The Second ACM International Workshop on Multimedia Databases (MMDB 2004) held at ACM CIKM 2004

Shu-Ching Chen
School of Computer Science
Florida International University
Miami, FL 33199, USA
chens@cs.fiu.edu

Mei-Ling Shyu
Department of Electrical & Computer Engineering
University of Miami
Coral Gables, FL 33124, USA
shyu@miami.edu

The Second ACM International Workshop on Multimedia Databases was held in Washington, DC, USA, November 13, 2004. Its aim is to bring together university researchers, scientists, industry professionals, software engineers and graduate students who need to become acquainted with new theories and technologies in multimedia databases, and to all those who wish to gain a detailed technical understanding of what multimedia databases involve.

We received thirty-one submissions. Each of them was reviewed by at least three referees. The program committee, which was constituted of researchers and practitioners from the multimedia database research areas, selected twelve papers to be presented at the workshop. The authors were asked to address each and every comment made by the referees to improve the quality of the papers. These papers reflect current trends in the multimedia database research areas with topics such as Content-Based Retrieval for Multimedia Databases, Multimedia Database Query Processing and Retrieval, Multimedia Data Indexing, and Multimedia Data Mining. From these twelve papers, the workshop offered three sessions for participants representing a variety of backgrounds and perspectives on multimedia databases.

In "Content-Based Sub-Image Retrieval Using Relevance Feedback," Jie Luo and Mario A. Nascimento presented their approach that uses relevance feedback in content-based sub-image retrieval (CBsIR) to improve the accuracy of successive retrievals. A tile re-weighting scheme assigns penalties to each tile of the database images and updates the tile penalties for all relevant images retrieved at each iteration using both the relevant (positive) and irrelevant (negative) images identified by the user. They evaluated the performance on a dataset of over 10,000 images by using 64 quantized colors in the RGB color space, their system can achieve a stable average recall value of 70% within the top 20 retrieved (and presented) images after only 5 iterations, where each iteration was about 2 seconds.

Liang-Tien Chia presented on behalf of Haoran Yi and Deepu Rajan on "A Motion based Scene Tree for Browsing and Retrieval of Compressed Videos." In this work, a fully automatic content-based approach for browsing and retrieval of MPEG-2 compressed video was proposed. The first step of their approach is shot boundary detection based on motion vectors available from the compressed video stream. The second step is to build a scene tree from the shots obtained earlier. The scene tree is shown to capture some

semantic information as well as to provide a construct for hierarchical browsing of compressed videos. Finally, a new model for video similarity based on global as well as local motion associated with each node in the scene tree was constructed via new approaches on camera motion and object motion estimation. Their experimental results demonstrated that the integration of the above techniques results in an efficient framework for browsing and searching large video databases.

Mei-Ling Shyu, Shu-Ching Chen, Min Chen, and Chengcui Zhang addressed the need to search and retrieve images efficiently and accurately from a large image database or a collection of image databases in their paper entitled "A Unified Framework for Image Database Clustering and Content-based Retrieval." A unified framework called Markov Model Mediators (MMMs) was proposed to facilitate database clustering and improve the query processing performance by analyzing the summarized knowledge. The unique characteristic of MMMs is that it provides the capabilities of exploring the affinity relations among the images and databases at the intra-database level and inter-database level using an effective data mining process. In this framework, at the intra-database level, each database is modeled by a local MMM which enables accurate image retrieval within the database; while at the inter-database level, conceptual database clustering is performed based on these local MMMs to group those related databases into a conceptual image database cluster, represented by an integrated MMM. The framework was tested using a set of image databases, which contain various numbers of images with different dimensions and concept categories, and the experimental results demonstrated that accurate retrieval results with low query processing cost can be achieved.

Marat Fayzullin, V.S. Subrahmanian, Massimiliano Albanese, and Antonio Picariello discussed the concept of a priority curve associated with a video in "The Priority Curve Algorithm for Video Summarization." They also developed an algorithm that can use the priority curve to create a summary (of a desired length) of any video. The created summary exhibits nice continuity properties and also avoids repetitions. The priority curve algorithm (PCA) was implemented and compared with other summarization algorithms in the literature. It is shown that PCA is faster than the existing algorithms and also produces better quality summaries. The quality of summaries was evaluated by a group of 200 students in Naples, Italy, who watched soccer videos.

In "Semantic Retrieval of Multimedia Data," Samira Hammiche, Salima Benbernou, Mohand-Said Hacid, and Athena Vakali proposed an approach that deals with the problem of finding multimedia data that fulfills the requirements of user queries. It assumes both the user query and the multimedia data are expressed by the MPEG-7 standard. However, the MPEG-7 formalism lacks the semantics and reasoning support in many ways, such as the search of the implicit data cannot be achieved since its description is based on XML schema. Therefore, they proposed a framework for querying multimedia data based on a tree embedding algorithm, combining the MPEG-7 standard and an external ontology.

"Complementary Information Retrieval for Cross-Media News Content" by Qiang Ma, Akiyo Nadamoto, and Katsumi Tanaka discussed a new way of integrating cross-media news content such as television programs and web pages. Cross-media news content was searched to find complementary items which can provide additional information to the users who are interested in a particular topic. The complementary news items not only are similar to the item the user is interested in, but also provide information in more detail or from a different perspective. A novel content-representation model called the "topic structure"

model was developed. Intuitively, a topic structure is made up of a pair of subject and content terms. Subject terms denote the dominant terms of a news item. A content term is a term having strong co-occurrence relationships with the subject terms. Based on the topic structure, the information related to a given news item (e.g., one in which the user is interested) from content, context, and media complementation perspectives can be searched. They also described an application system which concurrently presents a television news program along with complementary news articles to help the users understand news topics in greater detail and from multiple perspectives.

MPEG-7 documents, primarily for multimedia information exchange, are data-centric XML documents. Due to its advantages, the relational DBMS is the best choice for storing such XML documents. Storing XML data in relational DBMS can be classified into two classes of storage model: structure-mapping and model-mapping. However, the structure-mapping model cannot support complex Xpath-based query efficiently and model mapping approach lacks the flexible capability in representing all kinds of datatypes. To address this issue, Yang Chu, Liang-Tien Chia and Sourav Saha Bhowmick presented a new storage approach, called SM3, in "Looking at Mapping, Indexing & Querying of MPEG-7 Descriptors in RDBMS with SM3." As an XML document, MPEG-7 document can be viewed as XML tree. Such a tree graph, where the internal nodes are element types with element contents, represents the structure of documents and can be viewed as nodes which are meaningful only for document traversal. In their approach, the leaf node, which is a single-valued attribute or element type with text content, has little usage for XML tree routing as it is the end-point of Xpath. So it can be viewed as the special node which only holds the value. In this paper, SM3 was designed to use a model-mapping approach to store all internal nodes and a structure-mapping model to store all leaf nodes. SM3 integrates the advantages of those two models and hopefully avoids the main drawbacks from each method. In performance evaluation, SM3 with a pure model-mapping method (XParent) and SM3 with a pure structure-mapping method (XML-DBMS) were compared, and the initial experimental results are encouraging.

Multivariate time series (MTS) datasets are common in various multimedia, medical and financial applications. Kiyoung Yang and Cyrus Shahabi proposed a similarity measure for MTS datasets in their "A PCA-based Similarity Measure for Multivariate Time Series" paper. Eros (Extended Frobenius norm) applies Principal Component Analysis (PCA) to MTS datasets represented as matrices to generate principal components and associated eigenvalues. These principal components and eigenvalues are then used to compare the similarity between MTS matrices. Experiments that compared Eros with Euclidean Distance (ED), Dynamic Time Warping (DTW), Weighted Sum SVD (WSSVD) and PCA similarity factor (SPCA) in precision/recall on three datasets (2 real-world and 1 synthetic) were conducted to show its superiority for MTS datasets.

In "Indexing of Variable Length Multi-attribute Motion Data," Chuanjun Li, Gaurav Pradhan, Si-Qing Zheng, and Balakrishnan Prabhakaran exploited the singular value decomposition (SVD) properties of multi-attribute matrix data to obtain a representative vector for a multi-attribute matrix to address the multi-attribute data indexing issue. Based on this representative vector, they proposed a simple and efficient interval-tree based index structure for indexing motion data with dozens of attributes. At each tree level, only one component of the query vector needs to be checked during searching, comparing to all the components of the query vector that get involved for an R-tree or its variants. Searching time is independent of the number of pattern motions indexed by the tree, making the index structure well scalable

to large databases. Their experiments show that up to 91~93% irrelevant motions can be pruned for a query, and the query searching time is less than 30 microseconds with the existence of motion variations.

Video classification is an important step towards multimedia understanding. Most state-of-art approaches which apply HMM to capture the temporal information of videos have the limitation by assuming that the current state of a video depends only on the immediate previous state. Nevertheless, this assumption may not hold for videos of various categories. In "VRules: An Effective Association-based Classifier for Videos," Ling Chen, Sourav Saha Bhowmick, and Liang-Tien Chia presented an effective video classifier which employs the association rule mining technique to discover the actual dependence relationship between video states with promising classification performance. The discriminatory state transition patterns mined from different video categories are then used to perform classification. Besides capturing the association between states in the time space, their approach also captures the association between low-level features in spatial dimension to further distinguish the semantics of videos.

In their paper entitled "Automatic Classification of Speech and Music Using Neural Networks," Muhammad K. S. Khan, Wasfi G. Al-Khatib, and Muhammad Moinuddin addressed the need to classify audio into categories such as speech or music, which is an important aspect of many multimedia document retrieval systems. The motivation of their work is the evolving importance of automatic discrimination between speech signal and music signal. In this paper, they proposed the use of the mean and variance of the discrete wavelet transform in addition to other features that have been used previously for audio classification. A Multi-Layer Perceptron (MLP) Neural Networks was used as a classifier. Their initial testing has shown encouraging results that indicate the viability of their approach.

Finally, Latifur Khan on behalf of Lei Wang and Li Liu presented their proposed top-down, subspace clustering algorithm in "Automatic Image Annotation and Retrieval Using Subspace Clustering Algorithm." Their approach tries to address the problem that image data usually have a large number of dimensions but the traditional clustering algorithms assign equal weights to these dimensions, and become confounded in the process of dealing with these dimensions. In their approach, for a given cluster, the relevant features based on histogram analysis were determined and greater weights were assigned to the relevant features as compared to less relevant features. In their experiments, four different models to link the visual tokens with the keywords based on the clustering results of their clustering algorithm and the K-means algorithm were implemented and evaluated using the precision, recall and correspondence accuracy under a benchmark dataset. The results showed that their algorithm is better than the traditional ones for automatic image annotation and retrieval.

In addition to these spectacular sessions, the workshop also held one keynote speech by Prof. Arif Ghafoor at Purdue University. The title of this keynote speech was "Web-Based Multimedia Databases: Prospects and Challenges." This talk elaborated the key challenges of developing Web-based multimedia database applications, including (1) content analysis and indexing of distributed multimedia data and documents, (2) semantic modeling and knowledge-based representation of multimedia data, (3) transformation and organization of multimedia data semantics as a part of Semantic Web, (4) security, privacy and QoS related issues concerning Web-based multimedia database applications, and (5) emerging Web standards and their role in managing distributed multimedia databases. Several solutions and tools that have been developed for Web-based multimedia database systems were also discussed.

The long-term goal of MMDB is to build up a reputable and respectable workshop for the international community. Without the contribution and support from the international community, MMDB could never be classified as a significant and successful workshop as what the MMDB 2004 has achieved!

Furthermore, given the high quality of the papers, several selected papers are invited for the special issue on "Multimedia Databases" in the Information Systems Journal (Elsevier).

Many people contributed to the success of this event. We would like to thank all the participants, paper presenters, and session chairs for their valuable cooperation with the demands placed on them. We wish to thank the program committee members and the external referees for their great work. We are grateful to Professor Ghafoor for accepting to give a keynote presentation. Finally, we would like to express our gratitude towards the CIKM organizers.

Shu-Ching Chen and Mei-Ling Shyu
ACM MMDB 2004 Program co-chairs