Tenth International Workshop on Multimedia Data Mining

Valery A. Petrushin
The Nielsen Company
150 Martingale Rd
Schaumburg, IL 60173, USA
vapetr3@hotmail.com

Jia-Yu (Tim) Pan
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043, USA
jypan@google.com

Cees G.M. Snoek
University of Amsterdam
Science Park 904
1098 XH Amsterdam, The Netherlands
cgmsnoek@uva.nl

ABSTRACT

In this report we provide a summary of the tenth Multimedia Data Mining Workshop that was held in conjunction with the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2010), July 25-28 in Washington, DC.

Keywords

Multimedia data mining, multimedia information retrieval, knowledge discovery, semantic image classification, clustering, video analysis.

1. INTRODUCTION

As evidenced by the success of the previous editions of the Workshop, there is an increasing interest in new techniques and tools that can detect and discover patterns in multimedia data, that can lead to new knowledge. It is well known that multimedia information is ubiquitous and is often required, if not essential, in many applications. For example, in homeland security applications, we may need to mine data from an air traveler's credit history, traveling patterns, photo pictures, and videotapes from surveillance cameras in the airport. In the World Wide Web, user experience can be greatly improved if the abundant sources of text, images, audio, and video can be indexed and combined to satisfy a user's information need. In medicine, a disease might be predicted more accurately if the magnetic resonance imaging imagery is mined together with other information about the patient's condition. Similarly, in bioinformatics data are available in multiple formats. While mining on structured data or each type of modality of multimedia data such as text data, imagery data, and video data has been broadly addressed, there has not been much effort focusing on integrated approaches to mining information from multiple modalities, multiple sources and multiple formats.

For example, tools are needed for discovering relationships between objects or segments within images, managing multimedia data in the WWW, classifying images based on their content, extracting patterns in sound, categorizing speech and music, and recognizing and tracking objects in video streams. There is also an increasing interest in the analysis of multimedia data generated by different distributed applications, such as collaborative virtual environments, virtual communities, media sharing platforms, and multi-agent systems.

The Multimedia Data Mining (MDM) Workshop is one of the oldest workshops that accompany the KDD Conference. This year the Tenth Workshop has been conducted. The previous nine workshops on Multimedia Data Mining have been held in conjunction with KDD 2000 (Boston, MA), KDD 2001 (San Francisco, CA), KDD 2002 (Edmonton, Canada), KDD 2003 (Washington, DC), KDD 2004 (Seattle, WA), KDD 2005 (Chicago, IL), KDD 2006 (Philadelphia, PA), KDD 2007 (San Jose, CA) and KDD 2008 (Las Vegas, NV) respectively. The 2010 Workshop was held in a half day format.

The major topics of the workshop include the following:

- Data mining of media rich platforms (on-line communities, blogs, social networks, virtual communities, virtual worlds).
- Data mining of large datasets of user generated content (e.g., YouTube, Flickr, etc.).
- Integrated mining of different data formats (text, speech, video, images, relational data)
- Mining of data streams combined with structured data.
- Representation and reuse of discovered multimedia knowledge.
- Real-time multimedia data mining systems.
- Multimedia data mining techniques for specific domains and applications (e.g., biomedicine, education, entertainment, security, animation, art, etc.).
- Topic and event discovery in large multimedia repositories.

The submissions included 16 papers from the authors from ten countries: Australia, Canada, China, Germany, Hong Kong, Italy, Japan, South Africa, Tunisia, and USA. Each submission was reviewed at least by three program committee members. Ten papers were selected for publication and presentation at the workshop.

2. CONTRIBUTED PAPERS

The authors of the first paper "Large Scale Fingerprint Data Mining" (*Baughman, van Der Stockt, Greenland*) presented research conducted at IBM, which deals with enormous fingerprint databases of millions of individuals. The authors suggested using a support vector machine recognizer within a novel hyperspace structure. The SVM is formalized with a high dimensional hyperspace structure with an internal bootstrapped c-means clustering algorithm and probabilistic neural network. The suggested approach allows to reduce search space, improve performance and decrease true negatives while maintaining true positives at some level.

The second paper "Relevance Feature Mapping for Content-Based Image Retrieval" (*Zhou, Ting, Liu, Yin*) proposes a novel ranking framework for content-based image retrieval with a unique relevance feature mapping obtained from an ensemble of ranking models. Each relevance feature measures the relevance of an image to some profile underlying the image database. The framework assumes a two-stage process. In the on-line modeling stage, it constructs a collection of models which maps all images in the database to the relevance feature space. In the on-line retrieval stage, it assigns a weight to every relevance feature based on the query image, and then ranks images in the database according to their weighted average feature values. The framework also incorporates relevance feedback, which modifies the ranking based

on the feedbacks through re-weighted features. The proposed framework has the following unique characteristics: it does not use any distance or similarity measure and has linear time and space complexities with respect to the database size; it has constant on-line retrieval time and it can deal with high-dimensional image databases with constant time complexity.

The third paper "Bag of Visual Words Revisited - An Exploratory Study on Robust Image Retrieval Exploiting Fuzzy Codebooks" (*Kogler, Lux*) is also devoted to content-based image retrieval. The authors conducted an exploratory study revisiting the bag of visual words approach by applying a fuzzy clustering technique for visual words creation and visual words assignment. They show that fuzzy clustering leads to more robust results in terms of retrieval performance.

The authors of the fourth paper "Emerging Topic Detection on Twitter based on Temporal and Social Terms Evaluation" (*Cataldi, Di Caro, Schifanella*) analyzed short text messages that are published on the Twitter Web site to detect emerging topics. They propose a novel topic detection technique that permits to retrieve in real-time the most emergent topics expressed by the community. First, the contents (set of terms) of the tweets are extracted and a model the term life cycle is applied to detect emerging terms. A term is defined as emerging if it frequently occurs in the specified time interval and it was relatively rare in the past. Then the importance of information weighted taking into account its source by analyzing the social relationships in the network with the well-known Page Rank algorithm. Finally, emerging topics are detected using topic graph, which connects the emerging terms with other semantically related keywords. The authors show the validity of proposed approach in several different case studies.

The fifth paper "Large Scale Image Clustering with Support Vector Machine based on Visual Keywords" (*Chang, Ip, Feng*) continues the popular at the Workshop topic of image mining using visual keywords. The authors generalized the SVM clustering method to multi-class clustering via two different strategies - One-Against-All and hierarchical clustering and applied it to large scale image clustering based on the visual keywords representation and Histogram Intersection Kernel. Experiments on two benchmark databases show that proposed approach produces better clustering quality and more computationally efficient than the traditional SVM clustering method.

The sixth paper "Example-based Event Retrieval in Video Archive using Rough Set Theory and Video Ontology" (*Shirahama, Uehara*) deals with retrieving events of interest from a video archive. The authors combined two approaches for retrieving events - "rough set theory" to extract multiple classification rules, each of which correctly identifies a subset of shots of the event, and "video ontology", which is a formal and explicit specification of concepts, concept properties and relations among concepts. The combination of these approaches allows improving the retrieval quality, which was shown using TRECVID 2009 video data.

The authors of the seventh paper "DisIClass: Discriminative Frequent Pattern-Based Image Classification" (*Kim, Jin, Han*) developed new approaches to image classification. First they proposed a new image representation method B2S (Bag-to-Set) that keeps all frequency information but it is more discriminative than traditional histogram based bag representation. Then, based on B2S, they constructed two different image classification approaches. First, they used the Spatial Pyramid Matching algorithm and achieved 20% improvement in accuracy of classification. Second, they designed a Discriminative Frequent Pattern-Based Image Classification (DisIClass) framework to apply data mining algorithms to image classification. The .DisIClass adapts the locality property of image data, and applies sequential covering method to induce the most discriminative feature sets from a closed frequent item set mining method. The DisIClass approach outperformed both the classical- and B2S-based Spatial Pyramid Matching approaches.

The authors of the eighth paper "Measuring Performance of Web Image Context Extraction" (*Alcic, Conrad*) developed a framework for objective evaluation and comparison of the performance of image context extraction methods. They collected a large ground truth dataset consisting of diverse Web documents from real Web servers and defined performance measures adapted to the special properties of the context extraction task. The test dataset consists of about 13,000 documents that include more than 155,000 images The authors tested and reported results for eight extraction methods using precision, recall, F1 score and F1 score deviation as performance measures.

The ninth paper "Web Scale Computer Vision using MapReduce for Multimedia Data Mining" (White, Yeh, Lin, Davis) explores computer vision applications of the MapReduce framework. The authors discuss the MapReduce based implementation for several computer vision algorithms such as, classifier training, sliding windows, clustering, bag-of-features, background subtraction, and image registration. The also provided the experimental results for the *k*-means clustering and single Gaussian background subtraction algorithms that used a 410 node Hadoop cluster.

The tenth paper "Approximate Variable-Length Time Series Motifs Discovery using Grammar Inference" (Li, Lin) deals with the problem of identifying frequently occurring patterns or motifs in time series data. The authors propose a novel approach, based on grammar induction, for approximate variable-length time series motif discovery. The algorithm offers the advantage of discovering hierarchical structure, regularity and grammar from the data. The preliminary results show that the grammar-based approach is able to find some important motifs, and suggest that the new direction of using grammar-based algorithms for time series pattern discovery is worth exploring.

At the end of the Workshop Dr. Petrushin proposed all the participants to answer to the following question: "Do you believe that in five years the content-based image/video search will be used in commercial tools?" More than 70% of Workshop participants answered "yes" to this question.

3. ACKNOWLEDGMENTS

We thank the KDD Workshop Co-chairs Irwin King and Gabor Melli for their effort in staging this annual event. We also very grateful to the MDM Workshop Program Committee members who promptly provided high quality paper reviews during very restricted time period.

4. WORKSHOP CO-CHAIRS

Valery A. Petrushin, *The Nielsen Company*

Jia-Yu (Tim) Pan, *Google Inc.*

Cees G.M. Snoek, *University of Amsterdam*

5. PROGRAM COMMITTEE

Xian-Sheng Hua, *Microsoft Research*

Chong-Wah Ngo, *City University of Hong Kong*

Changsheng Xu, *Institute of Automation, Chinese Academy of Sciences*

Vasileios Mezaris, *CERTH*

Jinhui Tang, *National University of Singapore*

Qi Tian, *University of Texas at San Antonio*

Henning Muller, *University Hospitals of Geneva*

Yu-Gang Jiang, *Columbia University*

K. Selcuk Candan, *Arizona State University*

Chabane Djeraba, *LIFL - UMR CNRS*

William Grosky, *University of Michigan*

Shu-Ching Chen, *Florida International University*

Zhongfei (Mark) Zhang, *Binghamton University*

Maria Luisa Sapino, *University of Torino*

Vincent Tseng, *National Cheng Kung University*

Wei-Hao (Max) Lin, *Google Inc.*

Junfeng Pan, *Google Inc.*

Fatma Bouali, *University of Lille*

Latifur Khan, *University of Texas at Dallas*