About Events, Objects, and their Relationships: Human-centered Event Understanding from Multimedia

Ansgar Scherp
ZBW – Leibniz Information Centre for Economics and Kiel University
Kiel, Germany
asc@informatik.uni-kiel.de

Vasileios Mezaris
Centre for Research and Technology Hellas
Thessaloniki, Greece
bmezaris@iti.gr

Bogdan Ionescu
University Politehnica of Bucharest
Bucharest, Romania
bionescu@imag.pub.ro

Francesco De Natale
University of Trento
Trento, Italy
francesco.denatale@unitn.it

ABSTRACT

HuEvent'15 is a continuation of previous year's successful workshop on events in multimedia. It focuses on the human-centered aspects of understanding events from multimedia content. This includes the notion of objects and their relation to events. The workshop brings together researchers from the different areas in multimedia and beyond that are interested in understanding the concept of events.

1. INTRODUCTION

Events can be found from life-log applications to emergency response systems as well as in domains like cultural heritage, news, sports, and surveillance. Thus, we can understand events as natural abstraction of human experience [7]. The different applications and domains make use of different methods and approaches for detecting, representing, and using events. However, they share the common notion of considering events as important entities.

Events are generally understood as perduring entities that unfold over time. They are occurrences in which humans participate and may be subject to discussions and interpretations by humans. In contrast, objects are enduring entities that unfold over space. While some consider objects as 4D entities, i.e., extending across time just as they do in space, others consider both events and objects as first class entities that require each other.

Key research questions of the workshop series are:

- Discovering events and/or objects from media assets;
- Understanding objects' role in events;
- Understanding objects' interconnectivity through events;
- Understanding peoples' actions during events;
- Analysis of humans' reactions/emotions/comments to events and/or objects;
- Understanding humans' event interaction and its evolution in time;
- Analysis of human opinion and bias on event perception;
- Analysis of human factors for event evolution, spreading, and interpretation;
- Use of human behaviors for event prediction and trend analysis ("in" and "on" the event);
- Multiple views and multiple user event analysis;
- Time synchronization of multi-user event media;
- Multimedia event detection and recounting;
- Deep learning for multimedia event analysis and mining;
- Event-based processing and analysis of Big Multimedia Data;
- Event-based multimedia organization, clustering, summarization, and storytelling;
- Multimedia for human-centered event understanding;
- Novel applications of events and multimedia.

2. WORKSHOP GOAL

The goal of this workshop is to present and discuss the human-centered understanding of different aspects and notions of events and objects. This includes methods for detecting activities and low-level events and objects from media content and other sensory data. It also targets solutions and approaches for detecting and modeling the relationships between events and objects, as well as for organizing media according to events.

HuEvent'15, October 30, 2015, Brisbane, Australia.
ACM 978-1-4503-3748-9/15/10.
DOI: http://dx.doi.org/10.1145/2733373.2806413.

We aim to bring together researchers from the different areas in multimedia and beyond that are interested in understanding the concept of events. We invite original work in the areas of event modeling, detection of events from multimedia data, processing of events, organization of multimedia data using events as unifying mechanism, and applications of these techniques.

3. WORKSHOP HISTORY

The 2nd ACM International Workshop on Human-centered Event Understanding from Multimedia[1] (HuEvent'15) at the 23rd ACM Multimedia Conference in Brisbane, Australia is a continuation of a successful series of previous events. Most notably, it is a successor of the 1st Workshop on Human-centered Event Understanding from Multimedia[2] (HuEvent'14) that was co-located with ACM Multimedia in Orlando, Florida, USA [7].

However, the workshop finds its roots in earlier activities: There have been three successful Events in Multimedia (EiMM) workshops that were held in conjunction with the previous editions of ACM Multimedia in Beijing, China (2009) [3], Florence, Italy (2010) [5], and Scottsdale, AZ, USA (2011) [2]. In those editions, the focus was on events only and not on an a human-centered approach of event understanding that also includes the notion of objects and their relation to events - as it is pursued in this workshop. The Workshop on Event-based Media Integration and Processing[3] held in 2013 at ACM Multimedia in Barcelona is also related but focused on events and their extraction from multimedia content.

From these activities, also different special issues on the topic of events and objects in multimedia emerged [1, 4]. Finally, we released an extensive survey on modeling and indexing events in multimedia [6].

4. WORKSHOP PROGRAMME

This year's workshop programme is organized into a keynote, a set of paper presentations, and a final discussion.

The keynote is held by Cees Snoek[4] on the topic of "Recognizing events in videos without examples". The talk presents recent progress on recognizing events in videos, without the need for examples. The key to event recognition in such a challenging setting is to have a lingual video representation. Three lingual representations for zero-example event recognition are highlighted: covering concept, tag, and sentence embedding.

The accepted research papers at the workshop are continuing the discussions focusing on the following topics: The first paper, "Using Photo Similarity and Weighted Graphs for the Temporal Synchronization of Event-Centered Multi-User Photo Collections", addresses the issue of temporal synchronization of photo collections that have been created during the same event by different users using different (unsynchronized) devices. The method proposed by the authors employ multiple similarity measures to identify pairs of similar photos and then temporally align the photo collections by traversing a graph, whose nodes represent the collections, and edges represent the similar photo pairs between collections.

The second paper, "Media Synchronization and Sub-Event Detection in Multi-User Image Collections", continues the idea of media synchronization by conducting a thorough evaluation of the performance of several visual-based image synchronization techniques; but goes beyond and applies it to sub-event detection. Common clustering techniques are experimented for the detection of sub-events in the presence of synchronization misalignment.

Finally, the paper "Discovering Commonness and Specificness for Human Action Recognition" brings into discussion another dimension of human-centered event understanding, namely human action recognition. A discriminative dictionary, learning-based method that is specifically adapted to recognize commonness and specificness among different action classes is introduced. Experimental validation on standard benchmarking datasets show promising results for this approach.

Like in the previous year, the workshop programme is a kick-off for a final discussion. Topics addressed are among others the different understandings of events and objects, approaches for harmonization, and future research.

[1] http://huevent15.uni-kiel.de/en
[2] http://huevent14.uni-kiel.de/en
[3] http://ebmip.disi.unitn.it
[4] http://www.ceessnoek.info/

5. REFERENCES

[1] V. Mezaris, A. Scherp, R. Jain, and M. S. Kankanhalli. Real-life events in multimedia: detection, representation, retrieval, and applications. *Multimedia Tools Appl.*, 70(1):1–6, 2014.

[2] V. Mezaris, A. Scherp, R. Jain, M. S. Kankanhalli, H. Zhou, J. Zhang, L. Wang, and Z. Zhang. Modeling and representing events in multimedia. In K. S. Candan, S. Panchanathan, B. Prabhakaran, H. Sundaram, W. chi Feng, and N. Sebe, editors, *Proc. of the 19th Int. Conf. on Multimedia 2011, Scottsdale, AZ, USA, November 28 - December 1, 2011*, pages 613–614. ACM, 2011.

[3] A. Scherp, R. Jain, and M. S. Kankanhalli. Events in multimedia. In W. Gao, Y. Rui, A. Hanjalic, C. Xu, E. G. Steinbach, A. El-Saddik, and M. X. Zhou, editors, *Proc. of the 17th Int. Conf. on Multimedia 2009, Vancouver, British Columbia, Canada, October 19-24, 2009*, pages 1147–1148. ACM, 2009.

[4] A. Scherp, R. Jain, and M. S. Kankanhalli. Introduction to the special issue of the multimedia tools and applications journal on events in multimedia. *Multimedia Tools Appl.*, 57(1):127–129, 2012.

[5] A. Scherp, R. Jain, M. S. Kankanhalli, and V. Mezaris. Modeling, detecting, and processing events in multimedia. In A. D. Bimbo, S.-F. Chang, and A. W. M. Smeulders, editors, *Proc. of the 18th Int. Conf. on Multimedia 2010, Firenze, Italy, October 25-29, 2010*, pages 1739–1740. ACM, 2010.

[6] A. Scherp and V. Mezaris. Survey on modeling and indexing events in multimedia. *Multimedia Tools Appl.*, 70(1):7–23, 2014.

[7] A. Scherp, V. Mezaris, B. Ionescu, and F. G. B. D. Natale. Huevent'14: 2014 workshop on human-centered event understanding from multimedia. In K. A. Hua, Y. Rui, R. Steinmetz, A. Hanjalic, A. Natsev, and W. Zhu, editors, *Proc. of the ACM Int. Conf. on Multimedia, MM '14, Orlando, FL, USA, November 03 - 07, 2014*, pages 1253–1254. ACM, 2014.