MultiMediaMiner: A System Prototype for MultiMedia Data Mining *

Osmar R. Zaïane Jiawei Han Ze-Nian Li Sonny H. Chee Jenny Y. Chiang
Intelligent Database Systems Research Laboratory and Vision and Media Laboratory
School of Computing Science
Simon Fraser University
Burnaby, BC, Canada V5A 1S6
E-mail: {zaiane, han, li, schee, ychiang}@cs.sfu.ca

Abstract

Multimedia data mining is the mining of high-level multimedia information and knowledge from large multimedia databases. A multimedia data mining system prototype, MultiMediaMiner, has been designed and developed. It includes the construction of a multimedia data cube which facilitates multiple dimensional analysis of multimedia data, primarily based on visual content, and the mining of multiple kinds of knowledge, including *summarization, comparison, classification, association*, and *clustering.*

1 Introduction

There has been rapid progress in the field of data mining and data warehousing research, but nothing substantial in mining multimedia. Many data mining and data warehousing systems have been developed for mining knowledge in relational databases and data warehouses [4]. Multimedia has been the major focus for many researchers around the world. Many techniques for representing, storing, indexing, and retrieving multimedia data have been proposed. However, rare are the researchers who ventured in the multimedia data mining field. Most of the studies done are confined to the data filtering step of the KDD process. In [3], Czyzewski shows how KDD methods can be used to analyze audio data and remove noise from old recordings. Chien et al. in [2] use knowledge-based AI techniques to assist image processing in a large image database generated from the Galileo mission. Others use Multimedia to complement data mining systems. Bhandari et al., for instance, marries a data mining application with multimedia resources. His application does not claim to mine a multimedia database, but uses video clips to support the knowledge discovered from a numerical database [1].

Multimedia data mining is a subfield of data mining that deals with the extraction of implicit knowledge, multimedia data relationships, or other patterns not explicitly stored in multimedia databases.

Recent advances in the research on multimedia databases [7, 8, 5] enable creation of large multimedia databases which can be queried in an effective way. These advances, in combination with the research into multimedia database and advances in data mining in relational databases [4], created a possibility for the creation of multimedia data mining systems.

The current MultiMediaMiner system includes four data mining modules for mining knowledge in image and video databases: *characterization, comparison, classification*, and *association*. Additional modules are in the design and development stage.

A more detailed description of the MultiMediaMiner system is presented in Section 2. A summary and a discussion of our on-going research are in Section 3.

2 A database mining system prototype

The MultiMediaMiner system is based on our experiences in the development of an on-line analytical data mining system, DBMiner, and C-BIRD, a system for Content-Based Image Retrieval from Digital libraries.

The DBMiner system applies multi-dimensional database structures [6], attribute-oriented induction, multi-level association analysis, statistical data analysis, and machine learning approaches for mining different kinds of rules in relational databases and data warehouses. C-BIRD system contains 4 major components: (i) Image Excavator (a web agent) for the extraction of images and videos from multimedia repositories, (ii) a preprocessor for the extraction of image features and storing precomputed data in a database, (iii) a user interface, and (iv) a search kernel for matching queries

*Research is supported in part by the Natural Sciences and Engineering Research Council of Canada, the Canadian Network of Centres of Excellence (IRIS:HMI-5 and TL:NCE5.2), and MPR Teltech Ltd.

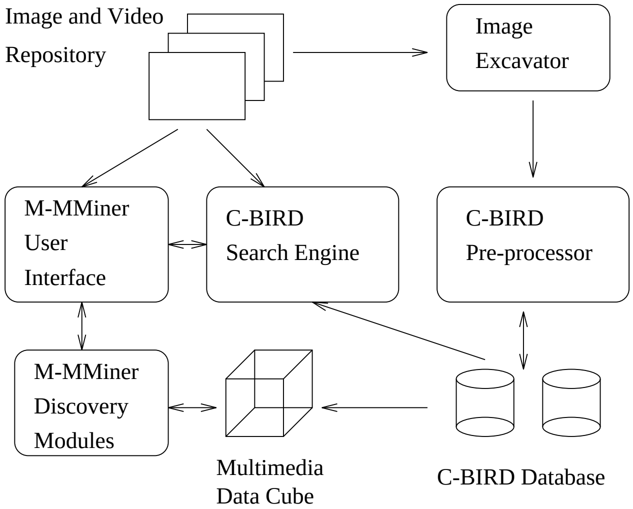

Figure 1: General Architecture of MultiMediaMiner.

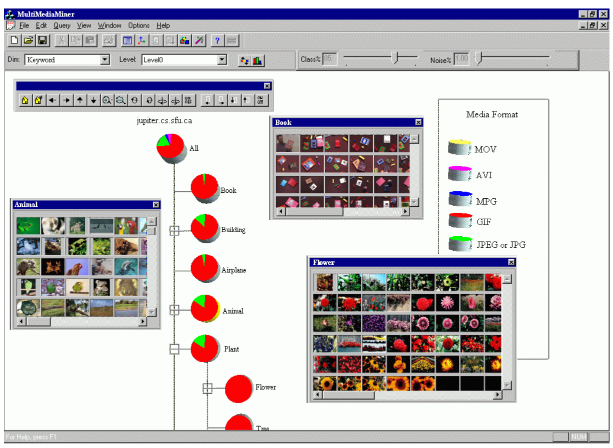

Figure 2: MultiMediaMiner Classifier user interface.

with image and video features in the database. The database used by C-BIRD is an addition to the image repository and contains mainly meta-data extracted by the pre-processor and the Image Excavator, like colour and texture characteristics and automatically generated keywords. MultiMediaMiner, the general architecture of which is shown in Figure 1, inherits the CBIRD database.

For each image collected, the database contains some description information, a feature descriptor, and a layout descriptor. The original image is not directly stored in the database; only its feature descriptors are stored. The description information encompasses fields like: image file name, image URL, image type (i.e. gif, jpeg, bmp, avi, mpeg, ...), a list of all known web pages referring to the image (i.e. parent URLs), a list of keywords, and a thumbnail used by the user interface for image and video browsing. The feature descriptor is a set of vectors for each visual characteristic. The main vectors are: a colour vector containing the colour histogram quantized to 512 colours ($8 \times 8 \times 8$ for $R \times G \times B$), MFC (Most frequent Colour) vector, and MFO (Most Frequent Orientation) vector. The MFC and MFO contain 5 colour centroids and 5 edge orientation centroids for the 5 most frequent colours and 5 most frequent orientations (the edge orientations used are: $0°$, $22.5°$, $45°$, $67.5°$, $90°$, etc.). The layout descriptor contains a colour layout vector and an edge layout vector. Regardless of their original size, all images are assigned an 8×8 grid. The most frequent colours for each of the 64 cells are stored in the colour layout vector and the number of edges for each orientation in each of the cells is stored in the edge layout vector. Other sizes of grids, like 4×4, 2×2 and 1×1, can be derived easily.

The Image Excavator uses image contextual information, like HTML tags in web pages, to derive keywords. By traversing on-line directory structures, like the Yahoo directory, it is possible to create hierarchies of keywords mapped on the directories in which the image was found. These graphs are used as concept hierarchies for the dimension "keyword" in the multimedia data cube.

The multimedia data cube we use has many dimensions. The following are some examples: (1) the size of the image or video in bytes with automatically generated numerical hierarchy; (2) the width and height of the frames (or picture) constitute 2 dimensions with automatically generated numerical hierarchy; (3) the date on which the image or video was created (or last modified) is another dimension on which a time hierarchy is built; (4) the format type of the image or video with two-level hierarchy containing all video and still image formats; (5) the frame sequence duration in seconds (0 seconds for still images) with numerical hierarchy; (6) the image or video Internet domain with a pre-defined domain hierarchy; (7) the Internet domain of pages referencing the image or video (parent URL) with a pre-defined domain hierarchy; (8) the keywords with a term hierarchy defined as described above; (9) a colour dimension with a pre-defined colour hierarchy; (10) an edge-orientation dimension with a pre-defined hierarchy, etc.

The mining modules of the MultiMediaMiner system include four functional modules, *characterizer, comparator, classifier*, and *associator*. The functionalities of these modules are described as follows:

- MM-Characterizer: This module discovers a set of characteristic features at multiple abstraction levels from a relevant set of data in a multimedia database. It provides users with a multiple-level view of the data in the database with roll-up and drill-down capabilities. For example, the module may describe the general characteristics of image sequences based on the topic of the video, the topic being a high level keyword defined in the concept hierarchy. The user can drill-down along the topic dimension to find characteristics of the image sequences based on more concrete topics.

- MM-Comparator: This module discovers a set of comparison characteristics contrasting the features of different classes of the relevant sets of data in a multimedia database. It compares and distinguishes the general features of one set of data, known as the target class, from the other set(s) of data, known as the contrasting class(es). For example, the module may show the differences in video duration and colour richness between videos served in the commercial Internet domain (com) and videos served on the education domain (edu) and created in July 1997.
- MM-Associator: This module finds a set of association rules from the relevant set(s) of data in an image and video database. An association rule shows the frequently occurring patterns (or relationships) of a set of data items in a database. A typical association rule is in the form of "$X \to Y[s\%, c\%]$" where X and Y are sets of predicates, $s\%$ is the support of the rule (the probability that X and Y hold together among all the possible cases), and c% is the confidence of the rule (the conditional probability that Y is true under the condition of X). For example, the module mines association rules like: "*what are relationships among still images, the frequent colours used in them, their size and the keyword* 'sky'*?*" One possible association rule among many to be found is "*if image is big and is related to* sky, *it is* blue *with a possibility of 68%*" or "*if image is small and is related to* sky, *it is* dark blue *with a possibility of 55%*".
- MM-Classifier: This module classifies multimedia data based on some provided class labels, such as topics (based on keywords). The result is an elegant classification of a large set of multimedia data and a characteristic description of each class. Figure 2 shows an output of this module where a classification of images and frames based on their topic, with reference to the distribution of image format, is made for a given Web site. By clicking on a class, a window displays the images pertaining to the class (ex. book, animal, flower in Figure 2).

3 On-going work and Conclusions

We[1] have designed and developed an interesting multimedia data mining system prototype, MultiMediaMiner, with the following features: (i) a multidimensional multimedia data cube, (ii) multiple data mining modules, including MM-Characterizer, MM-Comparator, and MM-Associator, and (iii) an interactive mining interface and display.

There are three major tasks calling for further research into the design and development of the MultiMediaMiner system.

The first task is the improvement of the design and construction of multimedia data cube. Our currrent implementation supports only limited number of intervals on the *colour* and *texture* dimensions in the data cube.

The second task is to enhance our data mining algorithms to take advantage of the MFC and MFO centroids in order to discover interesting spatial relationships.

The third task is the incremental addition of new data mining functionalities into the system. We plan to add a clusterer MM-Cluster-Analyser which would group images into different clusters based on their multiple dimensional features, including both multimedia features, such as colour and edge-orientation, and relational features, such as keywords, URL information, and duration.

[1] Acknowledgments: The authors would like to express their thanks to Jean Hou and Eli Hagen for their invaluable help in the implementation.

References

[1] I. Bhandari, E. Colet, J. Parker, Z. Pines, and R. Pratap. Advanced scout: Data mining and knowledge discovery in NBA data. *Data Mining and Knowledge Discovery*, 1(1):121–125, 1997.

[2] S. Chien, F. Fisher, H. Mortensen, E. Lo, and R. Greeley. Using artificial intelligence planning to automate science data analysis for large image databases. In *Proc. Third Int. Conf. on Knowledge Discovery and Data Mining*, pages 147–150, 1997.

[3] A. Czyzewski. Mining knowledge in noisy audio data. In *Proc. Second Int. Conf. on Knowledge Discovery and Data Mining*, pages 220–225, 1996.

[4] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy. *Advances in Knowledge Discovery and Data Mining.* AAAI/MIT Press, 1996.

[5] M. Flickner, H. Sawhney, W. Niblack, et al. Query by image and video content: The QBIC system. *IEEE Computer*, 28(9):23–32, 1995.

[6] J. Han, J. Chiang, S. Chee, J. Chen, Q. Chen, S. Cheng, W. Gong, M. Kamber, G. Liu, K. Koperski, Y. Lu, N. Stefanovic, L. Winstone, B. Xia, O. R. Zaïane, S. Zhang, and H. Zhu. DBMiner: A system for data mining in relational databases and data warehouses. In *Proc. CASCON'97: Meeting of Minds*, pages 249–260, Toronto, Canada, November 1997.

[7] S. Khoshafian and A. B. Baker. *Multimedia and Imaging Databases.* Morgan Kaufmann Publishers, 1996.

[8] B. Özden, A. Biliris, R. Rastogi, and A. Silberschatz. A low-cost storage server for movie on demand databases. In *Proc. 20th Int. Conf. on Very Large Data Bases*, pages 594–605, 1994.