Salient-Object-Based Image Query By Visual Content

Dawit Bulcha,
Solomon Atnafu
Dept. of Computer Science,
Addis Ababna University,
P.O.Box 1176, Addis Ababa,
Ethiopia
dawitbulcha@yahoo.com,
satnafu@cs.aau.edu.et

David Coquil,
Lionel Brunie
INSA, 20 Avenue A. Einstein
F-69621 Villeurbanne
CEDEX- France
{dcoquil, lbrunie}@insa-lyon.fr

Richard Chbeir
Laboratoire LE2I
Aile de l'Ingénieur,
Université de Bourgogne
BP 47870 21078 Dijon
CEDEX France
richard.chbeir@u-bourgogne.fr

Abstract

Content-Based Image Retrieval (CBIR) has attracted much attention of the research community. As exact matching is not possible with image retrieval, the approach is to use similarity-based matching using the global features of the entire image to compute a similarity score between two images. Equally important is the use of salient-objects: objects in an image that are of particular interest, as the basis of similarity-based computation. However, the current works on CBIR do not address very well the issues related to salient-objects.

In this work, we propose a data repository model so that spatial features of salient objects are captured. Moreover, we propose an extension to the similarity-based selection operator defined earlier to allow salient object based selection. We also propose spatial operators that can be used to compute spatial relations between an image and its contained salient objects.

To demonstrate the viability of our proposals, we extend a previous system named EMIMS, to develop EMIMS-S (Extended Medical Image Management System to support Salient objects). We also experimentally evaluate the retrieval effectiveness of salient-objects-based image queries.

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

1 Introduction

Image retrieval has been a topic of active research since the 1970s. The research communities that are involved in this area are mainly from the fields of computer vision and database [1].

Traditional database management systems mainly deal with textual data. These database systems effectively address the common issues of data integrity, transaction processing, concurrency... In image databases, the traditional approach is a heavy burden on the users as it is impossible to completely describe the content of an image with the features such as its color, shape, texture, and regions. As a result, retrieval of images from an image database requires techniques for processing image queries based on these low-level image features – a technique known as Content-Based Image Retrieval (CBIR). A typical CBIR involves two processes: the extraction of the low-level image features and the management and processing of these features for retrieval. Matching images is performed, by computing the closeness of the features of the images.

In the current state-of-the-art, similarity-based matching is performed by considering the content of the whole image. A comparison that considers part(s) of images for similarity would be a more effective approach of image retrieval, particularly in the application domains where only a part of the image is of interest. These regions of image that are of particular interest are termed as salient-objects.

Image retrieval based on salient objects is the particular focus of this work. 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 paper is outlined as follows: Section 2 discusses related work. Section 3 discusses the extended data repository model for salient objects. Section 4 presents the image query algebra and spatial operators supporting salient-objects. Section 5 discusses EMIMS-S, a prototype that demonstrates the use of salient-objects-based queries. Section 6 presents the conclusions and prospective areas of research.

2 Related Work

A lot of research works have been conducted on image retrieval, especially in the 1990s and later [1, 3]. Content-based retrieval using the visual content of images has been studied by the computer vision community to alleviate the problems of manual image annotations. Related issues such as image data modeling, image query processing have been studied by the database community [4, 3, 1].

Currently, several CBIR systems are in use. Most of these systems use low-level features (color, texture, shape). As stated in [2, 4], each system has made some contribution to this research field. However, these works mostly concentrated on retrieval using the entire image. Moreover, not much attention was given to the modeling and query processing of images that integrate salient objects.

In addition to the image data repository model, the works in [2, 4] have developed a similarity-based algebra. This is a major work that has formalized image data modeling and query processing in the context of an Object-Relational DBMS. Though this work laid a foundation for salient-object data repository, it does not integrate salient objects in similarity retrieval nor spatial relationships between salient objects.

A detailed analysis of spatial relationships is important for modeling the representation of the spatial behavior of salient objects. Spatial, directional and topological relationships between objects are studied. Objects of the real world are usually irregular in shape. As a result, they are approximated with some regular geometric objects.

Several approximations are proposed in the literature to represent these complex objects. Such approximations include: Minimum-bounding rectangle (MBR), Rotated minimum bounding box, Minimum bounding circle, etc. [6]. In this paper we have chosen to use the MBR approximation due to its simplicity, low storage requirement, and popularity in usage.

Another important point is the relation between an image and contained salient objects. The topological and directional relations are not sufficient to describe such relations. In section 4, we define important relations that can be used to describe such relations.

As image data is complex, a model that captures complete and meaningful information about an image is required. The model proposed in [8] provides a global view of an image. It supports both metadata and low level descriptions of images so that a multi-criteria query involving both metadata and low-level properties can be formulated.

S. Atnafu [4] proposed an image data repository model. The model can be used to describe both textual and content information. This model takes into account the storage and retrieval requirements of image data. It also complies with the abstract image model of R. Chbeir et. al. [8]. Based on [2, 4], 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 itself stored as a BLOB in the table or referenced as an external BFILE.*
- *F* *is a feature vector representation of the object O. Stored features include the color, texture, shape, and layout contents extracted from the image.*
- *A* *is an attribute component describing O using textual data annotations.*
- *P* *is a data structure that captures links to images of other tables as a result of a similarity-based join .*

This work emphasized the importance of salient objects and proposed a salient object repository model **S** $(\mathbf{id_s}, \mathbf{F_s}, \mathbf{A_s})$, where,

- id_s *is an identifier of a salient object.*
- F_s *is a feature vector extracted to represent the low-level features of the salient object.*
- A_s *is an attribute component that captures the semantic description of the salient object using textual data.*

This model defines the representation of the salient objects but does not specify the representation of the spatial features of the salient objects. 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 on the query predicate depending on the interest of the user and the application domain.

One of the strengths of the relational model is its strong mathematical foundation, the algebra. Due to their inherent complex properties, image data cannot be adequately managed under the relational systems. Therefore, there is a need for the formalization of a suitable algebra for the management of image data.

A major work in this direction is the work in [4, 8]. This work has developed and formalized a similarity-based image query algebra compatible with the object-relational DBMS. This work has defined important operators such as: *Similarity-based selection operator, similarity-based join operator, etc.*

The similarity-based algebra developed in this work is applied for image retrieval using the features of the entire image. However, the work did not address how similarity-based image retrieval based on salient-objects can be performed. In the next sections, we will explore how the spatial and physical features of salient-objects can be utilized in the similarity-based retrieval of images.

3 Image Data Repository Model for Salient Objects

In this section, we present an extension of the data repository model proposed in [2, 4]. In the original model the A component of the main image may be declared as an object, a set of objects, a table, or a set of tables linked to other tables. This specification makes it robust enough to extend it without violating the compatibility to support salient objects. Moreover, though salient objects are images by themselves, the fact that they are part of an image makes them another characteristics of the image that needs characterization.

The image data repository model discussed in section 2 has the format: M(id, O, F, A, P). In the extension of the model, we include a component MBR_m in A to enable us to characterize salient objects for storage and retrieval.

The storage of the MBR helps to capture the location of salient objects within the image. The A component can also contain other textual descriptions of the image. Whether the salient objects are identified manually or automated, textual or keyword information is an important description that is needed in most applications.

The MBR_S are used as the spatial descriptors of the salient objects within the image. This enables retrieval using the position of the salient object within the image. The link between the image and the salient objects can be captured 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 A_S.

In addition to the MBR_S, A_S captures all semantic descriptions of the salient object with textual description. Such description is important in addition to the use of the MBR_S. For example in a medical application, a physician would describe the nature of a tumor from a brain image.

4 Similarity-Based Algebra for Salient Object-based Image Queries

In the context of CBIR, similarity is the most important notion. Indeed in content-based image retrieval, search is based on similarity instead of exact matching. Therefore, operators that can be used for matching image similarity are required. Though there are several developments in this area, only the works in [4] has made a profound formalization of the notion of similarity. As mentioned in the previous sections, similarity can be matched based on the entire image using global features or using the features of salient objects, which is the main theme of this work.

In this section, we define operators that: match images using the features of the salient objects, determine the spatial position of salient objects within the image, and describe the spatial relationships of salient objects.

4.1 Salient-Object-Based Similarity Selection

We first re-state the similarity-based selection operator defined in [2, 4]. This operator is a unary operator on an image table M(id, O, F, A, P) performed on F. 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:

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

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

Given this definition, 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 $\delta_{O_s}^{\varepsilon}(M)$, 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,

$$\delta_{O_s}^{\varepsilon}(M) = \left\{ \left(id, o', f, a, p \right) \in M / O' \in \prod_{M.o} \left(\delta_{M.id \in I}(M) \right) \right\},$$

Where, $I = \prod_{S.A_s.id}\left(\delta^{\varepsilon}_{O_s}(S)\right)$, and

$$\delta^{\varepsilon}_{O_s}(S) = \left\{\left(id_s, f'_s, a_s\right) \in S / f' \in R^{\varepsilon}\left(S, f_s\right)\right\}$$

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 table of salient objects.

The extension of the similarity-based selection operator to salient-object-based selection involves two steps: similarity-based selection on the salient objects table followed by relational selection on the image table on condition that the salient objects 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. Next, the relational selection on the main image table M retrieves images from the table M whose ids 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 comparison instead of the entire image.

4.2 Spatial Query Operators

Studies on topological and directional relations between two image objects abound [5, 6, 7]. These relations can be used to describe the relation between two salient objects of an image. In addition to the topological and directional relations between several salient objects, equally important is the relation between an image and its contained salient objects. In this section, we 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 .

4.2.1 Main Image - salient object relation

Queries can usually involve positional predicates such as *top right, top left, bottom left, bottom right,* etc. In a medical application, a physician might for example be interested in brain images with a tumor at the top right. Such scenarios indicate the need for a scheme of computing the spatial position of a salient object within the main image.

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 (Figure 4.1).

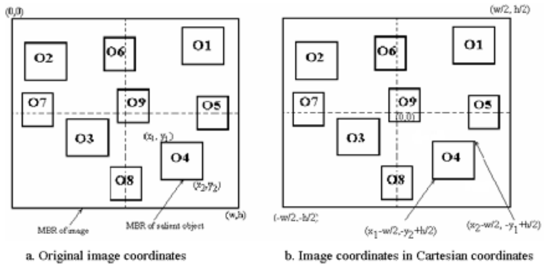

Figure 4.1 Salient objects positions within the main image

Assuming that {(0, 0), (w, h)} are the coordinates of the MBR of the main image and {(x1, y1), (x2,y2)} are the coordinates of the MBR of an arbitrary salient object, the nine positions can be expressed mathematically.

Once the MBRs of the image and of the contained salient objects are determined, we have sufficient information to process queries involving the positions of a salient objects such as:

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

Assuming that Sq, the salient object of the query image, is located at the top left of the image, 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)

4.2.2 Relation between salient objects

It is common to have more than one salient object within a single image. In these cases, it is of interest to describe the relationship between the salient objects. A query may require retrieval of all brain images with two tumors where one is located at the left of the other. This requires retrieval of brain images with salient objects with the relationship right or left. Such relationships are categorized into topological and directional relations.

In this section we will present refined mathematical formulations of how the topological and directional relations defined in [5, 6, 7] can be computed from the MBR_s of the salient objects. Directional relations defined as: *north, south, east, west, northeast, northwest, southwest, southeast, above, below, left* and *right* are considered.

4.2.3 Topological Relations

We considered six topological relations that can be used to describe the position of salient object with respect to each other. These relations are: equal, contains, covers, overlap, meet, and disjoint.

We outline the refined mathematical formulations of six of the topological relations defined in [5, 6, 7].

Let A and B represent arbitrary salient objects and their positions on the x and y axes denoted as AX, AY, and BX, BY. $\wedge$ and $\vee$ are the logical AND and OR operators. The notation { } is equal to the $\vee$ operator over relations.

Let $\{(AX._{x1}, AY._{y1}), (AX._{x2}, AY._{y2})\}$ and $\{(BX._{x1}, BY._{y1}), (BX._{x2}, BY._{y2})\}$ be the coordinates of the MBRs of the objects A and B in the coordinate system.

Then, we adopted the refined formulations of the six topological relations in [5, 6, 7].

5 EMIMS-S

EMIMS-S (Extended Medical Image Management System with Salient objects support) is an extension of EMIMS, which is presented in [4] as a prototype for similarity-based image data modeling and processing. EMIMS-S allows image data management that includes salient-objects-based queries. With EMIMS-S, we will:

- Implement the salient object data repository model,
- Demonstrate the extraction of salient objects from an image approximated by a MBR,
- Demonstrate the capture of spatial features of salient objects for retrieval and description purposes,
- Implement similarity-based retrieval of images by the visual features of their salient objects.

EMIM-S is developed as a client-server application. J2SE and Oracle 9i are used for the development of the user interfaces and for the storage of image data. JDBC is used for the communication between the client application and the Oracle database. The Oracle *inter*Media model is used for the management of image data and its features.

5.1 Structure of EMIMS-S

Extensions are made to EMIMS to support salient objects. Thus, the core classes, the data entry interfaces and query interfaces are extended to integrate salient objects and queries based on salient objects.

5.2 The user Interfaces

EMIMS-S includes two basic user interfaces; the data entry interface and the query interface. The data entry interface allows the user to insert an image and to define its salient objects. The data entry interface is extended to 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.

Once the image is inserted in the database, EMIMS-S allows specifying one or more salient-objects and storing them with their spatial and descriptive metadata information. After specifying the salient object and important metadata information, the user can save the salient objects information to the database. EMIMS-S extends the query interface of EMIMS by including the following additional functionalities:

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

The query interface allows image matching with the following functionalities:

- The entire image similarity,
- Similarity-based retrieval based on salient-objects,
- Possibility of using the spatial position of the salient object as a criteria in the query formulation.

With the EMIMS-S query interface, the user has the option to use the main image or to select a salient object and to use it for similarity comparison. When a salient object is used, the position of the salient object can be included in the query. The position of the salient object within the image is detected when the user selects a rectangular region of the image.

5.3 Experimental comparison of whole image- and salient-object-based image queries

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

The experimental steps

1. 112 different brain images are stored in the main image table, *M.*
2. 136 salient objects are extracted and stored in the salient objects table, S. For some of the images, more than one salient object is specified.

3. 8 images are selected as query images to test the retrieval effectiveness. For each image, sets of images are manually identified as relevant.
4. For each of the 8 images, the three types of queries are performed. A threshold value (ε) of 20 is used for each of the queries (maximum threshold: 100)
5. For each result image of each query, relevant retrieval and total retrieval are recorded. Returned images are counted as relevant when they are in the set of initially identified relevant images. These numbers are used to compute the precision and recall of the retrieval.

Figures 5.1, 5.2 and 5.3 show the comparison of precision, recall, and total retrieval of each query.

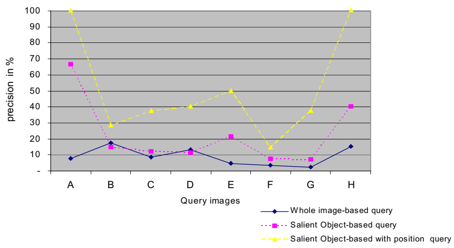

Figure 5.1 Comparative precision of the queries

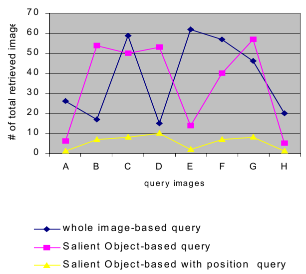

Figure 5.2 Comparative recall of the three types of queries

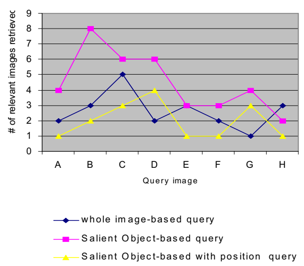

Figure 5.3 Total relevant retrieval

6 Conclusion

In this paper, we have proposed operators that integrate salient-objects-based image retrieval into image databases. The major contributions of this thesis are the following:

- We have made an extension to the data repository model proposed in [2, 4] so that spatial information of salient objects within the image is captured.
- We have extended the similarity-based selection operator proposed in [2, 4] so that 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.
- We have developed a 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 paper 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 References

[1] 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

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

[3] 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

[4] S. Atnafu. Modeling and Processing of Complex Image Queries, Ph.D Thesis. LISI; INSA de Lyon. July 2003.

[5] 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.

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

[7] Clementini, E., etal. Modeling topological spatial relations: strategies for query processing, Computers and Graphics, 18(6), pp. 815-822, 1994.

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