spatial relation between two salient objects of an image or an image and a salient object can be captured in the **A** component.

This model defines the representation of the salient objects but does not specify the representation of the spatial features of the salient objects. These spatial features enable retrieval using spatial relationship between the salient objects and the relationship between the salient objects and the image. The integration of spatial information into the model is very important as it results in a more efficient retrieval by restricting the result of a query by including additional restriction on the query predicate depending on the interest of the user and the application domain.

3.7 Similarity-based Image Query Algebra

Algebra is the basis of today's database management systems. One of the strengths of the relational system is its strong mathematical foundation. A query algebra is therefore an important part of a database system. With this regard, the relational system is well developed and as a result, commercial systems today provide satisfactory solution to business application requirements.

Most of the works on CBIR from computer vision and image processing concentrated on low level image feature extraction and the works in the database community concentrated on the management of alphanumeric types of data. Due to their inherent complex properties, image data can not be adequately managed under the relational systems. Therefore there is much work to be done in the formalization of a suitable algebra for the management of image data.

A major work in this direction is the work by S. Atnafu et. al [13, 39]. This work has developed and formalized similarity-based image query algebra important for the retrieval of image data under the object-relational DBMS.

This work has developed the following important operators:

- The similarity-based selection operator
- The *similarity-based join operator*
- The *Multi Similarity-Based Join operator*
- The *Symmetric Similarity-Based Join*
- The *Extract* and the *mine* operators

The similarity-Based Selection Operator

The similarity-based selection operator is a unary operator on an image table M (id, O, F, A, P) performed on the component F as defined below.

Given a query image *o* with its feature vector representation, an image table M(id, O, F, A, P), and a positive real number ε; the similarity-based selection operator, denoted by $\delta_o^{\varepsilon}(M)$, selects all the instances of M whose image objects are similar to the query image o based on the range query method.

Formally it is given as:

$$\delta_o^{\varepsilon}(M) = \{(id, o', f, a, p) \in M \,/\, o' \in R^{\varepsilon}(M, o)\}$$

where,

$R^{\varepsilon}(M, o)$ denotes the range query with respect to ε for the query image o and the set of images in the image table M.

The similarity-based selection operator operates on the feature component, F, of the image using the range query search method to select the images that are most similar to *o* from the objects in M. The result from the range query can be none or many depending on the value of ε and the feature similarity value of the query image *o* and the images in the table M.

The similarity-Based Join

Let M_1 (id_1, O_1, F_1, A_1, P_1) and M_2 (id_2,O_2, F_2, A_2, P_2) be two image tables and let ε be a positive real number. The similarity-based join operator on M_1 and M_2, denoted by $M_1 \otimes^{\varepsilon} M_2$, associates each object O_1 of M_1 to a set of similar objects in M_2 with respect to the F components of M_1 and M_2. The resulting table consists of the referring instances of M_1 (the table at the left) where P is modified by inserting a pointer pointing to the id's of the associated instances of M_2 (the table at the right side of the operation) with its corresponding similarity score.

Formally given as:

$$M_1 \otimes^{\varepsilon} M_2 = \{((id_1, o_1, f_1, a_1, p_1')/ (id_1, o_1, f_1, a_1, p_1) \in M_1 \text{ and}$$
$$p'_1 = p_1 \cup (M_2, , \{(id_2, \| o_1 – o_2 \|)\}) \text{ and } p'_1 \neq \text{Null}$$

- (id_2, o_2, f_2, a_2, p_2) $\in \delta^{\varepsilon}_{o_1}(M_2)$ (i.e., the instances of M_2 associated by the similarity-based selection $\delta^{\varepsilon}_{o_1}(M_2)$), and
- $\| o_1 – o_2 \|$ is distance between o_1 and o_2 in the feature space, also called the similarity score of o_2 and o_1 (also denoted as sim_Score (o_1, o_2)).

The similarity-based selection and similarity based join are the two basic operators developed in this work. Other operators were developed in addition to these to take advantage of some of their useful algebraic properties and query optimization benefits. These are the *Symmetric Similarity-Based Join*, the *Extract* operator, and the *Mine* operator mentioned above.

The similarity-based algebra developed in the work [13] is applied for image retrieval using the features of the entire image and it made a significant contribution to the area. Nevertheless, the work did not address how similarity-based image retrieval based on salient-objects can be integrated in the proposed system. Thus, the issue of addressing salient-object-based image retrieval is the main focus of this thesis. In the chapters that follow, we will explore how the spatial and physical features of salient-objects of an image can be utilized and integrated in the similarity-based retrieval of images.

Chapter 4

Image Data Repository Model Supporting Salient Objects

A data model is a model that describes, in an abstract way, how data is represented in an information system or database. A good data model allows capturing of sufficient and complete information about the entity to be modeled and allows better retrieval of information in the required format.

In sections that follow, we will present an elaboration of salient-objects within the generic data model of R. Chbeir et. al.[39], we then present an extension of the data repository model in [13] in a manner that spatial feature of salient objects are captured.

4.1 Image model with salient objects

Figure 4-1 below indicates the image model in [39] elaborating the placement of salient objects in the content space of the image. This presentation of the image model shows us that the content space of an image can be categorized into two sub-spaces as the features of the image as a whole (global features) and the features of each of the salient objects of interest.

The image feature (entire image): the image feature describes the physical, spatial, and semantic features of the entire image. These features describe the image as a whole without regard to constituent objects. In this description, we are referring to the aggregate features that are computed from the image considering it as a single entity. These include the physical, spatial, and semantic features as presented in [39]. Below we elaborate corresponding features for salient objects.

The Salient objects feature: The salient objects feature describes the physical, spatial, and semantic features of each salient object in the same manner as that of the image. Though a salient object is part of the image, it can be described with all of these three features:

Physical: The physical feature describes the low-level features such as color, texture, etc in a similar manner that it describes the image.

Spatial: The spatial feature of salient objects describes geometric aspects of the salient object such as shape and position with representation mechanisms used. It describes the position of the salient objects relative to the image and the position of the salient objects relative to each other.

Semantic: Describes the salient object separately with high-level description applicable in the domain of the image application. In the medical domain for example, such description could be the type of tumor in a brain image (Benign or malignant, primary or metastatic, grading or staging), the state of an anomaly (salient object), etc. These descriptions of salient objects help to integrate queries involving low-level features and keyword or semantic descriptions of images and salient objects.

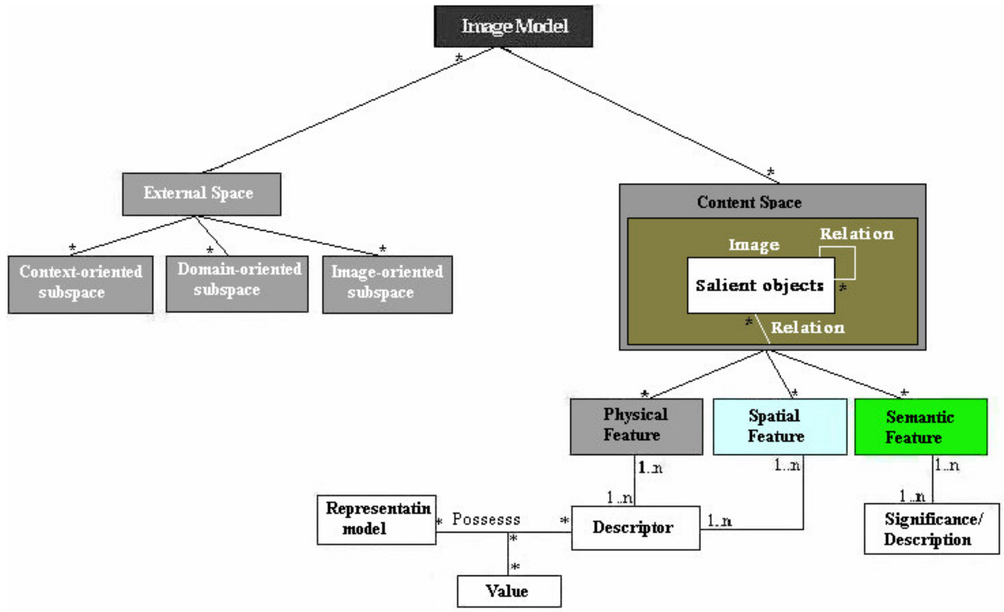


Figure 4-1 Elaboration of the salient objects within the data model of [39].

4.2 Extension of the general data repository model for salient objects

A data repository model is a conceptual model used for the storage of data. In an image database, it defines the structure and content of the image data to be stored. As described in [13], three data models are prevalent in the current database technology. These are: the popular relational model, the object-oriented model, and the Object relational model.

As described in earlier chapters, relational models are targeted towards alphanumeric types of data and do not have sufficient support for content-based image management. Nevertheless, the strength of relational model is its strong mathematical basis and maturity in the industry. Purely object-oriented model at its current state does not have rich capacity to handle complex data with complex queries as described by M. Stonebraker in [14], as a result, its success in the penetration of the current database industry is not significant. A solution for the DBMS need for image data management is the object-relational model as it combines the strengths of both the relational and object-oriented paradigm. Moreover, the OR paradigm is gaining popularity in the industry and is overshadowing the object-oriented approaches.

In the following sections, we present a repository model extended from the work in [13] in a manner that it supports the storage and retrieval of salient objects and related spatial information.

As discussed in section 3.6, in the original repository model, the **A** component of the main image captures semantic representation of the image and may be declared as object, set of objects, a table, or a set of tables linked to other relational tables. This specification makes it robust enough to extend it without violating compatibility. This flexibility allows us to extend the model so that it better supports salient objects. Moreover, though salient objects are images by themselves, the fact that they are part of the main image makes them just another characteristics (content) of the image that need characterization by themselves.

The image Data repository model discussed in section 3.6 has the following format containing five components:

M(id, O, F, A, P),

In the extension of the model, we include a required component MBR_m in the **A** component to enable us characterize an image for salient object storage and retrieval as follows:

A(MBR_m, …)

Where:

MBR_m is the minimum bounding rectangle for the main image.

The storage of the MBR for the main image helps during retrieval to characterize the spatial location of salient objects within the image. The **A** component can also contain other textual or keyword description of the image which can be specified in various forms depending on the application domain and requirement of the system under consideration. Whether the salient objects are identified manually or automated, textual or keyword information is an important high level description that is often needed in most applications.

4.3 Extension of the salient objects repository model

The salient object repository model has the following general structure:

S (id_s, F_s, A_s)

Where:

- **id_s**: The unique identifier of a salient object
- **F_s**: Feature vector representation of the salient object (as in the original image repository model)
- **A_s**: Spatial and textual/Keyword description of the salient object. This component has the following structure

To support the storage of spatial information for salient object, we extend the repository by including a required component **MBR_s** in **A_s** as follows

A_s(id, MBR_S, ...)

Where:

MBR_s is the minimum bounding rectangle for the salient object.

In addition to this required component, **A_s** stores the id of the containing image to be used as a liaison between the two. The **MBR_s** are used as the spatial descriptors of the salient objects within the scope of the main image. This will enable retrieval using the spatial position of the salient object within the main image.

Figure 4-2 below illustrates the relationship between the main image table and the salient objects table. The liaison between the main image and the salient objects can be implemented by storing the id components of the main images in each row of the corresponding salient objects table or as part of a separate object implementing $\mathbf{A_s}$.

id	**O**	**F**	**A**				**P**
id^1			MBR^1			…	
id^2			MBR^2			…	
. . .	. . .	. . .				…	. . .
id^n			MBR^n			…	

$\mathbf{id_s}$	$\mathbf{F_s}$	$\mathbf{A_s}$				
id_s^1		id^1	MBR_s^1			…
id_s^2		id^1	MBR_s^2			…
id_s^3		id^3				
id_s^4						
. . .	. . .	. . .	. . .			…
id_s^k			MBR_s^k			…

Figure 4-2 Relationship between image and salient object tables.

The two coordinates of the Minimum Bounding Rectangles identify the lower left and upper right corner or the upper left and lower right corners according to the representation scheme used.

In most application development tools, the MBR coordinates of an image are described using the left upper corner with a value of (0, 0) and right lower corner with a value of (w, h) (Figure 4-3 a.), where w and h are the width and height of the image in pixels respectively. Assuming that LU(0,0) and RL(w, h) are the coordinates of the MBR of the image and $LU_s(x_1, y_1)$ and $RL_s(x_2, y_2)$ are the coordinates of the MBR of a contained salient object, the following relation holds. This relation shows the fact that the salient objects MBRs are contained within the MBR of the image.

$$0 \le x_1 \le w,\ \ 0 \le x_2 \le w\ ,\ 0 \le y_1 \le h,\ 0 \le y_2 \le h$$

In most of the literatures dealing with spatial relations, the coordinate system used is the standard Cartesian coordinate with center (0, 0). In this case, the usual way of describing a minimum bounding rectangle is to use the lower left and upper right coordinates. To comply with the literature and have consistent definitions, we can translate the MBR coordinates of an image to the standard Cartesian coordinate system as shown in Figure 4-3 b.

With the above assumption that the MBR coordinates of an image are LU(0,0) and RL(w, h) respectively, for an arbitrary coordinate (x, y) from this region, we can translate the coordinates to the standard Cartesian coordinate (x', y') with center at the center of the image as follows (illustration Figure 4-3 b):

$$x' = x - w/2$$
$$y' = -y + h/2$$

With this representation, we can retain the coordinates obtained from image management tools, while complying with the literature from spatial relations.

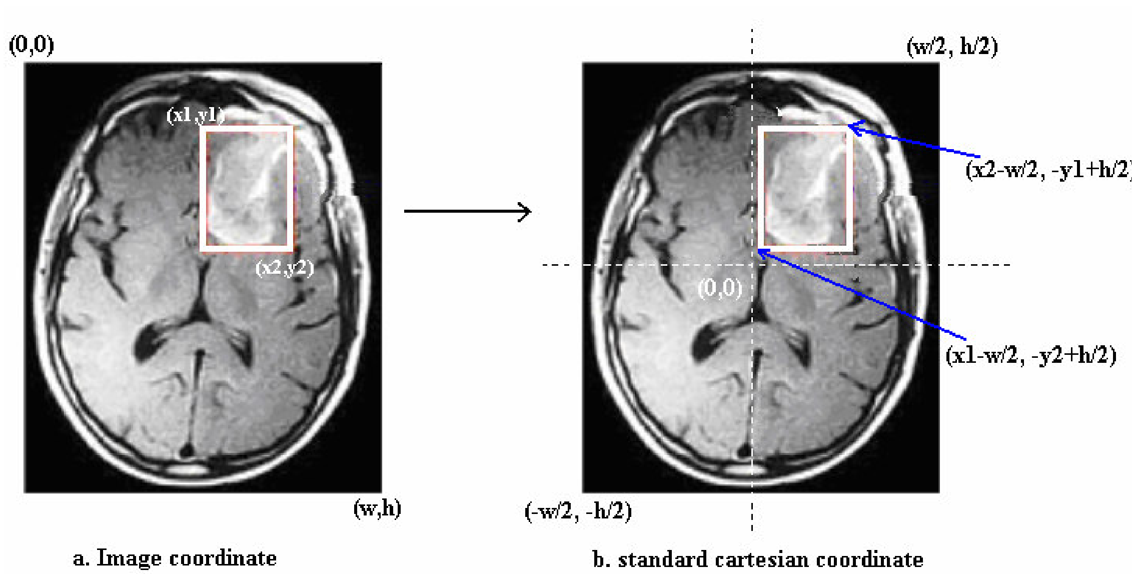


Figure 4-3 MBR representation of images and contained salient object(s)

The extended salient object repository model complies with the existing repository of S. Atnafu[13]. The inclusion of the MBR for both the image and contained salient objects helps to capture important spatial attributes. This information helps to compute positions of the salient objects in the image and spatial relations of salient objects during retrieval.

In addition to the $\mathbf{MBR_s}$, $\mathbf{A_s}$ captures all semantic descriptions of the salient object with textual/keyword description. Such description of the salient object is important in addition to the use of the MBRs. In a medical application, for example, these attribute components are useful to describe the location of observed anomalies in medical images. As examples, a physician might need to describe a tumor observed in a brain image or the characteristics of a lung cancer observed in a lung x-ray.

Chapter 5

Similarity-based Algebra for Salient Object-based image queries

In the context of CBIR, similarity is the most important notion. This is due to the fact that in a content-based image database, search is not based on exact matching, but on similarity-based matching. Therefore, it is important to have operators that can be used for matching image similarity. Though there are several developments in this area, only the works by S. Atnafu [13] has made a profound formalization and development of the notion of similarity in the context of image data management in a database environment. As has been mentioned in the previous chapters, similarity can be matched based on either the entire image using the global features or using the features of salient objects of interest, which is the main theme of this work.

In sections that follow, we will define important operators that: aid in matching image similarity using the features of the salient objects, determine the spatial position of salient objects within the image, and describe the spatial relationships of the contained salient objects.

5.1 Salient-Object-based Similarity Selection

Before defining the salient-object-based similarity selection, we re-state the similarity-based selection operator developed in [13].

The similarity-Based Selection Operator [13].

The similarity-based selection operator is a unary operator on an image table M (id, O, F, A, P) performed on the component F as defined below.

Given a query image o with its feature vector representation, an image table M(id, O, F, A, P), and a positive real number ε ; the similarity-based selection operator, denoted by $\delta_o^{\varepsilon}(M)$, selects all the instances of M whose image objects are similar to the query image o based on the range query method.

Formally it is given as:

$$\delta_o^{\varepsilon}(M) = \{(id, o', f, a, p) \in M \,/\, o' \in R^{\varepsilon}(M, o)\}$$

where ,

$R^{\varepsilon}(M, o)$ denotes the range query with respect to ε for the query image o and the set of images in the image table M.

The similarity-based selection operator operates on the feature component, F, of the image using the range query search method to select the images that are most similar to *o* from the objects in M. The result from the range query can be none or many depending on the value of ε and the feature similarity value of the query image *o* and the images in the table M.

Salient-Object-based Similarity Selection

Given the definition of the similarity-based selection operator and the range query discussed in chapter 3, we define the Salient-Object-based Similarity Selection operator as follows:

Given a query image O and its salient object O_s with its feature vector representation, an image table M (id, O, F, A, P), a salient Objects table S(id_s, F_s, A_s), and a positive real number ε; a salient-object-based similarity selection operator $\delta^{\varepsilon}_{O_s}(M)$, selects all instances of M whose image objects have salient objects similar to the salient object O_s of the query image O based on range query method.

Formally,

$$\delta^{\varepsilon}_{O_S}(M) = \left\{ \left(id, o', f, a, p \right) \in M \,/\, O' \in \prod\nolimits_{M.o} \left(\delta_{M.id \in I}(M) \right) \right\}$$

Where,

$$I = \prod\nolimits_{S.A_s.id} \left(\delta^{\varepsilon}_{O_s}(S) \right) \text{ and } \delta^{\varepsilon}_{O_s}(S) = \left\{ \left(id_s, f'_s, a_s \right) \in S \,/\, f' \in R^{\varepsilon}\left(S, f_s \right) \right\}$$

> $R^{\varepsilon}\left(S, f_s\right)$ denotes the range query with respect to ε for the salient object O_s whose feature vector is f_s and set of salient objects in the table S. Here, the feature vector f_s represents the salient object as we do not capture the salient object itself in the repository model. Hence, $\delta^{\varepsilon}_{O_s}(S)$ is a similarity-based selection operator applied to the salient objects table.

The extension of the similarity-based selection operator to salient-object-based similarity selection involves two steps; similarity-based selection on the salient objects table followed by relational selection on the main image table on condition that the salient objects of the images are retrieved with the similarity-based selection on the salient objects table.

The similarity-based selection on the salient objects table retrieves salient objects that are within the similarity threshold ε for the salient object of the query image and the salient objects table S. The next step, the relational selection on the main image table M retrieves images from the table M whose *id*s are returned from the projection over the *id* components on the result of the similarity-based selection operated on the salient objects table S.

The difference between this operator and the previous similarity-based selection operator on M is that, here, the salient objects are used for similarity computation instead of the entire image.

An important point to consider here is the situation where the retrieved image has more than one salient objects. That is, suppose that an image O_i is found to contain similar salient object to the query image salient object O_s. Suppose also that image O_i has two or more salient objects. In this case, the visualization of the resulting similar images has to provide a visual clue of which one of the salient objects is the cause for the similarity. To make this possible, we need to retrieve the salient objects together with their MBRs. Retrieving the MBRs of the salient objects enable us to visually locate the spatial position of the salient objects in the resulting images.

Another important issue is a case where the user can specify a query with more than one salient object. An example is where the user says:

> *Retrieve all images with two salient objects similar to that of the two salient objects specified for the query image as indicated on the query screen area.*

In such scenarios, the salient-object-based similarity selection can be extended by considering different query parameters such as the number of salient objects specified in the query image, the spatial relationship of the salient objects with respect to each other and with respect to the main image. Using the spatial relations discussed in chapter 3 and the refined formulations of topological and directional relations presented in section 5.2.2 below, is possible to respond to these types of queries.

5.2 Spatial Query Operators

As presented in chapter 3, many studies have been made on topological and directional relations between two objects [29, 31, 32]. These relations can be used to describe the relation between two salient objects of an image. In addition to the topological and directional relations, equally important is the relation between an image and the contained salient objects. The position of a salient object within an image is important in most applications that use content-based image database. In this section, we will classify and present spatial operators as those describing the relation between the salient objects and the image, and those describing the relation between the salient objects themselves.

In section 5.2.1, we will define spatial operators used for the computation of the relation between the image and contained salient objects. In section 5.2.2, we will present refined mathematical formulations of how the computation of the topological and directional relations of the spatial relations studied in [29, 31, 32] can be done given the MBRs of the salient objects.

The relations between an image and the contained salient objects are relations between objects that are always contained within another object. Therefore, the topological and directional relations do not suffice to describe these relations. In this regard, we need operators that can be used to state the position of a salient object approximated by the MBR relative to the main image. A problem in categorizing and defining such operators is the difficulty of identifying and naming possible partitioning of the space of an image. To simplify and resolve this problem, we have identified and defined nine operators that can be used to unambiguously describe the position of a salient object within the image using its MBR.

5.2.1 Main Image - salient object relation

As indicated in the example queries explained in chapter 2, queries can usually involve positional predicates such as *top left*, *bottom right*, *center*, and so on. In a medical application, a physician might for example be interested in brain images with a tumor at the top right part. These are scenarios that indicate the need for a scheme of computing the spatial position of a salient object within the main image.

In this work, we propose a scheme of describing the position of a salient object within the main image by partitioning the main image into four quadrants of equal size as indicated in the Figure 5-1 below.

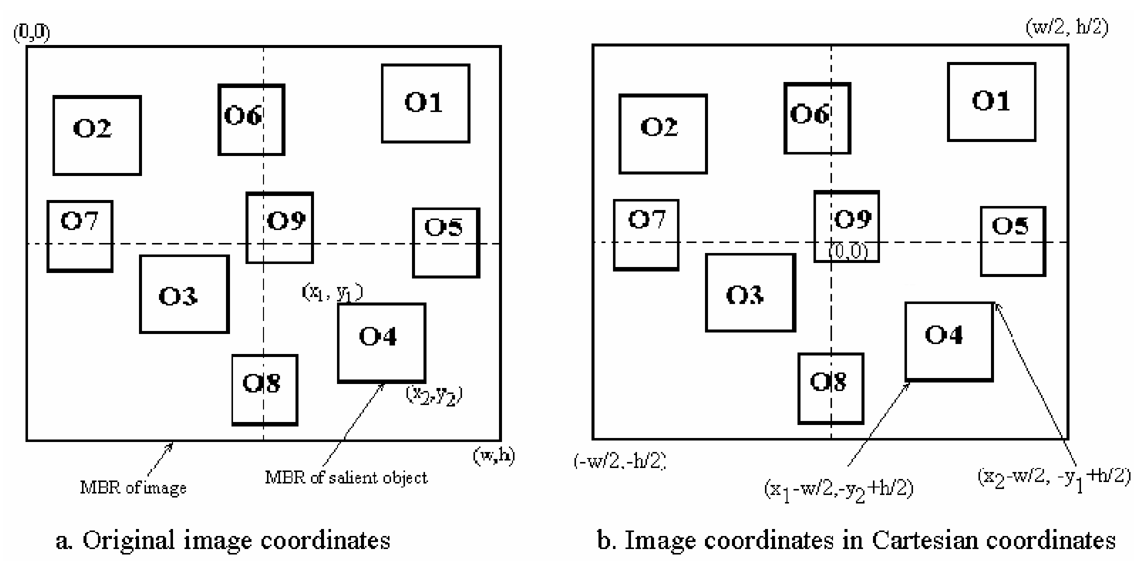


Figure 5-1 Salient objects positions within the main image

As indicated in Figure 5-1 above, we classify the position of a salient object within the image using nine positional descriptors. The coordinates in Figures a and b above show the usual coordinates used in image applications and the standardized Cartesian coordinates respectively. Table 7 below describes the corresponding nine positions.

Salient Object	Position description	Alternate description
O1	*top right*	*top right*
O2	*top left*	*top left*
O3	*bottom left*	*bottom left*
O4	*bottom right*	*bottom right*
O5	*center right*	*right*
O6	*top center*	*top*
O7	*center left*	*left*
O8	*bottom center*	*bottom*
O9	*center center*	*center*

Table 7 The nine positional description of a salient object within the main image

Assuming that {(0, 0), (w, h)} are the coordinate of the MBR of the main image and {(x_1, y_1), (x_2,y_2)} are the coordinates of the MBR of an arbitrary salient object within the image, the nine positions can be expressed mathematically as in the following table(Table 8). These descriptions hold equivalently when the coordinates are converted to the standard Cartesian coordinates.

Position description	Operator symbol	Mathematical description
top right	top_right	$w/2 \leq x_1 \wedge y_2 \leq h/2$
top left	top_left	$x_2 \leq w/2 \wedge y_2 \leq h/2$
bottom left	bottom_left	$x_2 \leq w/2 \wedge y_1 \geq h/2$
bottom right	bottom_right	$x_1 \geq w/2 \wedge y_1 \geq h/2$
right	right	$x_1 \geq w/2 \wedge (y_1 < h/2 \wedge y_2 > h/2)$
top	top	$y_2 \leq h/2 \wedge (x_1 < w/2 \wedge x_2 > w/2)$
left	left	$x_2 \leq w/2 \wedge (y_1 < h/2 \wedge y_2 > h/2)$
bottom	bottom	$y_1 \geq h/2 \wedge (x_1 < w/2 \wedge x_2 > w/2)$
center	center	$(x_1 < w/2 \wedge x_2 > w/2) \wedge (y_1 < h/2 \wedge y_2 > h/2)$

Table 8 Implementation of salient object main image relations

Once the MBRs of the image and the contained salient objects are determined and these operators are implemented, we have sufficient information and complete mechanism of responding to queries involving the positions of the salient objects as in the example queries 2 and 4 of chapter 2.

Query 2 of chapter 2 was stated as follows:

> *Find all brain images that contain a similar tumor, located at the same position as that of a sample image.*

Assuming that S_q, the salient object of the query image (the tumor), is located at the top left of the image, using the definition of Table 8 above, the query can be stated using the following SQL-like expression.

SELECT M.* FROM M, S
WHERE (M.id = $S.A_s$.id) AND ($Fv(S_q) \approx^{\varepsilon} S.Fv$)
AND *top_left* (O.A.MBR, $S.A_s$.MBR)

5.2.2 Relation between salient objects

In some of the cases, it is common to have more than one salient object within a single image. In these situations, it is of interest to describe the relationship between the salient objects themselves. Query 3 stated in chapter 2 requires retrieval of all brain images with two tumors (anomalies) where one is located at the left of the other. In other words, this requires retrieval of brain images with salient objects with the relationship *right* or *left*. As seen in chapter 3, such relationships are categorized into two as topological and directional relations.

As mentioned in chapter 3 and earlier in this chapter, in this section we will present refined mathematical formulations of how the topological and directional relations between objects defined in [29, 31, 32] can be computed from the MBRs of the salient objects. These formulations are not implemented in our prototype (EMIMS-S), but as EMIMS-S captures the necessary spatial attributes, it can be integrated in a similar way as that of the relation between the image and the salient objects.

5.2.2.1 Topological Relations

In chapter 3, we have stated eight topological relations that can be used to describe salient object position relative to each other [29, 31, 32]. These relations are: *equal*, *contains*, *inside*, *covers*, *covered by*, *overlap*, *meet*, and *disjoint*. Out of these, it suffices to define six of them as two of them can be derived from the others as follows:

- *Covered_by* is the inverse of *covers*
- *Contains* is the inverse of *inside*.

In the following, we outline the refined mathematical formulations of six of the topological relations defined in [29, 31, 32] between two objects using the MBRs.

Let A and B represent arbitrary salient objects and their projected intervals on the x and y axes denoted as *AX*, *AY*, and *BX*, *BY* respectively. $\wedge$ and $\vee$ are the logical AND and OR operators respectively. The notation { } is used to substitute the $\vee$ operator over relations. The symbols b, bi, m, mi, o, oi, d, di, s, si, f, fi, e are the basic temporal interval relations as discussed in chapter 3.

Let $\{(AX._{x1}, AY._{y1}),\ (AX._{x2}, AY._{y2})\}$ and $\{(BX._{x1},\ BY._{y1}\ ),\ (BX._{x2}, BY._{y2})\}$ be the respective coordinates of the MBRs of the objects A and B in the coordinate system. Figure 5-2 below illustrates the representation in terms of the MBRs.

Figure 5-2 MBR representation of the projection of objects in two dimensional coordinate plane

Then we can present the refined formulations of the six topological relations of [29, 31, 32] as follows:

Relation	**A** *equal* **B**
Definition	$AX\ \{e\}\ BX \wedge AY\ \{e\}\ BY$
Refined formulation	$(AX._{x1} = BX._{x1}) \wedge (AX._{x2} = BX._{x2}) \wedge (AY._{y1} = BY._{y1}) \wedge (AY._{y2} = BY._{y2})$

Relation	**A** *inside* **B**
Definition	$AX\ \{d\}\ BX \wedge AY\ \{d\}\ BY$
Refined formulation	$(AX._{x1} > BX._{x1} \wedge AX._{x2} < BX._{x2}) \wedge (AY._{y1} > BY._{y1} \wedge AY._{y2} < BY._{y2})$

Relation	**A** *cover* **B**
Definition	$(AX\{di\}BX \wedge AY\{fi, si, e\}BY) \vee (AX\{e\}BX \wedge AY\{di, fi, si\}BY) \vee$ $(AX\ \{fi,si\}BX \wedge AY\ \{di, fi, si, e\}\ BY)$
Refined formulation	$(BX._{x1} > AX._{x1} \wedge BX._{x2} < AX._{x2} \wedge ((BY._{y2} = AY._{y2} \wedge BY._{y1} > AY._{y1}) \vee$ $(BY._{y1} = AY._{y1} \wedge BY._{y2} < AY._{y2}) \vee (AY._{y1} = BY._{y1} \wedge AY._{y2} = BY._{y2}))$ $(AX._{x1} = BX._{x1} \wedge AX._{x2} = BX._{x2} \wedge ((BY._{y1} > AY._{y1} \wedge BY._{y2} < AY._{y2}) \vee$ $(BY._{y2} = AY._{y2} \wedge BY._{y1} > AY._{y1}) \vee (BY._{y1} = AY._{y1} \wedge BY._{y2} < AY._{y2}))) \vee$ $(((BX._{x2} = AX._{x2} \wedge BX._{x1} > AX._{x1}) \vee (BX._{x1} = AX._{x1} \wedge BX._{x2} < AX._{x2})) \wedge$ $((BY._{y1} > AY._{y1} \wedge BY._{y2} < AY._{y2}) \vee (BY._{y2} = AY._{y2} \wedge BY._{y1} > AY._{y1}) \vee$ $(BY._{y1} = AY._{y1} \wedge BY._{y2} < AY._{y2}) \vee (AY._{y1} = BY._{y1} \wedge AY._{y2} = BY._{y2})))$

Relation **A** *overlap* **B**

Definition AX {d, di, s, si, f, fi, o, oi, e} $BX \wedge AY$ {d, di, s, si, f, fi, o, oi, e}BY

Refined formulation

$$((AX._{x1} > BX._{x1} \wedge AX._{x2} < BX._{x2}) \vee (BX._{x1} > AX._{x1} \wedge BX._{x2} < AX._{x2}) \vee$$
$$(AX._{x1} = BX._{x1} \wedge AX._{x2} < BX._{x2}) \vee (BX._{x1} = AX._{x1} \wedge BX._{x2} < AX._{x2}) \vee$$
$$(AX._{x2} = BX._{x2} \wedge AX._{x1} > BX._{x1}) \vee (BX._{x2} = AX._{x2} \wedge BX._{x1} > AX._{x1}) \vee$$
$$(AX._{x1} < BX._{x1} \wedge BX._{x1} < AX._{x2}) \vee (BX._{x1} < AX._{x1} \wedge AX._{x1} < BX._{x2}) \vee$$
$$(AX._{x1} = BX._{x1} \wedge AX._{x2} = BX._{x2}))$$
$$\wedge$$
$$(AY._{y1} > BY._{y1} \wedge AY._{y2} < BY._{y2}) \vee (BY._{y1} > AY._{y1} \wedge BY._{y2} < AY._{y2}) \vee$$
$$(AY._{y1} = BY._{y1} \wedge AY._{y2} < BY._{y2}) \vee (BY._{y1} = AY._{y1} \wedge BY._{y2} < AY._{y2}) \vee$$
$$(AY._{y2} = BY._{y2} \wedge AY._{y1} > BY._{y1}) \vee (BY._{y2} = AY._{y2} \wedge BY._{y1} > AY._{y1}) \vee$$
$$(AY._{y1} < BY._{y1} \wedge BY._{y1} < AY._{y2}) \vee (BY._{y1} < AY._{y1} \wedge AY._{y1} < BY._{y2}) \vee$$
$$(AY._{y1} = BY._{y1} \wedge AY._{y2} = BY._{y2})$$

Relation **A** *meet* **B**

Definition (AX {m, mi} $BX \wedge AY$ {d, di, s, si, f, fi, o, oi, m, mi, e}BY) $\vee$
(AX {d, di, s, si, f, fi, o, oi, m, mi, e} $BX \wedge AY${m, mi}BY)

Refined formulation

$$($$
$$AX._{x2} = BX._{x1} \vee BX._{x2} = AX._{x1}) \wedge ((AY._{y1} > BY._{y1} \wedge AY._{y2} < BY._{y2}) \vee$$
$$(BY._{y1} > AY._{y1} \wedge BY._{y2} < AY._{y2}) \vee (AY._{y1} = BY._{y1} \wedge AY._{y2} < BY._{y2}) \vee$$
$$(BY._{y1} = AY._{y1} \wedge BY._{y2} < AY._{y2}) \vee (AY._{y2} = BY._{y2} \wedge AY._{y1} > BY._{y1}) \vee$$
$$(BY._{y2} = AY._{y2} \wedge BY._{y1} > AY._{y1}) \vee (AY._{y1} < BY._{y1} \wedge BY._{y1} < AY._{y2}) \vee$$
$$(BY._{y1} < AY._{y1} \wedge AY._{y1} < BY._{y2}) \vee (AY._{y2} = BY._{y1}) \vee (BY._{y2} = AY._{y1}) \vee$$
$$(AY._{y1} = BY._{y1} \wedge AY._{y2} = BY._{y2})$$
$$) \vee$$
$$($$
$$(AX._{x1} > BX._{x1} \wedge AX._{x2} < BX._{x2}) \vee (BX._{x1} > AX._{x1} \wedge BX._{x2} < AX._{x2}) \vee$$
$$(AX._{x1} = BX._{x1} \wedge AX._{x2} < BX._{x2}) \vee (BX._{x1} = AX._{x1} \wedge BX._{x2} < AX._{x2}) \vee$$
$$(AX._{x2} = BX._{x2} \wedge AX._{x1} > BX._{x1}) \vee (BX._{x2} = AX._{x2} \wedge BX._{x1} > AX._{x1}) \vee$$
$$(AX._{x1} < BX._{x1} \wedge BX._{x1} < AX._{x2}) \vee (BX._{x1} < AX._{x1} \wedge AX._{x1} < BX._{x2}) \vee$$
$$(AX._{x2} = BX._{x1}) \vee (BX._{x2} = AX._{x1}) \vee (AX._{x1} = BX._{x1} \wedge AX._{x2} = BX._{x2}) \wedge$$
$$(AY._{y2} = BY._{y1} \vee BY._{y2} = AY._{y1})$$
$$)$$

Relation **A** *disjoint* **B**

Definition AX {b, bi} $BX \vee AY$ {b, bi}BY

Refined formulation $AX._{x2} < BX._{x1} \vee BX._{x2} < AX._{x1} \vee AY._{y2} < BY._{y1} \vee BY._{y2} < AY._{y1}$

5.2.2.2 Directional Relations

As discussed in chapter 3, directional relations include the following: *north*, *south*, *west*, *east, northeast*, *northwest*, *southeast*, *southwest, above*, *below*, *left*, and *right* [29, 31, 32]. In the following, we present the original definitions and the refined formulations of these directional relations using similar notation used for topological relations.

Relation	**A** *south* **B**
Definition	AX {d, di, s, si, f, fi, e} $BX \wedge AY$ {b, m} BY
Refined formulation	$((AX\cdot_{x1} > BX\cdot_{x1} \wedge AX\cdot_{x2} < BX\cdot_{x2}) \vee (BX\cdot_{x1} > AX\cdot_{x1} \wedge BX\cdot_{x2} < AX\cdot_{x2}) \vee$ $(AX\cdot_{x1} = BX\cdot_{x1} \wedge AX\cdot_{x2} < BX\cdot_{x2}) \vee (BX\cdot_{x1} = AX\cdot_{x1} \wedge BX\cdot_{x2} < AX\cdot_{x2}) \vee$ $(AX\cdot_{x2} = BX\cdot_{x2} \wedge AX\cdot_{x1} > BX\cdot_{x1}) \vee (BX\cdot_{x2} = AX\cdot_{x2} \wedge BX\cdot_{x1} > AX\cdot_{x1} \vee$ $AX\cdot_{x1} = BX\cdot_{x1} \wedge AX\cdot_{x2} = BX\cdot_{x2})) \wedge \quad ((AY\cdot_{y2} < BY\cdot_{y1}) \vee (AY\cdot_{y2} = BY\cdot_{y1}))$

Relation	**A** *north* **B**
Definition	AX {d, di, s, si, f, fi, e} $BX \wedge AY$ {bi, mi} BY
Refined formulation	$(AX\cdot_{x1} > BX\cdot_{x1} \wedge AX\cdot_{x2} < BX\cdot_{x2}) \vee (BX\cdot_{x1} > AX\cdot_{x1} \wedge BX\cdot_{x2} < AX\cdot_{x2}) \vee$ $(AX\cdot_{x1} = BX\cdot_{x1} \wedge AX\cdot_{x2} < BX\cdot_{x2}) \vee (BX\cdot_{x1} = AX\cdot_{x1} \wedge BX\cdot_{x2} < AX\cdot_{x2}) \vee$ $(AX\cdot_{x2} = BX\cdot_{x2} \wedge AX\cdot_{x1} > BX\cdot_{x1}) \vee (BX\cdot_{x2} = AX\cdot_{x2} \wedge BX\cdot_{x1} > AX\cdot_{x1}) \vee$ $(AX\cdot_{x1} = BX\cdot_{x1} \wedge AX\cdot_{x2} = BX\cdot_{x2}) \wedge \quad (BY\cdot_{y2} < AY\cdot_{y1}) \vee (BY\cdot_{y2} = AY\cdot_{y1})$

Relation	**A** *west* **B**
Definition	AX {b, m} $BX \wedge AY$ { d, di, s, si, f, fi, e } BY
Refined formulation	$((AX\cdot_{x2} <= BX\cdot_{x1})) \wedge$ $((AY\cdot_{y1} = BY\cdot_{y1} \wedge AY\cdot_{y2} < BY\cdot_{y2}) \vee (BY\cdot_{y1} > AY\cdot_{y1} \wedge BY\cdot_{y2} < AY\cdot_{y2}) \vee$ $(AY\cdot_{y1} = BY\cdot_{y1} \wedge AY\cdot_{y2} < BY\cdot_{y2}) \vee (BY\cdot_{y1} = AY\cdot_{y1} \wedge BY\cdot_{y2} < AY\cdot_{y2}) \vee$ $(AY\cdot_{y2} = BY\cdot_{y2} \wedge AY\cdot_{y1} > BY\cdot_{y1}) \vee (BY\cdot_{y2} = AY\cdot_{y2} \wedge BY\cdot_{y1} > AY\cdot_{y1}) \vee$ $(AY\cdot_{y1} = BY\cdot_{y1} \wedge AY\cdot_{y2} = BY\cdot_{y2}))$

Relation	**A** *east* **B**
Definition	AX {bi, mi}BX $\wedge$ AY { d, di, s, si, f, fi, e }BY
Refined formulation	$((BX._{x2} <= AX._{x1})) \wedge$ $((AY._{y1} = BY._{y1} \wedge AY._{y2} < BY._{y2}) \vee (BY._{y1} > AY._{y1} \wedge BY._{y2} < AY._{y2}) \vee$ $(AY._{y1} = BY._{y1} \wedge AY._{y2} < BY._{y2}) \vee (BY._{y1} = AY._{y1} \wedge BY._{y2} < AY._{y2}) \vee$ $(AY._{y2} = BY._{y2} \wedge AY._{y1} > BY._{y1}) \vee (BY._{y2} = AY._{y2} \wedge BY._{y1} > AY._{y1}) \vee$ $(AY._{y1} = BY._{y1} \wedge AY._{y2} = BY._{y2}))$

Relation	**A** *northwest* **B**
Definition	(AX {b, m} BX $\wedge$ AY {bi, mi, oi} BY) $\vee$ (AX {o}BX $\wedge$ AY {bi, mi}BY)
Refined formulation	$(((AX._{x2} < BX._{x1}) \vee (AX._{x2} = BX._{x1})) \wedge ((BY._{y2} < AY._{y1}) \vee (BY._{y2} = AY._{y1})$ $\vee (BY._{y1} < AY._{y1} \wedge AY._{y1} < BY._{y2})) \vee$ $((AX._{x1} < BX._{x1} \wedge BX._{x1} < AX._{x2}) \wedge ((BY._{y2} < AY._{y1}) \vee (BY._{y2} = AY._{y1})))$

Relation	**A** *northeast* **B**
Definition	(AX {bi, mi} BX $\wedge$ AY {bi, mi, oi} BY) $\vee$ (AX {oi}BX $\wedge$ AY {bi, mi}BY)
Refined formulation	$(((BX._{x2} < AX._{x1}) \vee (BX._{x2} = AX._{x1})) \wedge ((BY._{y2} < AY._{y1}) \vee (BY._{y2} = AY._{y1})$ $\vee (BY._{y1} < AY._{y1} \wedge AY._{y1} < BY._{y2})) \vee$ $((BX._{x1} < AX._{x1} \wedge AX._{x1} < BX._{x2}) \wedge ((BY._{y2} < AY._{y1}) \vee (BY._{y2} = AY._{y1})))$

Relation	**A** *southwest* **B**
Definition	(AX {b, m} BX $\wedge$ AY {b, m, o} BY) $\vee$ (AX {o}BX $\wedge$ AY {b, m}BY)
Refined formulation	$((AX._{x2} < BX._{x1}) \vee (AX._{x2} = BX._{x1}) \wedge ((AY._{y2} < BY._{y1}) \vee (AY._{y2} = BY._{y1})$ $\vee (AY._{y1} < BY._{y1} \wedge BY._{y1} < AY._{y2}))) \vee$ $((AX._{x1} < BX._{x1} \wedge BX._{x1} < AX._{x2}) \wedge ((AY._{y2} < BY._{y1}) \vee (AY._{y2} = BY._{y1})))$

Relation	**A** *southeast* **B**
Definition	(AX {bi, mi} BX $\wedge$ AY {b, m, o} BY) $\vee$ (AX {oi} BX $\wedge$ AY {b, m}BY)
Refined formulation	$((BX._{x2} <= AX._{x1}) \wedge ((AY._{y2} < BY._{y1}) \vee (AY._{y2} = BY._{y1})$ $\vee (AY._{y1} < BY._{y1} \wedge BY._{y1} < AY._{y2}))) \vee$ $(((BX._{x1} < AX._{x1} \wedge AX._{x1} < BX._{x2})) \wedge (AY._{y2} <= BY._{y1}))$

Relation	**A** *left* **B**
Definition	*AX* {b, m} *BX*
Refined formulation	$(AX._{x2} <= BX._{x1})$

Relation	**A** *right* **B**
Definition	*AX* {bi, mi} *BX*
Refined formulation	$(BX._{x2} <= AX._{x1})$

Relation	**A** *below* **B**
Definition	*AY* {b, m} *BY*
Refined formulation	$(AY._{y2} <= BY._{y1})$

Relation	**A** *above* **B**
Definition	*AY* {bi, mi} *BY*
Refined formulation	$(BY._{y2} <= AY._{y1})$

Chapter 6

EMIMS-S (Extended Medical Image Management System with Salient Objects Support)

EMIMS-S (Extended Medical Image Management System with salient objects support) is an extension of EMIMS[13]. EMIMS (Extended Medical Image Management System) is presented in [13] as a prototype to demonstrate similarity-based image data modeling and processing by. EMIMS-S demonstrates image data management that also involves salient-objects-based queries. With EMIMS-S, we demonstrate the following issues discussed in this thesis.

- Implement the salient object data repository model
- Extraction of salient objects of interest from an image
- Capture spatial features of salient objects and use them for retrieval and description purposes
- Enable similarity-based retrieval of images by their salient objects

With EMIMS-S, retrieval is possible using either the image in its entirety or using the features of the salient-objects.

EMIM-S is developed as an application that can run in a client-sever environment. J2SE (Java 2 Platform, Standard Edition, v 1.4.2) and oracle 9i enterprise edition are used in the development in a windows 2000 environment. JDBC (Java Database Connectivity) is used for the communication between the client application and the Oracle database. The Oracle interMedia model is used for the storage and management of image data and its features.

Oracle interMedia is designed to manage media content in an Oracle8i and Oracle9i database. interMedia is a standard feature, enabling Oracle8i and Oracle9i to manage rich content, including text, documents, images, audio, video, and location information, in an integrated fashion with traditional business data [40].

6.1 Structure of EMIMS-S

The complete structure of EMIMS-S is shown in Figure 6-1 below. The shaded regions show the extensions made to the EMIMS implementation to integrate support for salient objects. In addition to the extension of the core classes, the data entry interfaces and query interfaces are extended to integrate salient objects specification and queries based on salient objects respectively.

The user interfaces

EMIMS-S consists of two basic user interfaces; the data entry interface and the query interface. These interfaces implement the image data entry and query integrating both the main images and salient objects of interest.

<u>The Data Entry Interface</u>

The Data entry interface provides an interface that allows the user to insert both the image and the salient objects.

<u>The Query Interface</u>

The query interface allows image matching with the following functionalities:

- The entire image similarity (the EMIMS implementation),
- Similarity-based retrieval based on salient-objects,
- Optional possibility of using the spatial position of the salient object as an additional criteria in the query formulation, and

- Locate the spatial position of salient objects that are the causes of similarity for images retrieved as a result of salient-objects-based similarity retrieval using their MBRs.

The classes

The Connection class

The connection class is migrated from EMIMS[13]. It establishes client connection to the database using JDBC interface and maintains the connection.

The QueryManager class

The query manager class is the EMIMS [13] class that implements the similarity-based selection operator, the join operator, and others discussed in the earlier chapters. These include: the similarity join/SimJoin, the query by example/QBE, Insert, Mine and other useful operators.

The QueryManager-s class

QueryManager-s is a class extended from the QueryManager class. It inherits all the methods of QueryManager (SimJoin, QBE, Insert, Mine, and others). In addition, it makes the following major extensions to allow salient-objects insertion and retrieval based on salient objects.

- **Insert Methods**: QueryManager-s implements three insert methods, one for the insertion of the main image, another for the insertion of the salient object, and a third one for the insertion of descriptive metadata on salient objects
    - **Insert**(table, imagePath, metadata, MBR): This method is used to insert the main image. It extends the insert method of the QueryManager class with an additional parameter, the MBR of the image to be inserted.

- **Insert**(salientImagesTable, salientImagePath): inserts the salient object (image) int the salient objects table. This method inserts only the image and its features.

- **InsertSalientDescription**: inserts the metadata descriptions of a salient object. These include the MBR and other descriptions specified for the salient object.

- **QBESalient**: this is a method that implements the salient-object-based similarity selection. It takes the salient object as input and retrieves images with similar salient objects. It also takes the position of the salient as an additional optional parameter and performs retrieval considering the position. As an example, if the salient object is at the *top left* of the image and retrieval considers position, only images with similar salient objects and at the *top left* position are returned. The final result is the same as that of QBE method of the QueryManager class, that is, the returns are still the main images.

The MBR Class

The MBR class implements the minimum bounding rectangle entity required both for the main image and the salient objects for use at the client side to process MBR related functionalities. It provides the following useful methods in the query operation:

- Methods *getHeight*, *getwidth*, and *getSize* are used to access the height, width, and size of the MBR respectively.

- The Method *getPosition* returns the position of an arbitrary MBR with reference to the MBR object. The result will be one of the nine positions discussed in chapter 5; these are one of *top*, *bottom*, *left*, *right*, *top right*, *top left*, *bottom right*, *bottom left*, and *center*.

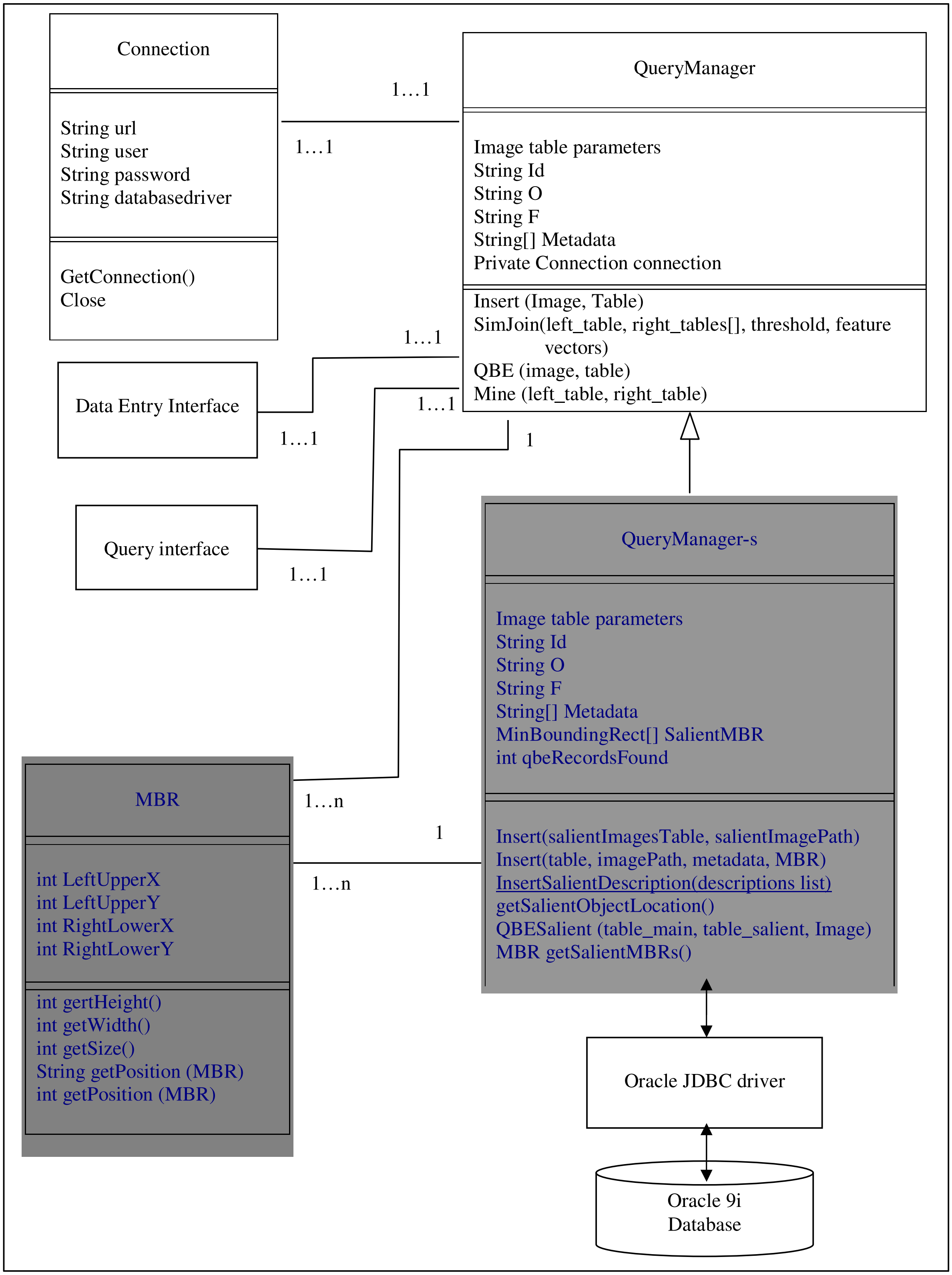


Figure 6-1 Structure of EMIMS-S

6.2 The Sample Database

EMIMS-S is implemented under oracle 9i with an application to medical images. The database implements the data repository model extension proposed for salient objects integration. It allows the storage of both the feature and spatial information of the main image and constituent salient objects. In addition to content information, it also allows capturing and storage of metadata information related to the salient objects.

6.2.1 EMIMS-S tables

The following tables are used for the implementation of EMIMS-S.

- **DOCTOR**(DSN, Name, Specialization, P_History)
  Basic Information on the medical Doctor

- **HOSPITAL**(H_CODE, NAME, ADDRESS, SECTIONS)
  Basic information on the hospital

- **MED_EXAM** (SSN, DSN, H_CODE, ME_CODE, DATEOFEXAM, C_PRESENTATION, CASE, M_HISTORY, FINDINGS, DIAGNOSIS, M_IMAGE)
  Detail Information on patient medical examination

- **M** (ID, O, F, RECT, ME_CODE, IMAGE_PATH, P)
  Main images table, uniquely identified by ID

- **S** (ID, O, F)
  Salient objects table, stores each salient object and its feature vector. ID is the unique identifier.

- **S_A**(SALIENT_ID, IDMAIN, RECT, ANOMALY_TYPE, CASE, DIAGNOSIS, FINDINGS, REMARK)
  Metadata description of the salient objects. This table stores semantic textual description of the salient objects and the MBRs.

- **PATIENT**(SSN, NAME, DATEOFBIRTH, R_ADDRESS, R_HISTORY, M_HISTORY
  Basic patient information, Uniquely identified by patient social security number (SSN)

6.2.2 Implementation of spatial operators

The MBR objects

The MBR objects are implemented in the Oracle database as object types with four attributes corresponding to the coordinates of the MBR. These MBR types are used as field types in the image tables and as parameters in the nine spatial operators discussed below.

The nine spatial operators

The spatial operators that determine the position of a salient object within the main image are implemented in the Oracle database with functions written using PL-SQL. The functions are: TOP_RIGHT, TOP_LEFT, BOTTOM_RIGHT, BOTTOM_LEFT, RIGHT, LEFT, TOP, BOTTOM, and CENTER. Each of these functions take two MBR objects (MBR of the salient object and MBR of main image) as parameters and return either 0 or 1. Thus, a return of 1 from the function TOP_LEFT indicates that the salient object is at the top left position. A return of 0 from the same function tells that the salient object is not at the top left position. This implementation allows the nine operators to be integrated into any queries submitted from clients.

6.3 The user Interfaces

The user interfaces of EMIMS-S is constituted of the data entry interface migrated from EMIMS, the salient object specification (data entry for salient objects), and the extended query interface (both main image-based and salient-object-based).

The EMIMS-S Data entry Interface

Main Image insertion interface

The EMIMS interface [13] allows insertion of the main image into the oracle database. In addition to its original functionality, this interface is extended to automatically generate and show the pixel coordinates of the main image as soon as it is retrieved from file. The MBR is then persisted as the spatial information of the image relative to which spatial position of salient objects can be captured. This extended interface is shown below (Figure 6-2).

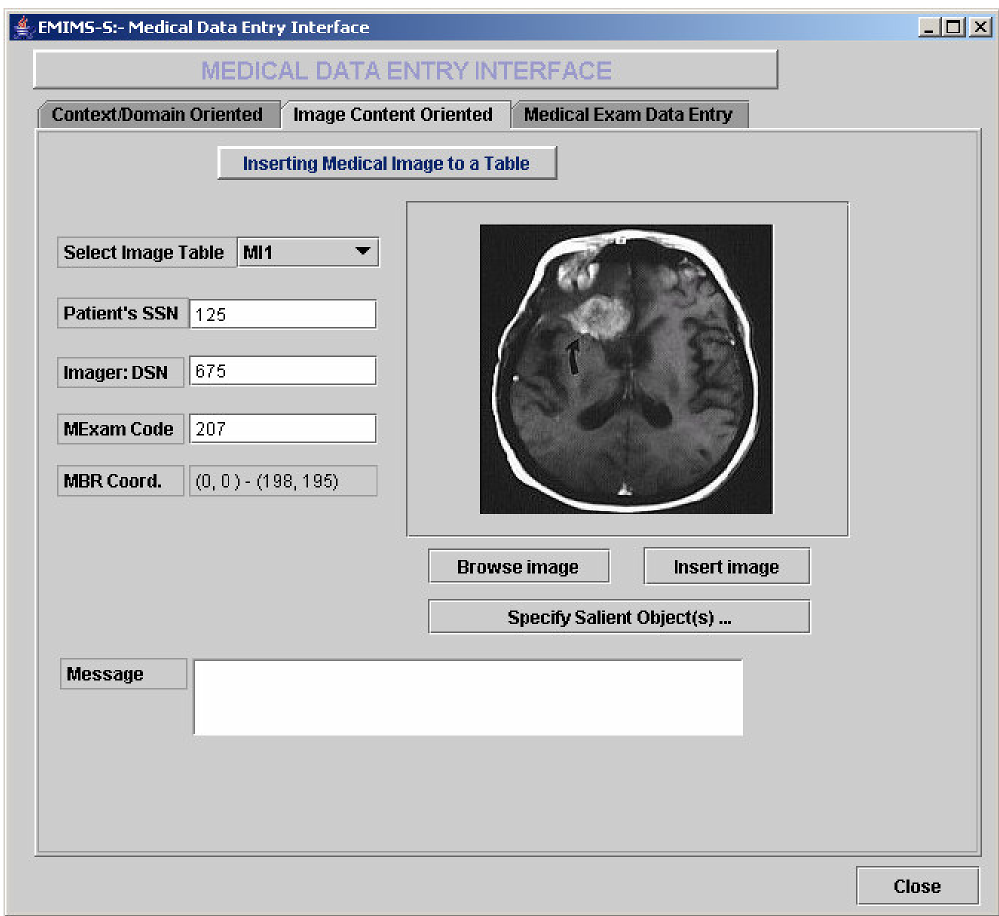


Figure 6-2 The Data entry interface of EMIMS extended with MBR inclusion

Once the main image is inserted, the interface allows specification and insertion of salient objects and their metadata information.

The salient object specification Interface

Once the main image is inserted to the database, EMIMS-S allows specifying one or more salient objects and storing with its spatial and descriptive metadata information. As shown in Figure 6-3 below, when the user selects a rectangular region of the image, the following are performed:

- The selected rectangular region (salient object) is extracted and treated as a separate image in a temporary file for insertion to the database.
- The corresponding MBR coordinates for the selected part corresponding to the pixel values are automatically generated and shown.
- The position of the selected salient object within the image is computed using our definitions of chapter 2 and displayed. Percent of the selected salient object is also shown

As in the discussion in the earlier chapters, the combination of content-based retrieval and metadata retrieval can result in a more efficient multi-criteria query. Describing an image or the salient object with high level semantics is very important specially in a medical application. Information such as the doctor's observation of the anomaly in the image (salient object) and the diagnosis need to be described using textual description. EMIMS-S allows describing the salient object with illustrative textual data. A physician can therefore select an anomalous part of the image (the salient object) and then give it a textual description (Fig 6-4). This allows capturing of both the text and content information.

After specifying the salient object and important metadata information, the user can click on the *insert salient object* button and save the salient objects information to the database. It is possible to select additional salient objects and insert to the database in case the user needs to specify more than one salient objects.

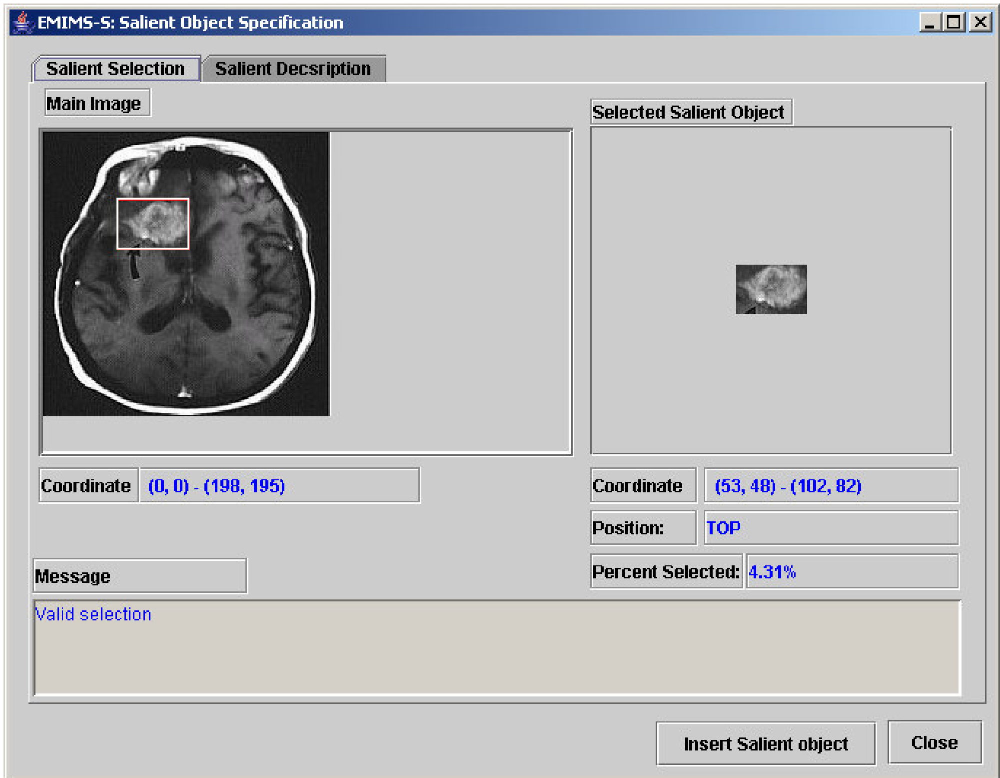


Figure 6-3 Salient Object specification interface

For the main image, the coordinate of the left upper corner will always have a value of (0,0) and the right lower corner will have a value of (w, h) where w and h correspond to the width and height of the image in pixels respectively. Therefore, an image with MBR {(0, 0), (198,195)} has a total number of pixels of 38,610.

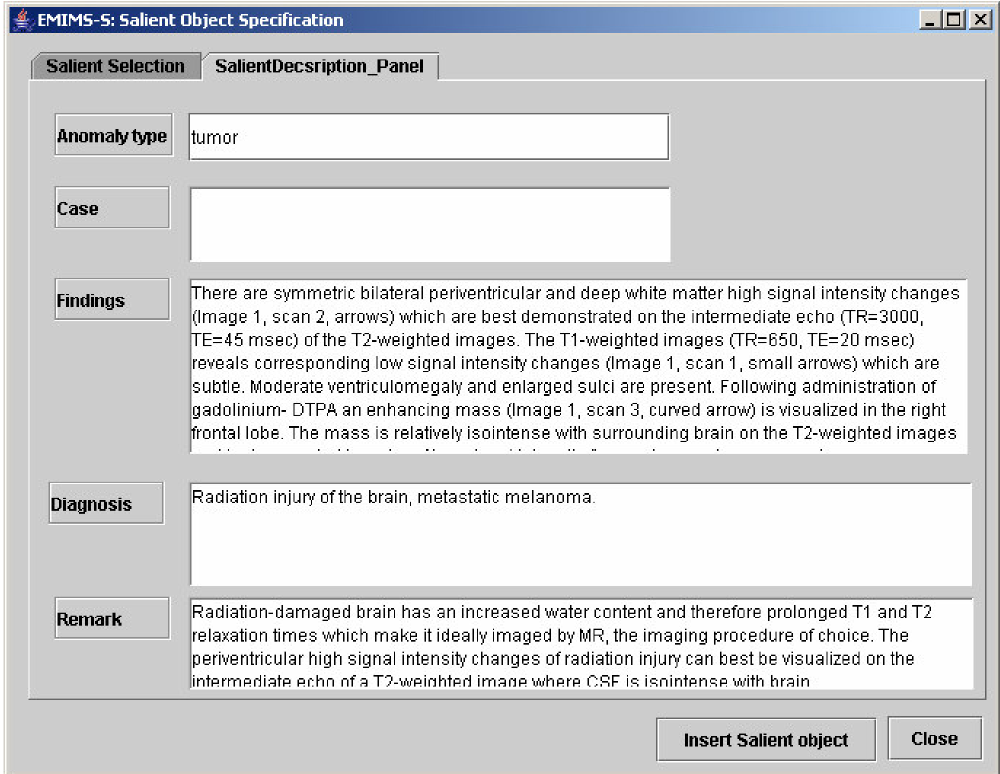


Figure 6-4 EMIMS-S Salient object metadata description interface

The Query interface

EMIMS-S extended the query interface of EMIMS(Figure 6-5) by including the following additional functionalities:

- Salient-object-based similarity matching,
- Combination of salient-object-based similarity and spatial position of the salient object within the image in retrieval,
- Visualization of the salient objects of resulting images that are the causes for the similarity, and
- Retrieval of metadata information used to describe the salient object.

With EMIMS-S query interface, the user has the option to use the main image or select a salient object of interest and use it for similarity comparison. When a salient object is used, the user has the option to consider the position of the salient object in the query (Figure 6-5).

The position of the salient object within the image (*top right*, *top left*, *bottom right*, *bottom left*, *right*, left, *top*, *bottom*, and *center*) is detected automatically when the user selects a rectangular region of the image. This information will determine the query when the user selects the option to consider salient-object position in the query. The following example queries show types of possible queries

1. *Find all images in table M, that have similar salient object to the salient object s_q of the query image q.*

Such a query can generally be formulated as:

```
SELECT SimScore(sq, s.o) score, m.A.MBR, s.As.MBR
FROM Images m, Sal_Objects s, S_A sa
WHERE (m.ID = sa.idMain) AND (sa.salient_id = s.id) AND
      isSimilar(sq, s.o, color, texture, shape,location, ε)
ORDER BY score
```

Below is an actual SQL generated when the query shown in Figure 6-5 is executed. In this query, M is the main images table, S is the salient objects table, S_A is the metadata description table for the salient objects corresponding to the A_s component of the salient objects repository, `QBE_TEMP` is a temporary table used to store the salient object of the query image.

```
SELECT ORDSYS.ImgScore(1) AS SCORE,
  m.ID, m.O, m.F, m.ME_CODE,m.IMAGE_PATH, s.id sal_Id,
  m.rect.lux m_lux, m.rect.luy m_luy,
  m.rect.rlx m_rlx,  m.rect.rly m_rly,
  sa.rect.lux s_lux , sa.rect.luy s_luy,
  sa.rect.rlx s_rlx, sa.rect.rly s_rly

FROM M m, S s, S_A  sa

WHERE  (m.ID = sa.idMain) AND (sa.salient_id=s.id)  AND
  ordsys.imgSimilar((SELECT QBE_TEMP.F FROM QBE_TEMP WHERE ID = 1),
  s.F,'color=1 texture=1 shape=1 location=1',45.0,1)=1
ORDER BY SCORE
```

2. *Find all images in table M, that have similar salient object to the salient object s_q of the query image q with the same position within the main image*

Assuming that the position of the salient object within the query image is *top left*, A general formulation of this query can look like:

```
SELECT SimScore(sq, s.o) score, m.A.MBR, s.As.MBR
FROM Images m, Sal_Objects s, S_A sa
WHERE (m.ID = sa.idMain) AND (sa.salient_id=s.id)
       AND TOP_LEFT (m.MBR, s.As.MBR) AND
      isSimilar(sq, s.o, color, texture, shape, location, ε)
ORDER BY score
```

Below is an actual SQL generated when a query performed with salient-object similarity and position consideration is executed. In this example case, the salient object of the query image is located at the top left of the image, therefore, the result will contain only images with similar salient objects and located at the top left position.

```
SELECT
ORDSYS.ImgScore(1) AS SCORE,
   m.ID, m.O, m.F, m.ME_CODE, m.IMAGE_PATH, s.id sal_Id,
   m.rect.lux m_lux , m.rect.luy m_luy,
   m.rect.rlx m_rlx,  m.rect.rly m_rly,
   sa.rect.lux s_lux , sa.rect.luy s_luy,
   sa.rect.rlx s_rlx, sa.rect.rly s_rly

FROM M m, S s, S_A  sa

WHERE  (m.ID = sa.idMain) AND (sa.salient_id=s.id)  AND
       AppAdmin.TOP_LEFT(m.rect, sa.rect) = 1 AND
       ordsys.imgSimilar((SELECT QBE_TEMP.F FROM QBE_TEMP WHERE ID = 1),
       s.F,'color=1 texture=1 shape=1 location=1',20.0,1)=1
ORDER BY SCORE
```

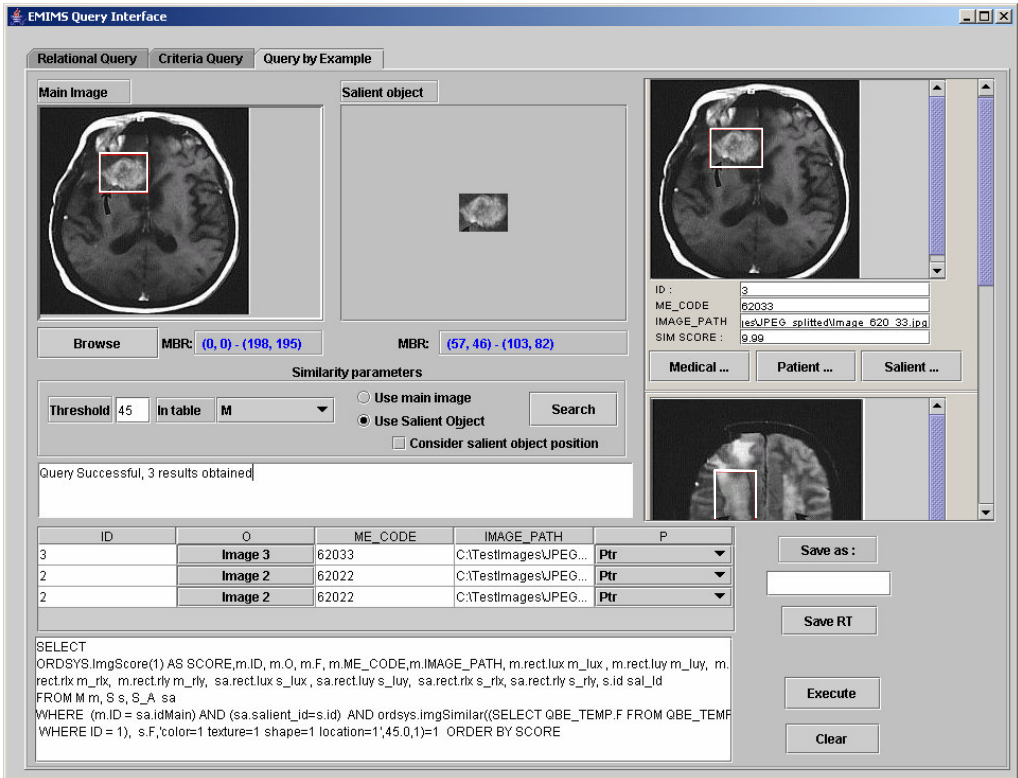


Figure 6-5 The query interface with salient-Object-based query integrated

An important benefit of considering the position of salient objects its discriminatory power, resulting in better selectivity. Salient objects with different size and position can result to be similar to the query salient object due to, for example, closeness of the distribution of the color in the color histogram. This result can be contrary to the human judgment in some scenarios, though it is found computationally similar. Therefore, considering the position of salient objects as additional search criteria complements the use of physical features (color, shape, texture, etc).

Once query results are retrieved using salient objects, EMIMS-S allows visualization of salient object metadata in addition to the EMIMS implementation of viewing patient and medical details. Clicking on the salient details button displays metadata information of the salient object (Figure 6-6).

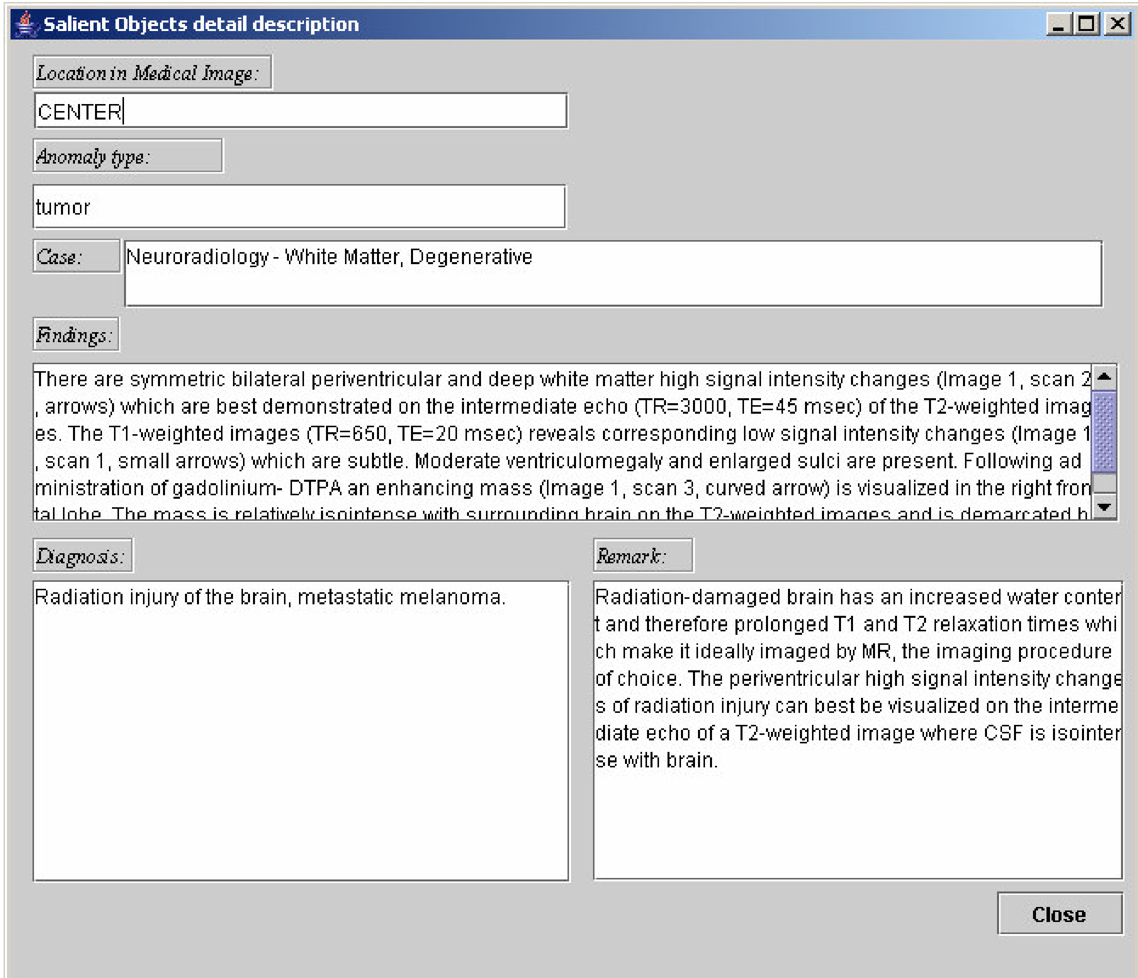


Figure 6-6 The salient Object details window

6.4 **Experimental** comparison of whole-image-based and salient-object-based image queries

Objective of the experiment

The objective of the experiment is to compare the retrieval efficiency of using the entire image, the salient object, and the salient object with position consideration. To compare these three forms of retrieval, precision and recall measurements are used.

Relevance

The relevance of the result of retrieval in this experiment is defined in terms of containing an object similar to a salient object of the query image.

Precision and recall

Recall is the ratio of the number of relevant records retrieved to the total number of relevant records in the database. Precision is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved. These are usually expressed as a percentage.

Precision and recall are concepts often used to measure the retrieval efficiency in text searches. Records must be considered either relevant or irrelevant when calculating precision and recall. This causes problems as individual perceptions differ: what is relevant to one person may not be relevant to another. Often, recall is estimated by identifying a pool of relevant records and then determining what proportion of the pool the search retrieved. In text retrieval, some of the ways of creating a pool of relevant records are: using all the relevant records found from different searches, and manually scanning several journals to identify a set of relevant papers.

The experimental steps

1. 112 different brain images are stored in the main images table, M. these images are learning files obtained from the American College of Radiology[1].

2. 136 salient objects were extracted and stored in the salient objects table, S. For some of the images, more than one salient objects are specified.

3. Eight images are selected as query images to test the retrieval effectiveness of the queries. For each of these images, different set of images are manually (visually) identified as relevant (Table 9).

4. For each of the eight images, the three types of queries (using the whole image, using salient objects, and using salient objects with position consideration) are performed, a total of 24 queries are run. The results shown in Table 9 are obtained. A threshold value (ε) of 20 is used for each of the queries performed.

5. For each of the resulting images of each query, relevant retrieval and total retrieval are recorded. Returned images are counted as relevant when they are found to be in the sent of initially identified relevant images. These numbers are used to compute the precision and recall of the retrieval (Table 10).

[1] http://www.learningfile.com (Last consulted: 15 May, 2004)

Query Image	# of relevant images (in M)	whole image-based query		Salient-object-based query		Salient-object-based query with position considered	
		Total retrieved	Relevant retrieved	Total retrieved	Relevant retrieved	Total retrieved	Relevant retrieved
A	6	26	2	6	4	1	1
B	8	17	3	54	8	7	2
C	6	59	5	50	6	8	3
D	7	15	2	53	6	10	4
E	3	62	3	14	3	2	1
F	3	57	2	40	3	7	1
G	4	46	1	57	4	8	3
H	5	20	3	5	2	1	1

Table 9 Relevant images of the 8 query images and results of retrieval

Query Image	whole image-based query		Salient-object-based query		Salient-object-based query with position considered	
	precision	recall	precision	recall	precision	recall
A	7.69	33.33	66.67	66.67	100.00	16.67
B	17.65	37.50	14.81	100.00	28.57	25.00
C	8.47	83.33	12.00	100.00	37.50	50.00
D	13.33	28.57	11.32	85.71	40.00	57.14
E	4.84	100.00	21.43	100.00	50.00	33.33
F	3.51	66.67	7.50	100.00	14.29	33.33
G	2.17	25.00	7.02	100.00	37.50	75.00
H	15.00	60.00	40.00	40.00	100.00	20.00

Table 10 Precision and recall from retrievals

Figures 6-7, 6-8, and 6-9 below show the comparison of precision, recall, and total retrieval of each of the queries.

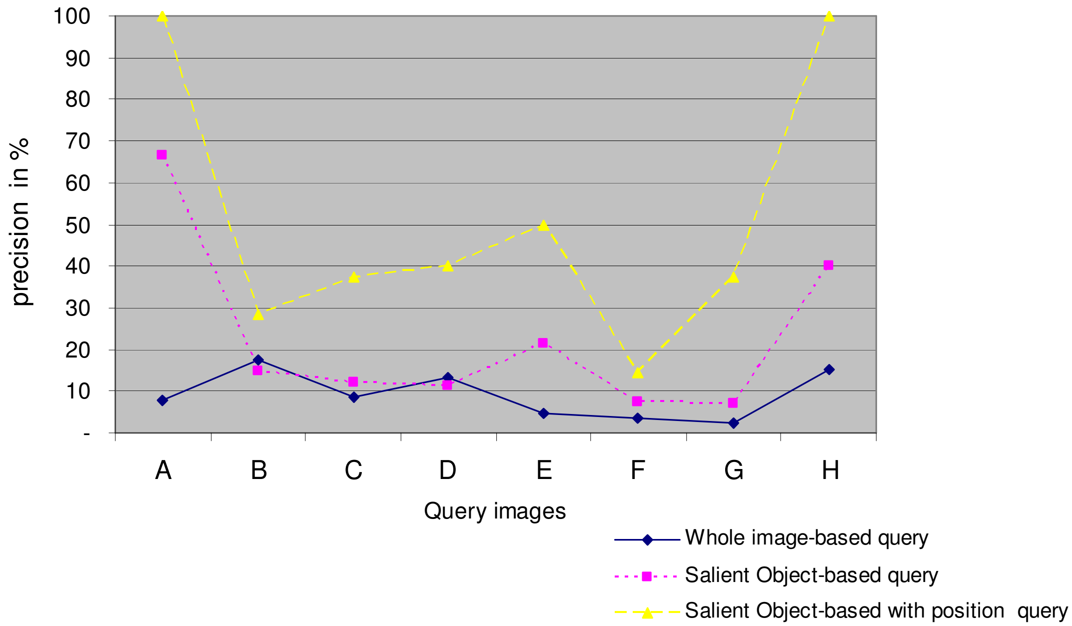


Figure 6-7 Comparative precision of the three types of queries

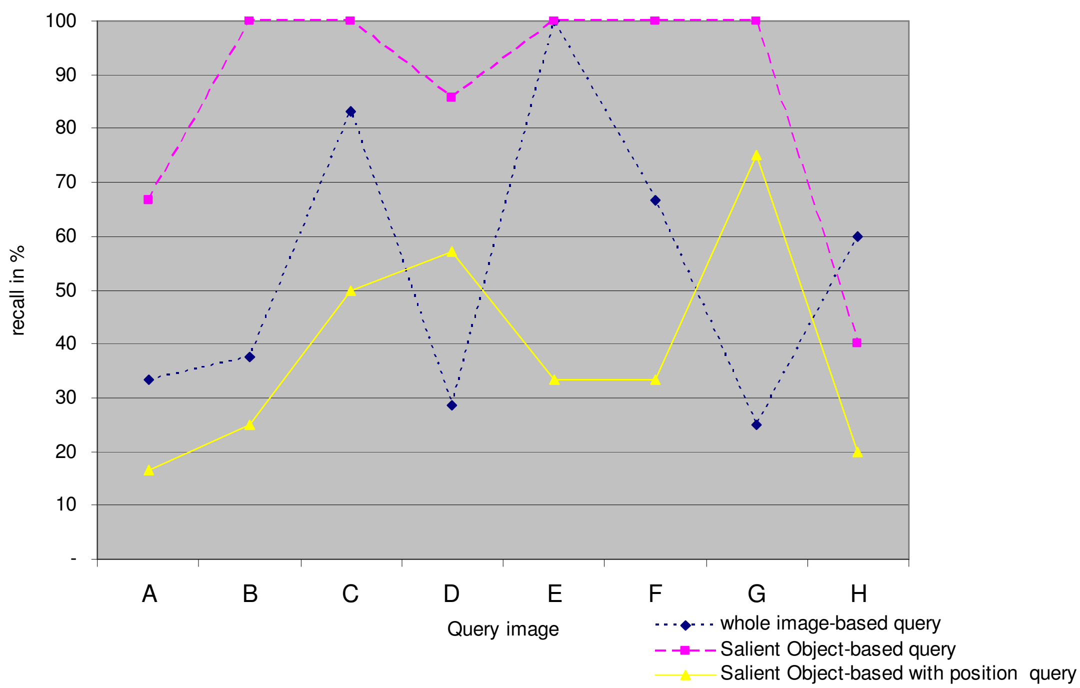
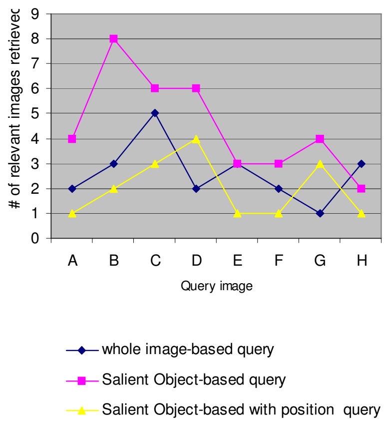
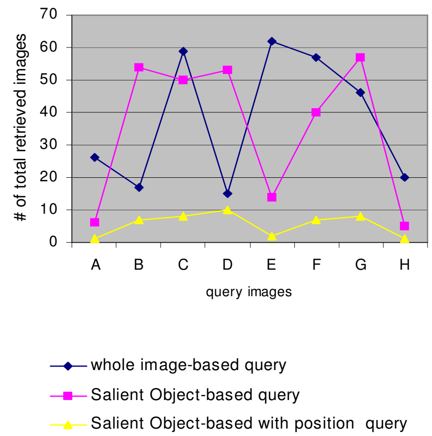


Figure 6-8 Comparative recall of the three types of queries



a. Total relevant retrievals



b. Total retrievals

Figure 6-9 Total relevant retrieval and total retrieval

<u>Discussion</u>

The graph in Figure 6-7 shows that salient-object-based retrieval with position consideration is more precise than the other two types of retrievals. This indicates that, when salient objects with position as an additional predicate are used as a basis of retrieval, the results obtained contain better proportion of relevant images as compared to the other queries though the number of images retrieved are relatively small(Figure 6-9, b). Whole image-based retrieval indicates less precision as compared to the two.

The recall graph of Figure 6-8 indicates salient-objects-based query at the highest position. This is an indication that salient-objects-based queries have better “knowledge” of the database. That means, they return generally higher number of relevant records as compared to the other two. This in fact, is due to the fact that relevance retrieval, in our case, is considered to be the one that contains a similar salient object. Figure 6-9 a also supports this idea.

Summarizing, our experiment shows that, in a similarity-based image retrieval where salient-objects are of more interest, the use of the entire image is a crude approach and will not result in a good retrieval. Salient-objects-based retrieval resulted in both better precision and recall. Moreover, the salient-object-based retrieval with the addition of positional predicate increased the selectivity by reducing the potential number of images to be retrieved.

As Figure 6-9 b indicates, it can not be deduced whether salient-objects-based (without position predicate) or whole-images-based retrieval has high selectivity. This results due to the nature of similarity-based retrieval itself. Therefore, generally, salient-objects-based retrieval has higher retrieval efficiency (recall and precision), but its selectivity can not be generally deduced. It is also worth noting that variation of the selection of the salient objects would result in a very different type of results in repetitive queries, as manual selection of salient object does not always result in exactly the same salient object between different queries.

In this experiment, queries are performed using 8 sample images. We therefore remark that, repeated experiments with higher number of images and more sample queries would result in a more comprehensive result.

6.5 Summary

The EMIMS-S prototype has demonstrated the viability of image retrieval by visual content that takes the salient objects and their spatial position into consideration. EMIMS-S implements the extended data repository model to capture and store the physical, semantic, and spatial information of the main images and salient objects.

The spatial information is captured using the Minimum Bounding Rectangles (MBRs) whose coordinates correspond to image pixels.

It has shown how salient objects can be integrated in the retrieval of images with the notion of similarity. Moreover, the prototype demonstrated the usefulness of the consideration of the spatial information of the salient objects and the benefits in application domains where the spatial location of the salient objects with respect to the main images is important.

The extended query manager class enables storage and retrieval of salient objects in addition to providing the full functionality of the original query manager class as it is extended by sub-classing the original query manager using additional functionalities (methods).

Chapter 7

Conclusions and Future works

7.1 Conclusions

The importance of salient-objects-based image queries has been discussed thoroughly in the preceding chapters. Image queries to-date were mainly based on the image in its entirety and no detailed study or work on formalizing similarity-based image retrieval by considering salient objects has been made.

In this thesis, we have assessed and proposed operators that integrate salient-objects-based image retrieval into content-based image databases. The major contributions that this thesis has made to content-based image databases are the following:

- We have made an extension to the data repository model proposed in [13] so that spatial information of salient objects within the image is captured.
- We have extended the similarity-based selection operator proposed in [13] in such a way that similarity-based image retrieval can be made based on salient-objects.
- We have developed spatial operators for the computation of the relation between a salient object and the image.
- We have presented a refined formulation of spatial relations between salient objects in compliance with our extended model for salient objects data repository model.
- We have developed an extended prototype that demonstrates the viability of salient-objects-based image queries.

One of the challenges in content-based image retrieval is the bridging of the semantic gap between the low-level image features and their higher level semantics. This thesis has demonstrated intermediate level image data utilization between the low-level (whole image) and a higher-level (salient-objects). A notable contribution is therefore, moving a step forward towards reducing of the semantic gap.

7.2 Further works

Segmentation is an important task in salient-object-based image queries to identify regions of interest. This is done either manually or automated. As there is no standard algorithm or tool to-date to perform automatic segmentation, the integration of automatic segmentation results into salient-object-based image queries remains an area to be explored.

As a regular geometric approximation to salient objects, we have used minimum bounding rectangles. Like most approximations, the relations between the minimum bounding rectangles do not always correspond to the actual relation between the salient objects. Therefore, refinement steps are needed to make further computations of the actual relation. Exploration of this task is also left further analysis.

Data structures involving minimum bounding rectangles are often organized into an index-structure to facilitate retrieval. Exploring the minimum bounding rectangles used in this work from the perspective of indexing is another area to be investigated further.

REFFERENCES

[1] V.S. Subrahmanian. *Principles of Multimedia Database Systems*. San Francisco, California: Morgan Kaufmann Publishers Inc. 442p., ISBN 81-7868-041-0, 1998.

[2] H. Koch, S. Atnafu. *Processing a multimedia join through the method of nearest neighbor search*. Information Processing Letters, 82(5) pp. 269 - 276 , June 2002.

[3] N. Roussopoulos, S. Kelly, and F. Vincent. *Nearest Neighbor Queries*. Proc. of ACM-SIGMOD, pp. 71-79, May 1995.

[4] Y. Rui, T.S. Huang, and S.-F. Chang. *Image retrieval: Past, Present, and Future*. Journal of Visual Communication and Image representation,10:1-23, 1999

[5] S. Berchtold, Daniel A. Keim, and Hans-Peter Kreigel. *The X-tree: An indexing structure for high-dimensional data*. In Proceedings of the VLDB Conference, pp. 28-39, Bombay, India, September 1996.

[6] V. Oria, M. T. Özsu, L. I. Cheng, P.J. Iglinski and Y. Leontiev, *Modeling Shapes in an Image Database System*. In *Proceedings of the 5th International Workshop on Multimedia Information System* , Indian Wells, Palm Springs Desert, California, USA, pp. 34-40, October 1999.

[7] S. Atnafu, L.Brunie, and H.Kosch. *Similarity-Based Operators and Query Optimization for Multimedia Database Systems*. Proceedings of the IEEE International Symposium on Database Engineering & Applications, p 346, 2001

[8] Jhon P. Eakins and Margaret E. Graham. *Content-Based Image Retrieval: A report to JISC Technology Applications Programme*. Inst for Image data Research, Univ. of Northumbria at Newcastle, January 1999.

[9] V. Oria, M.T. Özsu, L. Liu, X. Li, J.Z. Li, Y. Niu, and P.J. Iglinski. *Modeling Images for Content-Based Queries: The DISIMA Approach*. VIS'97, San Diego, pages 339-346, 1997.

[10] S. Nepal and M.V. Ramakrishna. *Query Processing Issues in Image (Multimedia) Databases*. Proc. Of the 15th International Conference on Data Engineering, Sydney, Austrialia, 22-29, 23-26 March 1999.

[11] J.Z. Li, M.T. Özsu, D. Szafron, and V. Oria. *MOQL: A multimedia object query language*. Proc. of the 3rd Int. Workshop on Multimedia Information Systems, pp. 19--28, Como, Italy, 1997.

[12] V. Oria, B. Xu, and M. T. Ozsu. *VisualMOQL: A visual query language for image databases*. Proceedings of 4th IFIP 2.6 Working Conference on Visual Database Systems - VDB 4, pp. 186---191, L'Aquila, Italy, May 1998.

[13] Solomon Atnafu. *Modeling and Processing of Complex Image Queries*, Ph.D Thesis. Laboratoire d'Ingénierie des Systèmes d'Information (LISI); INSA de Lyon. July 2003.

[14] Stonebreaker, M., *Object-Relational DBMSs*: The Next Great Wave. Morgan Kaufmann publishers, 1996.

[15] M. Ortega-Binderberger, K. Chakrabarti. *Database Support for Multimedia Applications*. University of Illinois at Urbana-Champaign. S. Mehrota. University of California at Irvine., August 2001.

[16] C. B ohm, S. Berchtold, and D. A. Keim. *Searching in high-dimensional spaces: Index structures for improving the performance of multimedia databases*. ACM Computing Surveys, 33(3):322 -- 373, September 2001.

[17] Remco C. Veltkamp, Mirela Tanase . *Content-Based Image Retrieval Systems: A Survey* ,2000. http://citeseer.nj.nec.com (consulted on 24 Nov, 2003)

[18] N. Sebe and M.S. Lew. ***Salient points** for **content-based retrieval***. In BMVC'01, pages 401-- 410, 2001.

[19] E.J. Pauwels and G. Frederix. *Finding salient regions in images: Nonparametric clustering for image segmentation and grouping*. Computer Vision and Image Understanding, 75(1,2):73--85, 1999.

[20] A.Dimai. *Unsupervised Extraction of Salient Region-Descriptors for Content Based Image Retrieval*. IEEE 10th International Conference on Image Analysis and Processing , Venice, Italy, p. 686, Sept. 27 - 29, 1999

[21] E. Loupias and N. Sebe, *Wavelet-based Salient Points for Image Retrieval* , RR 99.11, Laboratoire Reconnaissance de Formes et Vision, INSA Lyon, November 1999.

[22] C. Town and D. Sinclair. *Content based image retrieval using semantic visual categories*. AT&T Technical Report, 2001.

[23] J.K. Wu, A.D. Narasimhalu, B.M. Mehtre, J.P. Lam, and Y.J. Gao. *CORE: a content-based retrieval engine for multimedia information systems*. Multimedia Systems, 3(1):25-41, Feb 1995.

[24] M Das, R.Manmatha. *Automatic Segmentation and Indexing of Specialized Databases*. Multimedia Indexing and Retrieval Group, Technical Report, Center for Intelligent Information Retrieval, Department of Computer science, University of Massachusetts, Amherst, MA 01003, 2002.

[25] P. Lyman, H. R. Varian. *How Much Information?* A research at the School of Information Management and Systems at the University of California at Berkeley, October 2000.
http://www.sims.berkeley.edu/research/projects/how-much-info/
(Consulted on December 8, 2003)

[26] Myron Flickner, Harpreet sawhney, Wayne Niblack, et. al. *Query by Image and Video Content: The QBIC System*, IEEE Computer Society, vol28, no 9, pp. 23-32, September 1995

[27] A. Pentland, R. W. Picard, and S. Sclaro . *Photobook - content-based manipulation of image databases*. Int. Journal of Comp. Vision, 18(3):pp. 233-254, 1996.

[28] Alberto Del Bimbo. *Visual Information Retrieval*, San Francisco, California: Morgan Kaufmann Publishers Inc., 270p, ISBN 1-55860-624-6, 1999.

[29] J. Z. Li, M. T. Ozsu, and D. Szafron. *Spatial reasoning rules in multimedia management systems*. Technical Report TR-96-05, Department of Computing Science, University of Alberta, March 1996.

[30] Lei Chen , M. Tamer Ozsu , Vincent Oria. MINDEX: An Efficient Index Structure for Salient Object-based Queries in Video Databases

http://www.db.uwaterloo.ca/~ddbms/publications/multimedia/msj-mindex-leichen.pdf (Consulted on 30 March 2004)

[31] D. Papadias, Y. Theodoridis, T. Sellis, and M. Egenhofer. *Topological relations in the world of minimum bounding rectangles: A study with R-trees*. Proceedings of ACM SIGMOD International Conference on Management of Data, pp. 92--103, San Jose, CA, USA, 1996.

[32] Clementini, E., Sharma, J. and Egenhofer, M. J. *Modeling topological spatial relations: strategies for query processing*, Computers and Graphics, 18(6), pp. 815-822, 1994.

[33] M. Emre Celebi Y. Alp Aslandogan. *Content-based image retrieval incorporating models of human perception*. Proceedings of the IEEE International Conference on Information Technology: Coding and Computing (ITCC), pp.241- 245, 2004. Available on: http://www.cse.uta.edu/Research/Publications/Downloads/CSE-2003-21.pdf /Consulted on 30 March 2004/

[34] Safar M., and Shahabi C. *2D Topological and Directional Relations in the World of Minimum Bounding Circles*. Proc. of IEEE International Database Engineering and Applications Symposium (IDEAS), 1999, pp. 239-247, Montreal, Canada, August 2-4, 1999a.

[35] Thomas Brinkho , Hans-Peter Kriegel, and Ralf Schneider. *Comparison of approximations of complex objects used for approximation-based query processing in spatial database systems*. Proceedings of the 9th International Conference on Data Engineering, pp. 40-49, 1993.

[36] *MPEG-7 Overview* (version 9). International Organisation For Standardisation Iso/Iec Jtc1/Sc29/Wg11 Coding Of Moving Pictures And Audio Iso/Iec Jtc1/Sc29/Wg11n5525 Pattaya, March 2003.

http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm
(Consulted on 3 May, 2004)

[37] Rafael C. Gonzalez, Richard E. Woods. *Digital Image Processing*, Second Edition, Pearson Education, ISBN 81-7808-629-8, 2002.

[38] J. Chen. *Perceptually-Based Texture and Color Features for Image Segmentation and Retrieval* (Ph.D thessis). Northwestern university evanston, illinois, December 2003. http://www.ece.northwestern.edu/~jqchen/publication.html
(consulted on: 21 June 2004)

[39] R. Chbeir, S. Atnafu, L. Brunie. *Image Data Model for an Efficient Multi-Criteria Query: A Case in Medical Databases*. 14th International Conference on Scientific and Statistical Database Management (SSDBM'02), Edinburgh, Scotland, p. 165, July 24 - 26, 2002.

[40] *Oracle® interMedia. User's Guide and Reference*, Release 9.0.1