An Affinity-Based Image Retrieval System for Multimedia Authoring and Presentation

Shu-Ching Chen[1], Mei-Ling Shyu[2], Na Zhao[1], Chengcui Zhang[1]
[1]Distributed Multimedia Information System Laboratory, School of Computer Science
Florida International University, Miami, FL 33199, USA
[2]Department of Electrical & Computer Engineering
University of Miami, Coral Gables, FL, 33124, USA
[1]{chens, nzhao002, czhang02}@cs.fiu.edu, [2]shyu@miami.edu

ABSTRACT
In this demonstration, we present an image retrieval system to support multimedia authoring and presentation. An affinity-based mechanism, Markov Model Mediator (MMM), is used as the search engine for the system, which utilizes both the low-level image features and the learned high-level concepts via user access patterns and access frequencies. This system is one of the major components of MediaManager, a distributed multimedia management system developed by us. Both retrieval and learning facilities are supported in this system. The retrieval system also provides input information to the Multimedia Augmented Transition Network (MATN) environment for multimedia authoring and presentation.

Categories and Subject Descriptors
H.3.3 [**Information Systems**]: Information Storage and Retrieval – Information Search and Retrieval – *Retrieval models*.

General Terms: Experimentation, Human Factors.

Keywords: Content-based Image Retrieval, Markov Model Mediator, MediaManager.

1. INTRODUCTION
In this demonstration, the *Markov Model Mediator* (MMM) mechanism [3] is employed as the retrieval engine in our content-based image retrieval system. This mechanism not only considers the global image features, but also exploits the rich semantics contained in the user access patterns and access frequencies. This system is implemented as one of the major components of MediaManager, a prototype distributed multimedia management system developed by us. MediaManager not only utilizes the MMM mechanism in the common retrieval processing, but also provides the functionalities to learn user preferences. The demonstrated image retrieval and learning system is different from the common relevance feedback [2] CBIR systems in the following aspects:

- *Mining relative affinity between images*: Unlike the common relevance feedback which tries to refine the query results for one single image through several feedback iterations, in MMM, the user issued queries may have different query images. In other words, the MMM mechanism is able to mine both inter- and inner-query image affinities from the training data set. After a certain volume of trainings, the relative affinity relationships among the images can be obtained, which means that the more concurrently two images are accessed together, the more closely they are associated. Accordingly, the query results of most of the images in the database can be improved dramatically.
- *User-Friendliness*: In the MMM-based system, since the affinity between images can be discovered via a set of queries (they may involve different query images) issued by multiple users, there is no heavy burden for one single user to train the system in real-time via many iterations. In this system, the user has the choices to train the system either off-line or on-line. The calculated affinity matrix serves as the indications of users' high-level concepts in our system. After learning user semantic concepts, the well-trained system can have much better performance in image retrieval.
- *Robustness*: In relevance feedback, the user needs to take heavy responsibilities for the correctness of his/her feedbacks. Any mislabeled images will significantly damage the system's performance. On the other hand, since we consider the accumulative responses from multiple users in this system, small amount of errors from individual users can be compensated in most cases.

Further, since the trained CBIR subsystem can provide users with their desired images more accurately and more quickly, it will greatly facilitate the design of multimedia authoring and presentation by using the retrieved images. In MediaManager, a multimedia presentation authoring tool based on the Multimedia Augmented Transition Network (MATN) model [1] is developed.

2. SYSTEM OVERVIEW

2.1 Overall Architecture
With a multi-threaded client/server architecture employed, this system can support multiple clients to issue queries and offer feedbacks simultaneously. In the server side, a database engine is implemented to support image query processing, file supply, training computations, feature extraction and indexing of images. There are 10,000 color images as well as their feature set stored in our database. The client side provides a content-based image retrieval user interface, which allows the browsing, query and feedback of the image contents.

As shown in Figure 1, the overall architecture of MediaManager system is composed of 3 major subsystems, namely the image retrieval system, the video capture/indexing/retrieval system, and the multimedia presentation design/rendering system. In the MMM mechanism, the initial queries are issued based on the low level features (colors, object locations, etc.). The user can provide feedbacks for the query results of different query images. As the

MM'03, November 2-8, 2003, Berkeley, California, USA.
ACM 1-58113-722-2/03/0011.

indications for users' perceptions, the access patterns and access frequencies can be gathered and stored in the server side, and to be utilized in the training process. The relative affinity relationships are then refined via the MMM mechanism. Finally, both the low-level features and the captured user perceptions are used to refine the image retrieval system.

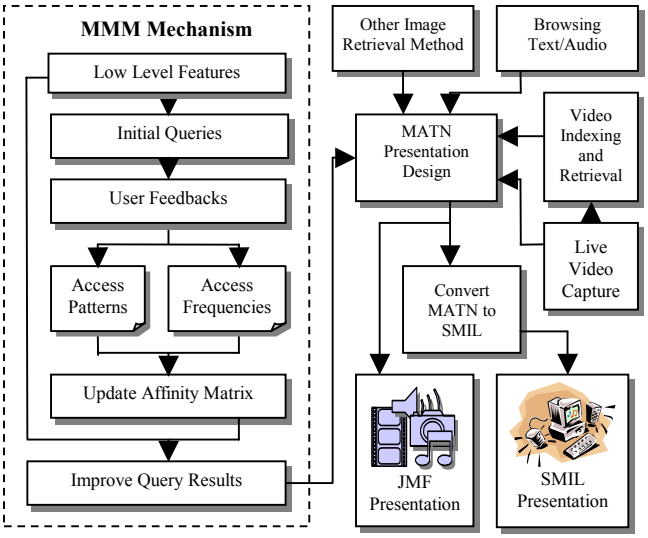

Figure 1. MMM Mechanism and its role in the MediaManager System

2.2 Retrieval Results

Different people may have different concerns when they look for relevant images. Taking the query image in Figure 2 as an example, some people may want to find the parade scene, some may want to find a horse, and some other people may think the human object looks more appealing to their interest. Hence, it is important that more users should be involved into the training process because fewer users may have bias on the query results. Currently, after 1,500 queries with feedbacks have been issued, 5,157 distinct images have been covered.

Figure 2. Improved query results by using the trained system

As shown in Figure 2, the query results are improved dramatically after training since most of the images being retrieved contain the parade scenes. Based on our experiments over randomly selected 80 query images with 5 categories, the overall accuracy within the top 20 images is 88%. We also found that with more images covered in training, more improved semantic retrieval accuracy can be obtained, and the system performance will be further boosted with accumulative learning. In addition, the average retrieval time for the current system is around 3 seconds.

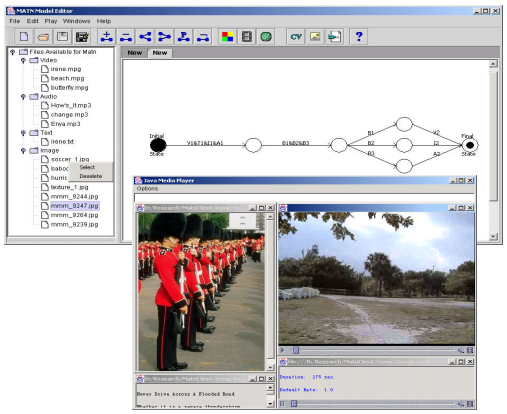

Figure 3. Multimedia presentation authoring tool

2.3 Multimedia Presentation Authoring

Another achievement in this implementation is to connect the MMM-based CBIR system with the multimedia presentation authoring tool. This module, as shown in Figure 3, is implemented to facilitate the design of multimedia presentations by using our MATN model [1]. After the user finds out the desired images via the image retrieval system, he/she may use the "Download" function to add them into the presentation material tree, which can be employed later in multimedia presentation designs. Other than rendering the presentation using JMF, another option is provided by the system to view it via web browser, in which the MATN model can be automatically translated to SMIL scripts. Basically, our CBIR system can help the users find their interested images more accurately and more quickly such that multimedia presentation design becomes much easier.

3. ACKNOWLEDGMENTS

For Shu-Ching Chen, this research was supported in part by NSF EIA-0220562 and the office of the Provost/FIU Foundation.

4. REFERENCES

[1] Chen, S.-C., Kashyap, R.L., and Ghafoor, A. Semantic Models for Multimedia Database Searching and Browsing, Kluwer, 2000.

[2] Rui, Y., Huang, T.S., and Mehrotra, S. Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval, *IEEE Transactions on Circuit and Video Technology*, Special Issue on Segmentation, Description, and Retrieval of Video Content, 18(5), pp. 644-655, 1998.

[3] Shyu, M.-L., Chen, S.-C., and Zhang, C. A Stochastic Content-Based Image Retrieval Mechanism, accepted for publication, Multimedia Systems and Content-based Image Retrieval, Edited by Sagarmay Deb, Idea Group Inc., publisher of the Idea Group Publishing, Information Science Publishing and IRM Press imprints.