

ACM SIGMM the First Workshop on Web-Scale Multimedia Corpus (WSMC09)

Benoit Huet
Eurecom
06904 Sophia-Antipolis - France
+33 (0)4.93.00.81.79
Benoit.Huet@eurecom.fr

Jinhui Tang
National University of Singapore
Computing 1, 117417, Singapore
+65-65161184
tangjh@comp.nus.edu.sg

Alex Hauptmann
Carnegie Mellon University
Pittsburgh, PA 15213 USA
+1-412-268-1448
alex@cs.cmu.edu

WORKSHOP OVERVIEW

The purpose of this workshop is to bring together researchers interested in the construction and analysis of Web Scale multimedia datasets and resources. The Workshop will provide a forum to consolidate key factors related to research on very large scale multimedia dataset such as the construction of dataset, creation of ground truth, sharing and extension of such resources in terms of ground truth, features, algorithms and tools etc. The Workshop will discuss and formulate action plan towards these goals.

Categories & Subject Descriptors: H.3.1 [Information Systems]: Information Storage and Retrieval, - Content Analysis and Indexing, H.3.7 Digital Libraries, H3.3. Information Storage and Retrieval,

General Terms: Algorithms, Experimentation, Measurement, Performance, Standardization,.

1. Introduction

Welcome to the 1st International Workshop on Web-Scale Multimedia Corpus (WSMC09) held at the ACM International Conference on Multimedia in Beijing, China.

Pivotal to many tasks in relation to multimedia research and development is the availability of sufficiently large dataset and its corresponding ground truth. Currently available datasets for multimedia research are either too small such as the Corel or Pascal datasets, too specific like the TRECVID dataset, or without ground truth, such as the several recent efforts by MIT and MSRA that gathered millions of Web images for testing. While it is relatively easy to crawl and store a huge amount of data, the creation of ground-truth necessary to systematically train, test, evaluate and compare the performance of various algorithms and systems is a major problem. For this reason, more and more research groups are individually putting efforts into the creation of such corpus in order to carry out research on Web-scale dataset. There is a need to unify these individual efforts into the creation of a unified web-scale repository which would benefit the entire multimedia research community.

Copyright is held by the author/owner(s).
MM'09, October 19–24, 2009, Beijing, China.
ACM 978-1-60558-608-3/09/10.

The purpose of the workshop is to bring together researchers interested in the construction and analysis of Web Scale multimedia datasets and resources. The Workshop provides a forum to consolidate key factors related to research on very large scale multimedia dataset such as the construction of dataset, creation of ground truth, sharing and extension of such resources in terms of ground truth, features, algorithms and tools etc. The Workshop discusses and formulates action plan towards these goals.

2. Topics of Interests

This workshop contains contributions from renowned multimedia researchers on the following topics:

1. Construction, Unification and Evolution of Corpus
2. Framework for sharing of dataset, ground truth, features, algorithms and tools
3. Web-scale Corpus analysis techniques
 - 3.1 Knowledge mining from web-scale multimedia corpus.
 - 3.2 Optimization techniques on web-scale multimedia data for efficiency.
 - 3.3 Techniques for web-scale content-based multimedia retrieval.
 - 3.4 Other related techniques.
4. Performance evaluation methodologies and standards

3. Summary of the Papers in the Workshop

The call for papers attracted submissions from Asia, North-America, Europe and Africa. The program committee only accepted 6 high quality papers, which was a difficult decision because some papers were very comparable in quality, but ultimately were rejected because they either had too much overlap with some accepted papers or extended the scope of the workshop too much. In addition, the program includes a panel on the topics addressed by the workshop and a keynote speech. We hope that these proceedings will serve as a valuable reference for multimedia researchers and developers as well as foster new research direction and results.

Looking over the papers accepted for the workshop, we start with a description of a large video dataset obtained from the creative commons. This paper by Foley on “Creating a web scale collection” discusses the problems and various issues with creating reference datasets at a large scale for researchers, and

offers some solutions with respect to collection building and copyright. Their intention is to support a next generation of benchmarking activities for content-based video operations.

This work is complemented by another dataset of web videos. In the paper entitled “MovieBase: A Movie Database for Event Detection and Behavioral Analysis” by Chua et al, they describe a 71 hour collection of full-length feature movies and YouTube videos that has been collected. This corpus comes with hand labeled annotations, as well as a set of low-level visual, motion and audio features together with baseline event detection results.

In order to improve the efficiency of manual labeling our next contribution proposes to utilize the visual features together with tag cloud information. In “Understanding Tag-Cloud and Visual Features for better Annotation of Concepts in NUS-WIDE Dataset” Gao and Chia propose an approach involving semi-manual annotation. A simple and intuitive method for selection improves an initial human-labeled annotation set by using either tag cloud information or visual features, via understanding of the WordNet taxonomy.

Kennedy et al., in their paper on “Reliable Tags Using Image Similarity: Mining Specificity and Expertise from Large-Scale Multimedia Databases” leverages the wisdom of the crowds. They show that when people agree that the tag applies to the image and when corresponding image pairs have high visual similarity in low-level feature space, the resulting concept classification method is likely to be much more reliable.

In the paper on a “Visual Tag Dictionary: Interpreting Tags with Visual Words” by Wang et al., the authors describe an effort to create a set of visual words as a dictionary for the “tag” labels. Similar to the phonetic lexicon of speech recognition systems, this visual word lexicon can be used to describe the visual content of images, and map these into a set of plausible tag labels.

Finally, when given a large collection of videos, it is important to categorize these automatically and meaningfully. This is addressed by Zhang et al in their paper on “Incremental web video categorization. They emphasize the need to build reliable concept classification algorithms and suggest that these will work best when there is a relatively small semantic gap between the low-level features and the abstract semantic concept to represent them. Their paper also provides an approach to incrementally learn the classification, as web scale data is to large to allow a one-shot complete learning approach.

4. Committee Members

Putting together WSMC’09 was a team effort. We are also heavily indebted to many individuals for their significant contribution. First of all, we would like to thank the authors, the keynote speaker and panelists for providing the content of the program. We thank the ACM Multimedia General Chairs (Wen Gao, Yong Rui and Alan Hanjalic) and Workshop Chairs (Ed Chang and Kiyoharu Aizawa) for giving us the opportunity to organize this workshop. We would like to express our gratitude to the program committee and external reviewers, who worked very hard in reviewing papers and providing suggestions for their improvements. We would also like to thank of Workshop Liaison (Tat-Seng Chua) for providing valuable input and guidance on crucial decisions.

We hope that you enjoy and benefit from this program,

WorkShop Co-Chairs

Benoit Huet, EURECOM
Jinhui Tang, National University of Singapore
Alex Hauptmann, Carnegie Mellon University

Workshop Liaison

Tat-Seng Chua, National University of Singapore

Technical Program Committee

Alberto Del Bimbo, University of Florence
Shih-Fu Chang, Columbia University
Liang-Tien Chia, Nanyang Technological University
Ajay Divakaran, Sarnoff Corporation
Xian-Sheng Hua, Microsoft Research Asia
Ramesh Jain, UC Irvine
Lyndon Kennedy, Yahoo!
Rainer Lienhart, Univ of Augsburg
Jiebo Luo, Eastman Kodak Company
Chong-Wah Ngo, City University of Hong Kong
Guo-Jun Qi, Univ. of Illinois at Urbana-Champaign
Shin'ichi Satoh, National Institute of Informatics
Alan Smeaton, Dublin City University
Cees Snoek, University of Amsterdam
Meng Wang, Microsoft Research Asia
Lexing Xie, IBM Watson Research Center
Lei Zhang, Microsoft Research Asia

External Reviewers:

Richang Hong, National University of Singapore
Xirong Li, University of Amsterdam
Stefan Romberg, Univ of Augsburg

References

[1] B. Huet, J. Tang, and A. Hauptmann. *Proceedings of the ACM SIGMM First Workshop on Web-Scale Multimedia Corpus*, in conjunction with the *ACM International Multimedia Conference*, Beijing, China, Oct. 2009.