

MDM/KDD: Multimedia Data Mining for the Second Time

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

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

Abstract

This brief report summarizes the presentations, conclusions and directions for future work that were discussed during the second edition of the International Workshop on Multimedia Data Mining. The report includes references to resources where one can find more information about the workshop format, the proceedings and the workshop participants.

Keywords

Multimedia data mining, knowledge discovery, databases, digital media, sound analysis, video analysis.

Introduction

The second edition of the International Workshop on Multimedia Data Mining was held in San Francisco, California on August 26th, 2001, in conjunction with the Seventh ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. The workshop organizing team (O. R. Zaiane and S. J. Simoff) received 20 submissions from 10 different countries: Australia, Canada, China, France, Germany, India, Japan, Singapore, United Kingdom, and United States of America. The international program committee, which, in comparison with the first edition of the workshop in year 2000, was significantly extended and included Terry Caelli (University of Alberta), Chabane Djeraba (University of Nantes), Chitra Dorai (IBM Thomas J. Watson Research Center), Alex Duffy (University of Strathclyde), Max J. Egenhofer (University of Maine), William Grosky (Wayne State University), Howard J. Hamilton (University of Regina), Jiawei Han (Simon Fraser University), Alexander G. Hauptmann (Carnegie Mellon University), Wynne Hsu (National University of Singapore), Odej Kao (Technical University of Clausthal), Nik Kasabov (University of Otago), Paul Kennedy (University of Technology-Sydney), Latifur Khan (University of Texas), Flip Korn (AT&T Laboratories), Brian Lovell (University of Queensland), Mark Maybury (MITRE Corporation), Mario Nacsimento (University of Alberta), Gholamreza Nakhaeizadeh (Daimler-Chrysler), Monique Noirhomme-Fraiture (Institut d'Informatique, FUNDP), Vincent Oria (New Jersey Institute of Technology), Jian Pei (Simon Fraser University), Simone Santini (University of California - San Diego), John R. Smith (IBM Thomas J. Watson Research Center), Duminda Wijesekera (George Mason University), Ian H. Witten (University of Waikato), and the workshop organizers themselves. All papers were thoroughly reviewed by three referees drawn from this committee and external reviewers. Reviewers scores were evaluated to a single grade and the first 12 papers with the highest grades were selected for presentation and inclusion in the workshop proceedings. The number of presentations has been decreased in favor of increased time for questions and discussions.

Motivation and Topics Covered this Year

The year since the 1st International Workshop on Multimedia Data Mining (MDM/KDD2000)¹ has seen significant advances in multimedia delivery technologies. The higher performance levels at competitive costs means that businesses are implementing these multimedia (or digital media) technologies on a broader spectrum, hence they are starting to enable and proliferate intelligent multimedia solutions in a spectrum of commercial and laboratory projects. Such intelligent solutions are usually based on some data mining and analysis techniques. The presentations at MDM/KDD2000 revealed that many researchers and developers in the areas of multimedia information systems and digital media turn to data mining and knowledge discovery methods for techniques that can improve indexing and retrieval in digital media. During the discussion at the end of the workshop participants identified that there is a need for (i) development and application of specific methods, techniques and tools for multimedia data mining; and (ii) frameworks that provide consistent methodology for multimedia data analysis and integration of discovered knowledge back in the system where it can be utilized. Consequently, the papers in this year's workshop were grouped in the following three streams: *Frameworks for Multimedia Mining*; *Multimedia Mining for Information Retrieval*; and *Applications of Multimedia Mining*.

The grouping of the papers reflects also the *aim of this year's workshop* - to bring together experts from a broader range of disciplines than last year, related to the analysis of digital media, unstructured data mining, and knowledge engineering and utilization. The papers in the workshop proceedings will be of particular interest to researchers in intelligent information and data analysis technologies. Below we present a brief overview of the workshop presentations.

Frameworks for Multimedia Mining

The first session of the workshop included two presentations. In *Image Mining: Issues, Frameworks and Techniques*, Ji Zhang Wynne Hsu and Mong Li Lee (National University of Singapore, Singapore) presented an integrated overview of the research issues and developments in the area of image mining. Image mining frameworks are grouped into two broad categories: function-driven and information-driven. Authors continue with an overview of the state-of-the-art techniques and systems, with references to 36 works. The presentation concluded with some future research directions for image mining, setting up the context

¹ For more information on MDM/KDD2000 see Simoff, S. and Zaiane, O. (2000). Report on MDM/KDD2000: The 1st International workshop on multimedia data mining, *SIGKDD Explorations*, 2 (2), 103-105.

for the other presentations in the area of image mining. In *Multimedia Mining of Collaborative Virtual Workspaces: An Integrative Framework for Extracting and Integrating Collaborative Process Knowledge*, Simeon J. Simoff (University of Technology Sydney, Australia) and Robert P. Biuk-Aghai (University of Macau, Macau S.A.R. China) addressed a new area of data mining in collaborative virtual environments. The paper presents a framework for integrating multimedia data mining techniques with collaborative virtual environments, starting from early conceptual development. The aim of presented research was to utilize the multimedia data about the actions and content of collaborative activities in projects conducted in virtual environments, extract meaningful insights out of it and feed discovered knowledge back into the environment. The ideas are illustrated with examples from the application of the framework to collaborative workspaces developed in LiveNet, a virtual workspace design system.

Multimedia Mining for Information Retrieval

In *The PERSEUS Project: Creating Personalized Multimedia News Portal*, Victor Kulesh (Accenture, Northbrook, IL, USA and Oakland University, Rochester, MI, USA), Valery A. Petrushin (Accenture, Northbrook, IL, USA) and Ishwark. Sethi (Oakland University, Rochester, MI, USA) presented the Perseus project. The project is devoted to developing techniques and tools for creating personalized multimedia news portals. The multimedia data mining techniques in this case are used for extracting video clips automatically from TV broadcasts, based on the user's preferences. The clips are then augmented with other relevant news from other sources on the Internet. Presenters discussed their approaches to event mining and tracking on the Internet, commercial detection and recognition in video and audio streams, and selection of relevant news video fragments, based on closed captioning and audio transcripts. In *Automatic Feature Mining for Personalized Digital Image Retrieval*, Kyoung-Mi Lee and W. Nick Street (The University of Iowa, Iowa City, IA, USA) addressed the important issue of measuring similarity based on feature representations of multimedia data. There is an agreement that in its current form the feature space approach does not necessarily represent the notion of similarity in human perception. One of the characteristics of human perception of similarity is that similarity does not vary in the same proportion in all directions in the feature space. Authors presented an incremental method to automatically obtain feature weights based on both the clustered database and on relevance feedback. They presented the results of shape-based indexing and retrieval, showing that using cluster information for an initial search gives better results than using the standard distance. In *Relationship Extraction from Large Image Databases*, Chabane Djeraba (IRIN, Ecole Polytechnique de l'Université de Nantes, France), addressed the semantic processing of image feature space. The paper presented an algorithm that discovers relationships between image features. Relationships are ranked based on confidence measures. Before the actual mining, the image features (e.g. similar colors and textures) for a particular database were summarized in a visual thesaurus. The relationships discovered assisted the automatic categorization of images during their insertion into image databases. At the retrieval stage these relationships were used to improve the accuracy in retrieving relevant images. The last presentation in this session was on

Semantic Content-Based Retrieval in a Video Database by Pramod K. Singh (University of Technology Sydney, Australia) and A. K. Majumdar (Indian Institute of Technology, Kharagpur, India). Authors discussed the issues in managing temporal information of video data that are common to many application areas. The echocardiogram video data management is the specific area addressed in the paper. The paper describes an approach of semantic content-based retrieval of video data using object state transition data model. The advantage of using this model is in allowing storage and indexing of echocardiogram video at different levels of abstraction based on semantic features of video objects. Authors presented briefly the system that utilized the proposed approach and discussed the issues in querying the video database.

Applications of Multimedia Mining

The applications session started with the presentation on *An Interactive Environment for Kansei Data Mining* by Nadia Bianchi-Berthouze (University of Aizu, Aizu Wakamatsu, Japan). The paper addressed the modeling of visual impressions from the point of view of multimedia data mining. Visual impressions are the impressions experienced when observing images. They are highly subjective, complex and difficult to be made explicit. Nadia presented an approach that took into account the diverse information involved in the mapping between images and visual impressions that those images triggered in an observer, and the way the observer expressed such impressions. The modeling process integrated different techniques of multimedia data mining to learn associations between image characteristics and impression words. The mining operations are assisted via mechanisms of externalization of user impressions – the user is offered tools to express his/her mental process and naïve models into a formal specification. Technologically a Web-based search engine, created for retrieving images by impression words, supports the modeling in this research. In *Data Mining for Typhoon Image Collection*, Asanobu Kitamoto (National Institute of Informatics, Tokyo, Japan) presented the application of image data mining methods to a narrow domain - the analysis and prediction of typhoons. The image analysis is based on a number of well-known techniques, such as principal component analysis, self-organizing maps and time-series analysis to characterize and visualize the statistical properties of typhoon cloud patterns. The prediction is based on the application of an instance-based learning method for analogy-based prediction using past similar patterns. Asanobu emphasized the fundamental problems in typhoon prediction from past similar patterns due to the chaotic nature of the atmosphere. The results tested for this research was the typhoon image collection that was established in the National Institute of Informatics. This medium-size, well-controlled, and richly-variational collection includes approximately 34,000 typhoon images created from satellite images of a geostationary meteorological satellite GMS-5. In *Multimedia Data Mining for Traffic Video Sequences*, Shu-Ching Chen (Florida International University, Miami, FL, USA), Mei-Ling Shyu (University of Miami, FL, USA), Chengcui Zhang and Jeff Strickott (Florida International University, Miami, FL, USA) presented a framework for multimedia data mining from traffic video sequences recorded at road intersections. Traffic video analysis can discover queues, vehicles identification, traffic flow, and spatio-temporal relations of the vehicles at the intersections including incidents. Several methods are used to

analyze the traffic video sequence - background subtraction (a technique to remove nonmoving components from a video sequence), image/video segmentation and object tracking. The spatio-temporal relationships of the vehicle objects in each frame are identified and modeled using multimedia augmented transition networks and multimedia input strings. The multimedia transition networks (labeled directed graphs) are used to derive hierarchical representations of the video clips. The multimedia string represents the transition path in symbolic form that can be processed with some grammar rules. The effectiveness of the proposed methodology was demonstrated with the results of its application on a real-life traffic video sequence. In *A Bayesian Learning Algorithm of Discrete Variables for Automatically Mining Irregular Features of Pattern Images*, Hanchuan Peng (Johns Hopkins University, Baltimore, MD, USA) and Fuhui Long (Duke University, Durham, NC, USA) presented an image mining algorithm for irregular feature extraction, i.e. for extracting features that can hardly be described in a quantitative way. The recognition of unconstrained handwritten characters is an example of the irregular feature extraction problem. The authors applied a Bayesian learning algorithm to mine the unknown features. Proposed algorithm is based on learning belief networks of pattern image pixels, each of which is regarded as a discrete variable with a limited number of states. The probability of belief network, i.e. Bayesian metrics, is chosen to measure the associations between image pixels and the pattern image category. The process includes also learning the structure of the belief network, where clusters of “equivalent” pixels are regarded as the irregular features. In the unconstrained handwritten character feature extraction experiments, a hierarchy of statistically optimal feature vectors was obtained by averaging the pixel clusters over many independent experiments. The presented algorithm outperformed several well-known techniques applied to the problem, including template-based matching methods, statistical methods, and few classes of neural network models. In *Application of Data Mining Techniques for Medical Image Classification*, Maria-Luiza Antonie, Osmar R. Zaiane and Alexandru Coman (University of Alberta, Edmonton, Canada) presented the results of comparison of two data mining techniques – backpropagation neural network and constraint association rule mining, to classify digital mammograms for assisting medical specialists in the detection of breast cancer. Mammograms as images are difficult to interpret and the authors presented the preprocessing procedures applied to improve the quality of the images, hence, to improve the results of the feature extraction and the effectiveness of the data mining techniques. According to the results the backpropagation neural network was less sensitive to the imbalance (normal and abnormal cases) of the data set than the association rule mining method. The results showed that the two approaches performed well, obtaining a classification accuracy reaching over 70% percent for both techniques. Conducted experiments demonstrated the use and effectiveness of association rule mining in image categorization. In the last presentation at the workshop, *A Computer-Aided Visual Exploration System for Knowledge Discovery from Images*, Yusuke Uehara, Susumu Endo, Shuichi Shiitani, Daiki Masumoto, and Shigemi Nagata (Fujitsu Laboratories Ltd., Chiba, Japan) presented a hypothesis-driven multimedia mining process, which combined human cognitive abilities for feature detection

with computing methods for feature extraction in texts and images. On the computing side the text feature extraction was based on the vector space model, the image feature extraction support operated with color features, shape features and texture features. The overall exploration process included, hypothesis formulation and hypothesis verification steps. During the hypothesis formulation stage, the human interacts with the images in consideration and formulates hypotheses (for example, association rules between visual features of images and content of collateral text, e.g. if popular handbags have a specific design, an association rule can represent the relation between a visual feature of these handbags and trend in sales data described in the text). At this stage computer aid is in locating and displaying images appropriately. During the verification stage computer can aid in calculating different characteristics of the hypothesis (in the example with the association rule that can be the support and confidence estimates). Once verified, a hypothesis can be accepted as a chunk of knowledge. Authors presented an experimental system named MIRACLES (Multimedia Information Retrieval, Classification, and Exploration System), designed to aid humans during the hypothesis formulation stage.

Conclusion

The workshop ended with a discussion, which revised the scope of multimedia data mining identified during the previous workshop and identified the directions for the research in the area. With regard to the scope of multimedia data mining, it was clearly identified the need to approach multimedia data as a “single unit” rather than ignoring some layers in favor of others. The participants of the workshop confirmed that multimedia data mining is a distinct area of research with high potential for the application of its methods in the medical, design, art and entertainment domains. The research in multimedia mining should be extended in the area of collaborative virtual environments, 3D virtual reality systems, musical domain and e-business. All participants emphasized the importance to ensure that this event will continue on an annual basis.

Overall, it was a successful workshop, with different focus from MDM/KDD2000 workshop. It demonstrated the increase in theoretical research in methods for multimedia mining and a growing area of application of such methods. The workshop was a great and unique opportunity for exchanging ideas and extending the research community in this data mining area.

Discussion of topics related to multimedia data mining can continue on the multimedia data mining research list with address: mdm-kdd@cs.ualberta.ca.

The complete proceedings of the workshop can be found at http://www.cs.ualberta.ca/~zaiane/mdm_kdd2001/papers/mdm01.pdf.gz

Acknowledgements

We are thankful to the members of the KDD workshop committee for providing the opportunity to stage this event. Special thanks go to the workshop program committee for promptly providing the organizers with high quality paper reviews.