Report on MDM/KDD2000: The 1st International Workshop on Multimedia Data Mining

Simeon J. Simoff
University of Technology Sydney
Sydney, NSW 2007
Australia
simeon@it.uts.edu.au

Osmar R. Zaïane
University of Alberta
Edmonton, Alberta, T6G 2H1
Canada
zaiane@cs.ualberta.ca

Abstract
This short report on the First International Workshop on Multimedia Data Mining summarizes the activities that took place during the workshop and gives pointers to resources where more information can be found.

Keywords
Multimedia, knowledge discovery, databases, image repositories, sound, speech, video, data mining.

Introduction
The First International Workshop on Multimedia Data Mining is no more an organizer dream, but a reality - it was held on August 20th, 2000, in conjunction with the Sixth ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. The workshop organizing team (S. J. Simoff and O. R. Zaïane) received 27 submissions from 12 different countries: Australia, Belgium, Canada, China, France, Germany, Japan, Malaysia, Switzerland, Taiwan, United Kingdom, and United States of America. All papers were extensively reviewed by three referees drawn from the international program committee and external reviewers. The program committee consisted of Max Bramer (University of Portsmouth), Alex Duffy (University of Strathclyde), Max J. Egenhofer (University of Maine), Tom Gedeon (Murdoch University), William Grosky (Wayne State University), Howard J. Hamilton (University of Regina), Jiawei Han (Simon Fraser University), Odej Kao (Technical University of Clausthal), Nik Kasabov, (University of Ottago), Raymond Ng (University of British Columbia), Timothy K. Shih (Tamkang University) and Jaideep Srivastava (University of Minnesota). 14 submitted papers and 1 invited paper were selected for presentation and inclusion in the workshop proceedings.

Topics Covered
Vinton G. Cerf, President of the Internet Society and Senior Vice-President of MCI Data Services compares the activities surrounding multimedia ideas with the metaphor of a disturbed ant hill, in which the inhabitants run hither and yon to discover the cause of disturbance and, perhaps, to do something about it. No wonder researchers and developers in multimedia information systems turn to data mining and knowledge discovery methods looking for techniques for improving the indexing and retrieval of necessary information out of these data sets. The aim of the workshop was to bring together experts in multimedia information systems, state-of-the-art data mining and knowledge discovery in multimedia database systems, and domain experts from various applied disciplines with potential in multimedia data mining. Accepted papers were presented in three sessions: *Mining spatial multimedia data*; *Mining audio data and multimedia support*; and *Mining image and video data*.

The submissions to the workshop revealed a variety of topics that come under the umbrella of multimedia data mining. The time limit of a one-day workshop limited the acceptance of several other good submissions. Further, the underlined name indicates the person who presented the paper during the workshop.

Mining Spatial Multimedia Data
The first session started with a presentation by one of the "fathers" in spatial data mining – Jiawei Han (Simon Fraser University, Canada). He presented a joint work with Anthony K. H. Tung from the same university, Raymond T. Ng (University of British Columbia, Canada) and Laks V. S. Lakshmanan (IIT, Bombay, India & Concordia University, Canada) on *Geo-Spatial Clustering with User-Specified Constraints*. In real world spatial data such constraints can occur from the representations of physical obstacles that exist in a region of clustering. In other cases such constraints may be defined from the semantics of the problem. The paper addresses these two types of constrained clustering, and proposes some formalisms and techniques for solving them. In *Multi-level Indexing and GIS Enhanced Learning for Satellite Imageries*, Krzysztof Koperski and Giovanni B. Marchisio (Data Analysis Division of MathSoft Inc., Seattle, USA) addressed two issues in the data analysis of remotely sensed images - the extraction of information that enables reduction of the data from multi-spectral images into a number of features and the organisation of these features under a schema that would allow flexible and scalable discovery of the knowledge from the databases of remotely sensed images. The paper presents the concepts behind the *GeoBrowse* project - a data mining system for the analysis of satellite images, and some preliminary results of the experiments with the collection of LANDSAT images. The idea revolves around hierarchical feature extraction schema (extracted features are at the level of pixel, region and tile), enhanced with some a priori knowledge about the models that are based on data from satellite images. In *Predicting Locations Using Map Similarity (PLUMS): A Framework for Spatial Data Mining*, Sanjay Chawla (Vignette Corporation, Massachusetts, USA), Shashi Shekhar and Weili Wu (University of Minnesota, Minneapolis, USA), and Uygar Ozesmi (Ericyes University, Kayseri, Turkey) proposed an approach for supervised spatial data mining problems (PLUMS - Predicting Locations Using Map Similarity), which incorporates spatial autocorrelation. Authors claimed that at a fraction of the computational costs, PLUMS achieves accuracy comparable to the state-of-the-art spatial statistics methods.

Mining Audio Data and Multimedia Support

In *Learning Prosodic Patterns for Mandarin Speech Synthesis*, Yiqiang Chen and Wen Gao (Chinese Academy of Sciences, Beijing, China), and Tingshao Zhu (University of Alberta, Edmonton, Canada) considered a machine learning approach to the extraction of prosodic patterns in speech synthesis. Prosody includes the phrase and accent structure of speech, which makes human speech so distinct. Consequently learning and applying the appropriate prosodic patterns (as a variation of pitch in speech) is expected to improve the quality of the speech synthesisers. Unfortunately, for visa reasons, the author who was supposed to present the paper could not attend. In *Unsupervised Classification of Sound for Multimedia Indexing*, Bruce Matichuk and Osmar R. Zaïane (University of Alberta, Edmonton, Canada) presented preliminary results in the clustering of audio segments based on frequency and harmonic analysis. The ROCK algorithm used by the authors as the clustering engine, is very sensitive to the threshold value. The authors proposed to ease this dependence by replacing the manual threshold selection procedure with adaptive estimation of the threshold value during sound classification. In *Effective Retrieval of Audio Information from Annotated Text Using Ontologies*, Latifur Khan and Dennis McLeod (University of Southern California, Los Angeles, USA) proposed an approach for retrieval of audio information, which includes ontology-based model for the generation of metadata for audio and the selection of audio information in user-customised manner. Ontology in this work is used in the sense of domain-dependent taxonomy of terms.

In *Incorporating Domain Knowledge with Video and Voice Data Analysis in News Broadcasts*, Kim Shearer (IDIAP, Switzerland), Chitra Dorai (IBM T. J. Watson Research Center, NY, USA) and Svetha Venkatesh (Curtin University of Technology, Australia) complement Khan and McLeod's approach in the area of audio and video data with the fusion of several low-level image and aural analysis techniques, together with an a priori knowledge of the shot syntax for a particular domain. The resultant techniques aim to improve existing methods in the area of video annotation, indexing and retrieval of video information. In *Multimedia Support for Complex Multidimensional Data Mining*, Monique Noirhomme-Fraiture (Institut d'Informatique, FUNDP, Belgium) brings multimedia to extend the dimensions of visual and other human sensations support for data mining. The work includes a variation of the radial graph data visualisation technique extended with semantic use of audio fragments and animation where necessary. This multi-modal approach aims at the decrease of the load on human visual processing system in a visual data mining session.

Mining Image and Video Data

In *A Self Organizing Map (SOM) Extended Model for Information Discovery in a Digital Library Context*, Jean-Charles Lamirel and Hager Kammoun (LORIA, Cedex, France), and Jacques Ducloy (INIST, Cedex, France) presented a knowledge discovery tool for multimedia databases (or digital libraries), which learning model is an extension of the basic Kohonen self-organising map techniques. The developments are developed and tested on a version of an iconographic digital library server developed under the BIBIAN (Bibliographic and Iconographic Base Art Nouveau) project. In *Learning Feature Weights from User Behavior in Content-Based Image Retrieval*, Henning Müller, Wolfgang Müller, Stephane Marchand-Maillet and Thierry Pun (University of Geneva, Switzerland), and David McG Squire (Monash University, Melbourne, Australia) described an algorithm for obtaining knowledge about the importance of features from analysing user log files of a content-based image retrieval system. The work is based on log files obtained from the use of Viper image retrieval system, operating with four different groups of image features: global color histogram based on HSV color space; local color blocks at different scales; global texture characteristics and local Gabor filters. In *When Image Indexing Meets Knowledge Discovery*, Chabane Djeraba (IRIN, Ecole Polytechnique de l'Universite de Nantes, Cedex, France) challenges the automatic building of semantic descriptions of image contents. The idea is to apply data mining techniques to discover implicit relations between basic image features and then use discovered relations to improve extracted semantic descriptions for content-based image retrieval. In their study *Semantic Indexing and Temporal Rule Discovery For Time-Series Satellite Images*, Rie Honda, Hirokazu Takimoto and Osamu Konoshi (Kochi University, Kochi, Japan) presented a framework for data mining from a set of time-series satellite images, including moving objects. The proposed framework is based on the application of self-organising maps for image classification and an original approach towards the discovery of time-dependent association rules based on time-series pattern analysis. In *Data Mining from Functional Brain Images,* Mitsuru Kakimoto, Chie Morita, Yoshiaki Kikuchi and Hiroshi Tsukimoto (Toshiba Corporation, Japan) studied the discovery of relations between brain areas and brain functions from functional brain images. In addition to the common difficulty of reducing images to a symbolic description, functional brain images challenge the data mining field with the possible existence of correlation between adjacent pixels in a image and the limited number of samples available from a single object. The authors applied to real brain images a data mining algorithm developed by Tsukimoto and Morita. The paper presents the result including the discovery of certain rules for finger tapping action and speech related action. In *Mining Cinematic Knowledge*, Duminda Wijesekera and Daniel Barbara (George Mason University) reported current investigation in an on-going effort to create a movie mining system. The emphasis of this project is to examine the suitability and applicability of existing data mining concepts to multimedia data, where semantic content is time sensitive and constructed by fusing data obtained from component streams.

Finally, Simeon J. Simoff presented *Variations on Multimedia Data Mining*, aiming at setting up the background for the discussion section. Rather than making an overview of existing works and projects related to multimedia data mining, the presentation outlined some facets of the domain along two areas where multimedia data mining is becoming popular - case-based reasoning where cases are represented as multimedia and environments for computer supported collaborative work.

Conclusion

The workshop ended with a discussion on the scope of multimedia data mining and the need to organize a similar event in the future. With regard to the scope of multimedia data mining, it was clear through the discussion that the extent of multimedia is not widely approved. While text, maps, video, sound and images

typically fall into the realm of multimedia, research fields such as spatial data mining and text mining are already known active disciplines. The participants of the workshop reach the consensus that multimedia data mining is emerging as its own distinct area of research. This new research area should mainly include, but not exclusively, mining from images, sound and video streams. All concurred on the importance to organize another workshop on the topic.

Overall, it was a successful workshop, where interesting insight into current research projects was given and a variety of topics related to multimedia data mining were covered. The workshop was a good opportunity for exchange of ideas and networking.

It was decided to create a mailing list to continue the discussions and create a forum on multimedia data mining research related issues. The mailing list address is: **mdm-kdd@cs.ualberta.ca**.

The proceedings of the workshop are available in electronic form at the workshop web site. Check out:

http://www.cs.ualberta.ca/~zaiane/mdm_kdd/

or

http://www.acm.org/sigkdd/proceedings/mdmkdd00/

Acknowledgements

We are thankful to the members of the program committee who helped us review the submissions, and we are particularly grateful to Philip Chan, the Workshops Chair at SIGKDD'2000, for coordinating the workshops and doing a tremendous job with the organization.