01-F1

Shu-Ching Chen, Mei-Ling Shyu, Xia Jin, Qiong Chen, Chengcui Zhang, and Jeff Strickrott, "A Flexible Image Retrieval and Multimedia Presentation Management System for Multimedia Databases," ACM Multimedia 2001 conference, pp. 601-602, September 30 - October 5, 2001, Ottawa, CANADA.

A Flexible Image Retrieval and Multimedia Presentation Management System for Multimedia Databases

Shu-Ching Chen[1*], Mei-Ling Shyu[2], Xia Jin[1], Qiong Chen[1], Chengcui Zhang[1], Jeff Strickrott[1]

[1]School of Computer Sciences, Florida International University,
Miami, FL, 33199, USA

[2]Department of Electrical & Computer Engineering, University of Miami,
Coral Gables, FL, 33124, USA

[1]{chens, xjin01, qchen01, czhang02, jstric01}@cs.fiu.edu, [2]shyu@miami.edu

ABSTRACT

In today's fast-growing information age, multimedia data is becoming more and more common in daily applications, however this data is nearly useless if there is no computer-aided browsing, searching, and retrieving mechanism to obtain the desired contents. In this demonstration, we present a multimedia query and presentation management system for multimedia databases. It uses an unsupervised segmentation method for image and video feature extraction. A presentation design interface systems is provided that allows query results to be used to create multimedia presentations.

1. INTRODUCTION

An efficient multimedia database management system (MDBMS) must:

- Have the capability to query data including both media data and textual data represented in different formats.
- Have the capability to take the answer generated by a query and develop a presentation of that answer in terms of audio-visual media.
- Have the capability to deliver this presentation in a way that satisfies various Quality-of-Service requirements.
- Have the capability to handle time-dependency and synchronized presentation of multimedia data.

For the query capability, since images are not only the most widely used multimedia type, but also the base for other multimedia types such as video, the need for efficient content-based image retrieval has become necessary to accomplish the query functionality of the MDBMS. Besides query capability, the development of abstract semantic models is also essential for a robust MDBMS. Much research has been done in the area of modeling the temporal and/or spatial relations among multimedia objects. The model should be powerful enough to support multimedia presentation synchronization and there should be a good programming data structure for the implementation to handle multimedia presentations. The multimedia augmented transition network (MATN) is one such model [2]; a semantic model that can model multimedia presentations, the temporal and/or spatial relations of different media streams, and multimedia browsing.

Our system is a client/server architecture and includes two modules: 1) content-based image retrieval module and 2) MATN module. The following sections will describe the architecture and technical details about the proposed system.

2. SYSTEM OVERVIEW

In our system, a novel unsupervised segmentation method called simultaneous partition and class parameter estimation (SPCPE) [3] is applied for extracting the spatial features of image contents. The client is comprised of two modules: the image retrieval module and the MATN module. The image retrieval module provides a query interface, while the MATN module provides an MATN interface. The client also has a mechanism that uniquely combines multimedia queries and multimedia presentations.

On the server side: The novel SPCPE algorithm is adopted to develop an efficient image segmentation method. The proposed image segmentation method achieves the unsupervised identification of perceptually salient regions in images. The algorithm obtains an optimal class partition of image pixels by minimizing a probability function derived from Bayes' theorem. The algorithm starts with a semi-random initial partition and refines the partition recursively based on computed class parameters obtained in the previous iteration. Therefore, the class partition and the computation of class parameters are carried out simultaneously and they form the basis for the computation of each other

On the client side:

- The major functionality of the image retrieval module is to carry out query-by-image, which includes image loading, image query, image browsing, and the connection with the multimedia database server. The image retrieval module is also designed to connect with the MATN module for providing the input data for multimedia presentation design.
- The MATN module carries out the MATN model composition and multimedia presentations. Structuring and modeling multimedia data is very important in the design of a multimedia information system. The MATN model fully captures the two-way communication between users and the

*This research was supported in part by NSF CDA-9711582.

system so that they can compose their desired multimedia presentations flexibly. In order to visually present the multimedia data, a JMF based media player and a image/text displayer are integrated into the client side.

For the client and server communication:
In order to meet the requirement of processing the distributed multimedia data, the system is built upon a client/server architecture. A transfer protocol based on UDP protocol is developed to connect the client and the server to transmit the GIF and JPEG images, as well as the MPEG video files. Each packet consists of header information and a segment of the image data. To visually display an image sent from the server, the client has to collect all the necessary image data and then restore them into an image again.

3. SYSTEM ARCHITECTURE

This system is based on a two-tier client-server architecture in which the user interface runs on the client and the database is stored on the server. It is composed of a Java client and a database server implemented in C++. The system supports both Windows and Unix platforms. The client initiates the queries and sends the queries to the database server. After receiving the queries, the server searches and retrieves the relevant results, and returns the query results to the client. Then the client displays all the query results.

3.1 The Client Tier

The client of the system plays two important roles. One role is to interact with users and the other role is to communicate with the server. Since the client is the front-end of the system, graphical user interfaces that are easy-to-use, user-friendly, and fully functional must be developed. In addition, the client must be structured in the way that future updates can be easily done.

The Client is broken down into three modules: the query module, the MATN module and the multimedia presentation module (Figure 1 depicts these). The query module provides a query interface for users to query by image and serves as the front-end of the content-based retrieval. Nine images can be displayed simultaneously; one image can be selected at a time as a query image. The query environment provides input information to the MATN environment for presentation construction. The MATN module provides a set of tools for automatically generating a MATN model and creating a multimedia presentation environment. A MATN model consists of a set of states represented by filled circles, a set of arcs, and a set of arc labels (multimedia input strings [2]). Multimedia presentations directly involve the MATN model; users are allowed to select any part of the presentation to display by selecting a starting state and an ending state in the model. The multimedia presentation module implements the interface with the user. Appropriate windows, text, image and audio/video are generated based upon the arc labels of the MATN.

3.2 The Database Server Tier

The task of a multimedia database server is to build up a system that can manage huge collections of multimedia data in an efficient way. The main function is to serve multimedia queries. Instead of manually and subjectively associating keywords to each image stored in the database, our system is able to conduct unsupervised indexing based on similarity features. The queries are answered by comparing the feature sets of those images to that of the query image. The main functions of the database server in this implementation are image class partitioning, image segment tracking, and content-based image retrieval. Images segmentation (feature extraction) is executed offline and is stored in the database before run time.

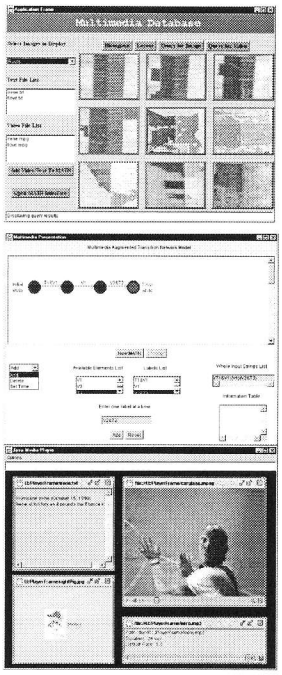

Figure 1: Image query interface (top), MATN interface, and media player window

The content-based image retrieval process implemented in this system is to query the images that are similar to the query example image. The parameters captured for each class in the class partition step are stored in the database and are used as indexes for the images. At the query time, a multi-filter query strategy is introduced [1] for content-based retrieval. Several stored parameters that are obtained during the segmentation process are used for database searching. By comparing those parameters, similar images can be retrieved. The parameters are calculated based on the mathematical and probability functions of the spatial coordinates of pixels so they can represent the spatial distribution of the pixels in the image.

4. CONCLUSIONS

Our research focuses on the design of an MDBMS that can efficiently organize, store, manage, and retrieve multimedia information from the underlying multimedia database, as well as the development of an abstract semantic model (MATN) that is capable of effectively modeling the varieties of multimedia data in terms of their structure, behavior, and function, and supporting multimedia presentations. A content-based image retrieval framework for a multimedia database server is implemented. The MATN design and presentation module provided by our system allows users to compose their desired MATN models for multimedia presentations. Our system adopts a client/server architecture for a distributed environment.

5. REFERENCES

[1] S.-C. Chen, X. Jin, and Q. Chen, "Client and Server Design for Content-Based Retrieval in Multimedia Database Systems," Technical Report 2001-04, School of Computer Science, Florida International University, April 2001.

[2] S.-C. Chen and R. L. Kashyap, "A Spatio-Temporal Semantic Model for Multimedia Database Systems and Multimedia Information Systems," to appear in *IEEE Trans. on Knowledge and Data Engineering*.

[3] S.-C. Chen, S. Sista, M.-L. Shyu, and R. L. Kashyap, "An Indexing and Searching Structure for Multimedia Database Systems," *IS&T/SPIE conference on Storage and Retrieval for Media Databases 2000*, pp. 262-270, January 23-28, 2000, San Jose, CA, U.S.A.

01-FI

A Flexible Image Retrieval and Multimedia Presentation Management System for Multimedia Databases

Shu-Ching Chen[1*], Mei-Ling Shyu[2], Xia Jin[1], Qiong Chen[1], Chengcui Zhang[1], Jeff Strickrott[1]

[1]School of Computer Sciences, Florida International University,
Miami, FL, 33199, USA

[2]Department of Electrical & Computer Engineering, University of Miami,
Coral Gables, FL, 33124, USA

[1]{chens, xjin01, qchen01, czhang02, jstric01}@cs.fiu.edu, [2]shyu@miami.edu

ABSTRACT

In today's fast-growing information age, multimedia data is becoming more and more common in daily applications, however this data is nearly useless if there is no computer-aided browsing, searching, and retrieving mechanism to obtain the desired contents. In this demonstration, we present a multimedia query and presentation management system for multimedia databases. It uses an unsupervised segmentation method for image and video feature extraction. A presentation design interface systems is provided that allows query results to be used to create multimedia presentations.

1. INTRODUCTION

An efficient multimedia database management system (MDBMS) must:

- Have the capability to query data including both media data and textual data represented in different formats.
- Have the capability to take the answer generated by a query and develop a presentation of that answer in terms of audio-visual media.
- Have the capability to deliver this presentation in a way that satisfies various Quality-of-Service requirements.
- Have the capability to handle time-dependency and synchronized presentation of multimedia data.

For the query capability, since images are not only the most widely used multimedia type, but also the base for other multimedia types such as video, the need for efficient content-based image retrieval has become necessary to accomplish the query functionality of the MDBMS. Besides query capability, the development of abstract semantic models is also essential for a robust MDBMS. Much research has been done in the area of modeling the temporal and/or spatial relations among multimedia objects. The model should be powerful enough to support multimedia presentation synchronization and there should be a good programming data structure for the implementation to handle multimedia presentations. The multimedia augmented transition network (MATN) is one such model [2]; a semantic model that can model multimedia presentations, the temporal and/or spatial relations of different media streams, and multimedia browsing.

Our system is a client/server architecture and includes two modules: 1) content-based image retrieval module and 2) MATN module. The following sections will describe the architecture and technical details about the proposed system.

2. SYSTEM OVERVIEW

In our system, a novel unsupervised segmentation method called simultaneous partition and class parameter estimation (SPCPE) [3] is applied for extracting the spatial features of image contents. The client is comprised of two modules: the image retrieval module and the MATN module. The image retrieval module provides a query interface, while the MATN module provides an MATN interface. The client also has a mechanism that uniquely combines multimedia queries and multimedia presentations.

On the server side: The novel SPCPE algorithm is adopted to develop an efficient image segmentation method. The proposed image segmentation method achieves the unsupervised identification of perceptually salient regions in images. The algorithm obtains an optimal class partition of image pixels by minimizing a probability function derived from Bayes' theorem. The algorithm starts with a semi-random initial partition and refines the partition recursively based on computed class parameters obtained in the previous iteration. Therefore, the class partition and the computation of class parameters are carried out simultaneously and they form the basis for the computation of each other

On the client side:

- The major functionality of the image retrieval module is to carry out query-by-image, which includes image loading, image query, image browsing, and the connection with the multimedia database server. The image retrieval module is also designed to connect with the MATN module for providing the input data for multimedia presentation design.
- The MATN module carries out the MATN model composition and multimedia presentations. Structuring and modeling multimedia data is very important in the design of a multimedia information system. The MATN model fully captures the two-way communication between users and the

*This research was supported in part by NSF CDA-9711582.

system so that they can compose their desired multimedia presentations flexibly. In order to visually present the multimedia data, a JMF based media player and a image/text displayer are integrated into the client side.

For the client and server communication:
In order to meet the requirement of processing the distributed multimedia data, the system is built upon a client/server architecture. A transfer protocol based on UDP protocol is developed to connect the client and the server to transmit the GIF and JPEG images, as well as the MPEG video files. Each packet consists of header information and a segment of the image data. To visually display an image sent from the server, the client has to collect all the necessary image data and then restore them into an image again.

3. SYSTEM ARCHITECTURE

This system is based on a two-tier client-server architecture in which the user interface runs on the client and the database is stored on the server. It is composed of a Java client and a database server implemented in C++. The system supports both Windows and Unix platforms. The client initiates the queries and sends the queries to the database server. After receiving the queries, the server searches and retrieves the relevant results, and returns the query results to the client. Then the client displays all the query results.

3.1 The Client Tier

The client of the system plays two important roles. One role is to interact with users and the other role is to communicate with the server. Since the client is the front-end of the system, graphical user interfaces that are easy-to-use, user-friendly, and fully functional must be developed. In addition, the client must be structured in the way that future updates can be easily done.

The Client is broken down into three modules: the query module, the MATN module and the multimedia presentation module (Figure 1 depicts these). The query module provides a query interface for users to query by image and serves as the front-end of the content-based retrieval. Nine images can be displayed simultaneously; one image can be selected at a time as a query image. The query environment provides input information to the MATN environment for presentation construction. The MATN module provides a set of tools for automatically generating a MATN model and creating a multimedia presentation environment. A MATN model consists of a set of states represented by filled circles, a set of arcs, and a set of arc labels (multimedia input strings [2]). Multimedia presentations directly involve the MATN model; users are allowed to select any part of the presentation to display by selecting a starting state and an ending state in the model. The multimedia presentation module implements the interface with the user. Appropriate windows, text, image and audio/video are generated based upon the arc labels of the MATN.

3.2 The Database Server Tier

The task of a multimedia database server is to build up a system that can manage huge collections of multimedia data in an efficient way. The main function is to serve multimedia queries. Instead of manually and subjectively associating keywords to each image stored in the database, our system is able to conduct unsupervised indexing based on similarity features. The queries are answered by comparing the feature sets of those images to that of the query image. The main functions of the database server in this implementation are image class partitioning, image segment tracking, and content-based image retrieval. Images segmentation (feature extraction) is executed offline and is stored in the database before run time.

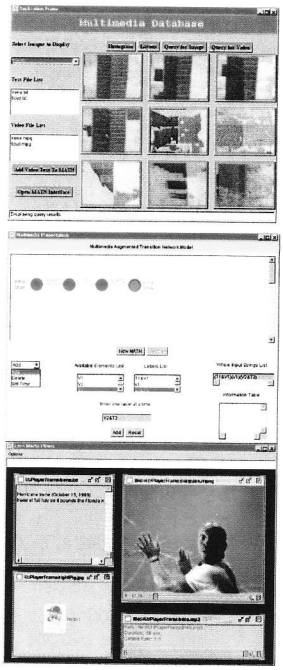

Figure 1: Image query interface (top), MATN interface, and media player window

The content-based image retrieval process implemented in this system is to query the images that are similar to the query example image. The parameters captured for each class in the class partition step are stored in the database and are used as indexes for the images. At the query time, a multi-filter query strategy is introduced [1] for content-based retrieval. Several stored parameters that are obtained during the segmentation process are used for database searching. By comparing those parameters, similar images can be retrieved. The parameters are calculated based on the mathematical and probability functions of the spatial coordinates of pixels so they can represent the spatial distribution of the pixels in the image.

4. CONCLUSIONS

Our research focuses on the design of an MDBMS that can efficiently organize, store, manage, and retrieve multimedia information from the underlying multimedia database, as well as the development of an abstract semantic model (MATN) that is capable of effectively modeling the varieties of multimedia data in terms of their structure, behavior, and function, and supporting multimedia presentations. A content-based image retrieval framework for a multimedia database server is implemented. The MATN design and presentation module provided by our system allows users to compose their desired MATN models for multimedia presentations. Our system adopts a client/server architecture for a distributed environment.

5. REFERENCES

[1] S.-C. Chen, X. Jin, and Q. Chen, "Client and Server Design for Content-Based Retrieval in Multimedia Database Systems," Technical Report 2001-04, School of Computer Science, Florida International University, April 2001.

[2] S.-C. Chen and R. L. Kashyap, "A Spatio-Temporal Semantic Model for Multimedia Database Systems and Multimedia Information Systems," to appear in *IEEE Trans. on Knowledge and Data Engineering*.

[3] S.-C. Chen, S. Sista, M.-L. Shyu, and R. L. Kashyap, "An Indexing and Searching Structure for Multimedia Database Systems," *IS&T/SPIE conference on Storage and Retrieval for Media Databases 2000*, pp. 262-270, January 23-28, 2000, San Jose, CA, U.S.A.