Modeling and Representing Events in Multimedia

Vasileios Mezaris
Centre for Research and
Technology Hellas
Informatics and Telematics
Institute
Thermi 57001, Greece
bmezaris@iti.gr

Ansgar Scherp
University of Koblenz-Landau
Information Systems and
Semantic Web
Universitätsstr. 1
Koblenz, Germany
scherp@uni-koblenz.de

Ramesh Jain
University of California
Donald Bren School of
Information and Computer
Sciences
Irvine, CA 92697, USA
jain@ics.uci.edu

Mohan Kankanhalli
National University of
Singapore
School of Computing
Singapore 117590
mohan@comp.nus.edu.sg

Huiyu Zhou
Queen's University Belfast
ECIT Institute
Belfast, BT3 9DT, United
Kingdom
h.zhou@ecit.qub.ac.uk

Jianguo Zhang
University of Dundee
School of Computing
United Kingdom
jgzhang@computing.
dundee.ac.uk

Liang Wang
National Lab of Pattern
Recognition
Beijing, China
wangliang@nlpr.ia.ac.cn

Zhengyou Zhang
Microsoft Research
One Microsoft Way, Redmond
WA, USA
zhang@microsoft.com

ABSTRACT

This paper presents an overview of the Joint Workshop on Modeling and Representing Events (JMRE), which is held as part of ACM Multimedia 2011. JMRE is concerned with the understanding of events from multimedia, and with using events in order to better organize and consume multimedia.

Categories and Subject Descriptors

H.1.0 [**Information Systems**]: Models and Principles

General Terms

Design, Human Factors, Management

Keywords

Multimedia, events, objects, event detection, event-based applications, event models, sparse representation

1. MOTIVATION

Events are ubiquitous in multimedia. We find them in the contents that we all have created, shared or simply encountered on the Internet: for example, images from a concert that we recently attended or an interesting location that we visited, amateur video footage from a friend's wedding ceremony, official footage from an important game of our favorite football team, or eye-witness descriptions and infrared satellite images of a devastating tornado that made it to the headlines. All these contents come in different media types (e.g. image, video, text) created by different devices under different conditions. What is common though for all these media items is that they all have captured and convey information about real-life events. Efficiently and effectively understanding these heterogeneous events from the available multimedia contents and using them in order to better organize, share, retrieve and consume the contents in any possible way raises a significant challenge in today's world.

The goals of the Joint Workshop on Modeling and Representing Events (JMRE), held as part of ACM Multimedia 2011, are documenting and advancing the state of the art in relation to the above challenge. JMRE comprises two parts, the International Workshop on Events in Multimedia (EiMM) and the International Workshop on Sparse Representation for Event Detection in Multimedia (SRED). JMRE'11 also features a keynote talk by Prof. Nack, titled "Event and story - an intricate relationship" [6].

2. EVENTS IN MULTIMEDIA (EIMM)

This year's *International Workshop on Events in Multimedia*[1] (EiMM'11) is the third edition of the series of EiMM workshops, following the very successful workshops of this series, EiMM'09[2], held in Beijing, China, as part of ACM Multimedia 2009 [7], and EiMM'10[3], held in Florence, Italy, as part of ACM Multimedia 2010 [8].

The focus of the EiMM series lies on how to detect, model, and process domain-level events, and on how to develop applications that make use of domain-level events in the context of multimedia data. The EiMM'11 workshop brings together researchers from the different areas of the multimedia research community, who are interested in understanding the concept of events. It presents research progress in the

---

[1] http://www.uni-koblenz.de/confsec/eimm11/
[2] http://www.uni-koblenz.de/confsec/eimm09/
[3] http://www.uni-koblenz.de/confsec/eimm10/

areas of event detection and processing in multimedia data, event representation and event models, events in the context of Web 2.0, architectures for event management, and interactive event-based applications and tools.

Specifically, the papers included in the EiMM'11 program address the following topics:

- Event-based organization of personal photo archives. Treating the images acquired by a single person during a period of several months as a photo archive, Tankoyeu et al. [9] use various contextual cues for organizing the photos into different events.

- Recognition of events and sub-events in photo albums. Assuming that an album contains photos of a single event, Mattivi et al. [4] exploit the multitude of photos in the album for detecting its event type (e.g. Baseball, Concert, Wedding), and then further organize the album's photos into pre-defined sub-events (e.g. for Wedding, into ceremony, party-eating, etc.).

- Recognition of dynamic events (e.g. running, walking) in video using visual cues. Ballas and Delezoide [1] build on recent work on local feature trajectories, to investigate the impact of a new trajectory filtering scheme and two new trajectory descriptors to the detection of such video events.

- Recognition of complex events (e.g. feeding an animal) in video using audio cues. Mertens et al. [5] exploit non-speech audio features for building "acoustic super-models" that detect complex events from low-level audio features, and show that even using audio alone they can achieve high recognition rates.

- Event-based analysis of the game flow in sports video. In [2], mid-level and possibly incomplete annotations of individual shots (expressing camera views, play types, etc.) are used for finding and representing the overall game flow as the series of major events (e.g. "scoring", "turn-over") of the game. This approach is tested in the domain of American football.

- Lifelog event analysis. In [3], several aspects of lifelog research and related technical challenges are discussed, and a complete web-based event management service for collaborative lifelog event analysis is presented.

3. SPARSE REPRESENTATION FOR EVENT DETECTION IN MULTIMEDIA (SRED)

The *International Workshop on Sparse Representation for Event Detection in Multimedia*[4] (SRED'11) is focused on modeling, detection and recognition of events using sparsity analysis and applications that make use of sparse learning for event analysis in the context of multimedia. It presents research highlights in the area of sparse representation for event analysis and crowd behaviour analysis.

Specifically, the papers included in the SRED'11 program address the following topics:

- Amateur video summarisation. In [10], Wang et al. present a kernel based sparse representation approach

for directly summarising consumer videos, where a dictionary of subsequences is generated using the clustered frames with importance ranking.

- Unusual event detection in crowded scenes. Xu et al. introduce in [11] an approach for unusual event detection via sparse reconstruction of dynamic textures over an overcomplete basis set, with the dynamic texture described by local binary patterns from three orthogonal planes (LBPTOP).

4. ACKNOWLEDGMENTS

5. REFERENCES

[1] N. Ballas and B. Delezoide. Trajectories based descriptor for dynamic events annotation. In *Proc. JMRE'11 @ ACM Multimedia 2011*, Scottsdale, AZ, USA, Nov.-Dec. 2011.

[2] Y. Ding and G. Fan. Finding the Game Flow from Sports Video. In *Proc. JMRE'11 @ ACM Multimedia 2011*, Scottsdale, AZ, USA, Nov.-Dec. 2011.

[3] P. H. Kim and F. Giunchiglia. Lifelog Event Management: Crowd Research Case Study. In *Proc. JMRE'11 @ ACM Multimedia 2011*, Scottsdale, AZ, USA, Nov.-Dec. 2011.

[4] R. Mattivi, J. Uijlings, F. D. Natale, and N. Sebe. Exploitation of time constrains for (sub-)event recognition. In *Proc. JMRE'11 @ ACM Multimedia 2011*, Scottsdale, AZ, USA, Nov.-Dec. 2011.

[5] R. Mertens, H. Lei, L. Gottlieb, G. Friedland, and A. Divakarian. Acoustic Super Models for Large Scale Video Event Detection. In *Proc. JMRE'11 @ ACM Multimedia 2011*, Scottsdale, AZ, USA, Nov.-Dec. 2011.

[6] F. Nack. Event and story - an intricate relationship. Keynote talk. In *Proc. JMRE'11 @ ACM Multimedia 2011*, Scottsdale, AZ, USA, Nov.-Dec. 2011.

[7] A. Scherp, R. Jain, and M. Kankanhalli. Events in multimedia. In *Proc. ACM Multimedia 2009*, pages 1147–1148, Beijing, China, October 2009.

[8] A. Scherp, R. Jain, M. Kankanhalli, and V. Mezaris. Modeling, Detecting, and Processing Events in Multimedia. In *Proc. ACM Multimedia 2010*, pages 1739–1740, Firenze, Italy, October 2010.

[9] I. Tankoyeu, J. Paniagua, J. Stottinger, and F. Giunchiglia. Event Detection and Scene Attraction by Very Simple Contextual Cues. In *Proc. JMRE'11 @ ACM Multimedia 2011*, Scottsdale, AZ, USA, Nov.-Dec. 2011.

[10] Z. Wang, M. Kumar, J. Luo, and B. Li. Sequence-Kernel Based Sparse Representation for Amateur Video Summarization. In *Proc. JMRE'11 @ ACM Multimedia 2011*, Scottsdale, AZ, USA, Nov.-Dec. 2011.

[11] J. Xu, S. Denman, S. Sridaran, C. Fookes, and R. Rana. Dynamic Texture Reconstruction from Sparse Codes for Unusual Event Detection in Crowded Scenes. In *Proc. JMRE'11 @ ACM Multimedia 2011*, Scottsdale, AZ, USA, Nov.-Dec. 2011.

---

[4]`http://research.microsoft.com/en-us/um/people/zhang/SRED11/`