Recently, there has been widespread interest in various kinds of database management systems for managing formatted information. Depending upon the domain and the nature of formatted data, these systems are variously referred to as Multimedia Information Systems, Spatial Databases, Pictorial Information Systems, and Image Database Systems. However, the image data models employed in these systems are not based on any general framework. The model is rather extracted often from the implemented system and hence these data models are shaped by the idiosyncratic characteristics of the domains. Similar kinds of problems plague the query language design.

By studying the application requirements and the limitations of the proposed approaches, we envision a multi-layered structure for retrieval. The various layers in the scheme are: Physical Layer, Spatial and Shape Layer, Iconic and Attribute Layer, and Conceptual Layer. These layers are not designed to operate in isolation but rather work in cooperation. To avoid redundancy in representation the layers are structured to form a lattice. The layers can also be viewed as multiple representations for the same object. This framework is expected to be highly flexible enough so that it can be useful across several application areas.
APPLICATION OF MACHINE LEARNING TECHNIQUES FOR EFFECTIVE RETRIEVAL IN IMAGE DATABASES

FINAL REPORT

V. NAIDU GUDIVADA
VIJAY V. RAGHAVAN

APRIL 30, 1992

U.S. ARMY RESEARCH OFFICE

CONTRACT / GRANT NUMBER: DAAL03-89-G-0118

JACKSON STATE UNIVERSITY
JACKSON, MS 39217

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED.
THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.
# Table of Contents

1. Statement of the Problem Studied  
   1.1 Image Data Models  
   1.2 Query Languages  
   1.3 The Proposed Framework for Image Retrieval  
2. Summary of the Most Important Results  
3. List of all Publications and Technical Reports  
4. List of all Participating Scientific Personnel  
5. Bibliography
1. Statement of the Problem Studied

Recently, there has been widespread interest in various kinds of database management systems for managing formatted information. Depending upon the domain and the nature of formatted data, these systems are variously referred to as Multimedia Information Systems [33, 31, 34], Spatial Databases [18, 19], Pictorial Information Systems [3, 5, 6, 8], and Image Database Systems [7, 4, 21, 17]. The application areas for these systems include but not limited to Medical Imaging [40], Medical Information Systems [26, 27], Document Image Processing and Office Information Systems [15, 16, 29], Remote Sensing and Management of Earth Resources [14, 32], Geographic Information Systems and Cartographic Modeling [23, 43], Mapping, Land Information Systems [12], Robotics [24], Interactive Computer-Aided Design (CAD) and Computer-Integrated Manufacturing Systems (CAM) [24, 11], and Image Understanding Systems [28]. Though these application areas are diverse, they all view image data as a principal resource which needs to be integratedly managed with other types of data such as voice and conventional formatted data. As diverse as the applications are, there seem to be no agreement as to what exactly is meant by the term Image Databases [7, 17, 18, 19, 21, 41].

In the early 1970s, research began in related areas such as image acquisition and registration, image processing, pattern recognition, image restoration, data structures and compression methods for image storage, techniques for image retrieval including retrieval by similarity in fairly independent directions. Only recently researchers have recognized the potential advantages in streamlining and integrating all these independent research results in building future image database systems. Tamura and Yokoya [41] provide an excellent survey of image database systems that were in practice around early 1980s. Chock [10] also provides a good survey and comparison of functionalities of several image database systems for geographic applications.

The functionality of current image database systems range from simple cataloging used for the distribution of remotely sensed imagery to image understanding and similarity measures required in medical imaging and medical information systems. However, the image data models employed in these systems are not based on any general framework. The model is rather extracted often from the implemented system and hence these data models are shaped by the idiosyncratic characteristics of the domains. Similar kind of problems plague the query language design. We provide a classification for current image database systems in section 1.1 and give a brief account on the major features and limitations for each class. Section 1.2 treats the same issues for query languages. Our proposed framework for image retrieval is presented in section 1.3.

1.1 Image Data Models

Current image data models and approaches to implementation can be classified into five different categories. The five categories are:

1. Conventional Database Systems as Image Database Systems
2. Image Processing Systems Enhanced with Advanced File System/Database Functionality
3. Building Extensions and Extensibility into the Conventional Database Systems
4. General Purpose Image Database Systems
5. Ad hoc Approaches from Application Domains

In the first category of systems, an image is fragmented and then represented as tuples in a relational table. The fragmentation process is not based on any semantic considerations and is purely based on partitioning the spatial extent of an object in terms of points, line segments, and regions. For example, a road object is approximated by a series of line segments and each line segment is represented as an independent tuple in a relation. The information that all these tuples are in fact represent the same road object can only be derived by introducing additional field(s) into the relation and then by performing the relational algebraic operation select. This fragmented representation leads to large semantic gap between the user's view of the image data and the actual view that this model provides. In addition, severe performance problems have been observed and hence this approach is not well received.

The second approach was primarily advocated by image processing and vision community. The emphasis here is on domain specific and generic low-level image representations and similarity based retrieval for template and model matching and not so much on several issues that surround the image databases. Hence, the scope and functionality of these systems are very limited.

The shortcomings of conventional relational model for image data management as noted above as well as its unsuitability to non-commercial applications has led to various extensions to it. The relational model with various extensions came to be known as Engineering Databases, Geometric Databases, and Spatial Databases. The extensions include repeating fields, procedural fields, and extensibility through abstract data types [24]. Though the data model has become more expressive now, the semantic gap between the user's view of the data and the actual view provided by system continues to exist with reduced severity especially in the query aspects of the model.

Very recently a class of spatial database systems have emerged where a clear distinction between the spatial and formatted data is maintained. The formatted processor specializes in processing the non-spatial data and the spatial processor is dedicated for manipulating the spatial data. Several variations of this architecture can be envisioned based on the coupling that exists between the formatted and the spatial processor. Usually the formatted processor is based on an extensible relational database system. Except for the availability of several spatial indexing schemes, there are no general purpose logical data models and query specification and processing techniques available to the spatial processor.

Finally, there are several image database systems that have purely evolved from very narrow application domains such as medical imaging, face and fingerprint recognition. Typically these systems are characterized by feature extraction algorithms based on domain-specific knowledge. The nature and the extent of processing and querying needs widely differ from one application area to another. As such their use as general purpose image database system is quite limited. The following section discusses similar issues that affect the query aspects of the system.
1.2 Query Languages

Powerful and flexible querying mechanisms are provided as an integral part of the conventional database management systems for formatted data so that they can be highly useful to a casual database user. Querying schemes assume even more critical prominence in systems used for image data management. Most of the existing schemes for querying image databases are variations of either Query-by-Example (QBE) [44] or Structured Query Language (SQL) [2]. We feel that this approach to querying image databases is very unsatisfactory in terms of their expressive power and also being not natural to the image data. The approaches that have been taken in the existing systems can be classified into the following three categories:

1. Extensions to the host database system query language.
2. Command Language designed to specifically suit the application requirements.
3. Logic-based query languages.

The extended query language approach is typically found in majority of the systems which are recently built on top of an existing conventional database system. It is quite natural to take advantage of host database system for querying the image data also. If the images are converted to symbolic representations and stored in relations, the only extension needed to the query language is display management for image output. On the other hand, in systems where image data is stored in a separate database with spatial indexing, if the query language allows query specification in pictorial form, the extensions needed may not be trivial. In general, a query may reference only the image data, only the non-image data, or a combination of image and non-image data. An ideal query language should provide a consistent user interface for both image and non-image data and be able to command and coordinate both the query processor for the formatted data and the spatial query processor in a transparent way to the user. This approach is taken in PSQL [37] and in PICQUERY [22] to some degree. PSQL is an extension of SQL and PICQUERY has a flavor similar to QBE and QPE [4].

The KBGIS-II system uses a logic-based query language called spatial object language (SOL) [30]. The query language also allows for spatial constraint specification. However, its practicability remains to be seen in large spatial databases given the exorbitant computational requirement that usually surrounds many logic based query languages. Chang et al. [9] proposed a 2D G-string based query language. This query language processes queries using string matching and spatial reasoning. Both the expressive power and the naturalness of these languages for specifying spatial queries is not established.

User Interface requirements should be carefully evaluated and incorporated into the design of pictorial query languages. Eigenhofer and Frank [13] discuss user interface considerations in designing a pictorial/spatial query language. The query languages designed for conventional formatted databases are clearly unsuitable for specifying pictorial queries. There are several types of spatial/image queries and each type may require a potentially different method of specification. For example, a user may wish to indicate an area of interest on a map to search for a specific feature using an interactive input device such as a mouse. Users specify point queries to obtain information on all the
objects that occupy a specified location in space. Region queries are used to obtain information on all the objects that exist in space enclosed by a hyper rectangle called query window. A query window can be conveniently specified in two dimensions by indicating two points using a pointing device.

The point and region queries can be combined with SQL queries to increase the expressive power of a query language. For example, a query can be specified to select only those objects within the query window that satisfy given non-spatial predicates. This type of query specification is highly suitable for applications dealing with geographic data. However, for querying image databases of electronic product catalogues, browsing techniques are effective.

In some applications involving retrieval based on similarity, iconic query interface may be essential. The user specifies a query by placing icons designating real objects in the domain at certain desired locations in a query window and then assigns attribute properties to each of these icons. By doing so, the user is able to specify the spatial relationships among domain objects as well as their attribute values in the query. Retrieval based on conceptual similarity [36] between images may require yet another method of query specification. If the response to a query involves the display of some images, the query language must also provide for the specification of an appropriate output device of user's choice. These problems compound and make the design of a pictorial query language a difficult task. Many of the proposed approaches to querying image databases address only one aspect this complex task mostly driven by specific application needs. In essence, it may not be unreasonable to associate a different method of query specification with different classes of queries but under a consistent user interface. We describe our unified framework for retrieval in general purpose image database systems in the next section.

3. The Proposed Framework for Image Retrieval

The discussion in the previous section suggests that a general purpose image database system should provide several classes of retrieval needs within a unified framework and it should involve a consistent user interface. The application areas should be able to choose only those functionalities that are both natural and useful to the domain, much along the lines of EXODUS object oriented database generator. An image represented as an array of pixels (raster format) or as a collection of line segments (vector format) is considered to be at physical level representation. Since physical level representation provides very little information on the image contents without extensive image processing and understanding, it is desirable for an image database model to be able to provide multiple logical representations to facilitate interactive image retrieval. Logical representations denote abstractions of the image at various desired levels and are derived only once when an image is added to the database.

Most of the commercial systems operate at the physical level representation and build ad hoc logical representations for answering certain types of queries. These ad hoc logical representations vanish as soon as a query is processed and the whole process starts all over again when a similar query arrives next time. To avoid the exorbitant computational cost involved in building these ad hoc logical representations repeatedly
some systems precompute and store important results which can be derived from such logical representations. However, it should be noted that this is not a solution since we cannot in general anticipate the precise nature of queries and even if were to anticipate, simply it would be too voluminous and uneconomical to explicitly store all such precomputed data. Hence for practical and large spatial databases, multiple logical representations are not only necessary but also required to meet performance requirements for interactive processing [46].

Each logical representation in this hierarchy can be viewed as an ideal representation for efficiently processing one or more classes of queries. Moreover, these logical representations can be very useful in restricting access to the image data by assigning user access privileges to one or more layers in the hierarchy. In essence, this representation can also be regarded as a view defining mechanism for image databases.

Now, we have, at one end of the spectrum, the physical layer at bit-level representation of the image. At the other end of the spectrum, we have the logical image which is an extremely abstracted version of the physical image. In between, we can conceive several logical layers corresponding to varying degrees of abstraction. The layers at lower levels embody more accurate representations of the image than the layers at the higher levels which provide a course representation by suppressing several insignificant details. By studying the application requirements and the limitations of the proposed approaches, we envision a multi-layered structure for retrieval. The various layers in the scheme are:

1. Physical Layer
2. Spatial and Shape Layer
3. Iconic and Attribute Layer
4. Conceptual Layer

These layers are not designed to operate in isolation but rather work in cooperation. To avoid redundancy in representation the layers are structured to form a lattice. The layers can also be viewed as multiple representations for the same object. A detailed description of this framework can be found in [46].

2. Summary of the Most Important Results

The major accomplishments of the project are listed below.

1. Developed algorithms for retrieval in image databases based on spatial similarity.

2. Developed three sets of test bed of images from three different domains for testing spatial similarity algorithms and for discovering domain concepts using Personal Construct Theory (PCT). The domain concepts are expected to provide the basis for concept-based retrieval of images.

3. Developed a novel spatial query specification technique, called the iconic query interface, for specifying certain classes of spatial queries.
4. A prototype system, using the iconic query interface, to demonstrate the effectiveness of the spatial similarity algorithms on the test bed of images has been developed.

5. A novel method for discovering domain concepts by the application of Personal Construct Theory has been devised. This method has been successfully demonstrated on two test beds of images.

6. The design of a comprehensive multi-layered object-based framework for retrieval in image databases is near completion. The framework provides a flexible data model and query specification techniques suitable to several classes of image queries.

3. List of All Publications and Technical Reports


4. List of All Participating Scientific Personnel

V. Gudivada is the Principal Investigator at Jackson State University (JSU) and V. Raghavan is the Co-Principal Investigator at the University of SW Louisiana (USL). The Principal Investigator was provided 25% release time during regular academic semesters and two months of full support during summer semesters. The Co-Principal Investigator was supported one month during summer semesters.

D. Carr, V. Tummalapally, V. Griddalur, X. Bao, B. Panda are the students supported at JSU. Except X. Bao, all other students obtained their masters degrees in Computer Science. D. Carr was supported throughout the project period, V. Griddalur and X. Bao were supported through substantial part of the project period, and B. Panda was supported though a summer semester. They contributed to results on the creation of the
two image databases, the design of a retrieval function for retrieving images by spatial similarity, and the development of iconic query interface.

V. Elayavalli, Y. Zhang, S. Mishra, J. Alsabbagh, J. Bhuyan, S. Sridharan and G. Jung are the students at USL who received partial/full support to varying periods of time. Y. Zhang was supported for two semesters; J. Alsabbagh and V. Elayavalli was supported one semester with full support and another semester with partial support. S. Sridharan was supported through one regular semester and J. Bhuyan and G. Jung were supported through one summer semester. S. Mishra and J. Alsabbagh received partial support for short periods of time. These students contributed in setting up the PC-based software environment, eliciting concepts from images using PCT, providing analysis tools for repertory grids, and establishing analogies between text and image domains.

5. Bibliography


46. Gudivada, V.N. and Raghavan, V.V. (1991), An Iconic Query Interface for an Image Database, Technical Report: 91-2, Department of Computer Science, Jackson State University, Jackson, MS.