Human-centered Event Understanding from Multimedia

Ansgar Scherp
Kiel University and Leibniz
Information Center for
Economics
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@alpha.
imag.pub.ro

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

ABSTRACT

This workshop 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 are everywhere! We find them in life-log applications and 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. 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.[1]

---

[1] http://www.slideshare.net/ascherp/slideshare-events-in-multimedia-theory-model-application-v07, last accessed: July 11, 2014.

2. WORKSHOP GOAL

The goal of this workshop is to present and discuss the 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. Finally, we invite submissions of novel applications that are based on the notion of events and objects and that make use of events and objects as first-class entities.

We 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 PROGRAMME

The workshop programme is organized into a keynote, a set of paper presentations, and a final discussion. The keynote is held by Ramesh Jain, University of California at Irvine, CA, USA with the title the "Objective Self".

The submissions accepted for the workshop are organized in two sessions[2]: A first set of papers deals with the detection of events in video. In this case, an event is most often defined as a complex interaction pattern that involves people and objects and is depicted in the video. In "Entity centric Feature Pooling for Complex Event Detection", the authors tackle the problem of detecting such events in YouTube-like videos. Their approach is based on understanding the spatial arrangement of people and objects that participate to an event. In "Skeleton-augmented Human Action Understanding by Learning with Progressively Refined Data", the authors focus on the human actors of the event, specifically

---

[2] http://www.huevent14.uni-kiel.de/en/technical-programme, last accessed: August 8, 2014.

on inferring skeletons that describe the human actions depicted in the video. In "Using Minute-by-Minute Match Report for Semantic Event Annotation in Soccer Video", the event-based analysis of video belonging to a very specific domain (soccer matches) is addressed. The soccer video is annotated using the match report; however, without assuming that perfect time synchronization exists between the video content and textual description of the match. Considering again a very specific but radically different domain, the authors of "Event Understanding in Endoscopic Surgery Videos" examine the segmentation of surgery video in sub-events.

The second set of papers takes a somewhat different view on events: Instead of looking at events as short-term human actions or human-object interactions that are captured in video, they focus on broader personal and social events or on the user's interaction with the image / video content that relates to such events. The latter interaction can provide useful cues for the event-based organization of the media items. In "Concept-based image clustering and summarization of event-related image collections", the authors deal with the problem of summarizing image collections that correspond to a single event. To this end, they propose using trained visual concept detectors in combination with clustering methods. In "Sentiment Flow for Video Interestingness Prediction", the authors develop a method for predicting how interesting a video is, employing a mid-level sentiment representation for the video content. In "User Emotion Sensing in Search Process based on Chromatic Sensation", the authors propose a model for sensing the user's emotions by examining the colors (of photos, icons etc.) browsed or selected by the user. In "Investigating Human Factors in Forgery Detection Process", the authors deal with the difficult problem of detecting if an image is forged or not (which is a question that often arises when an image of an event such as a natural disaster or a war incident surfaces on the Web). They conduct a subjective evaluation to investigate human factors that are associated with this problem.

Finally, in the position paper titled "On the Personalization of Event-based Systems", the authors describe their position about personalization as a paradigm shift and discuss its relation to the event-based processing and organization of signals and digital content.

The keynote, author presentations, and position paper are used as kick-off for a final discussion about the different understandings of events and objects from the various fields. The questions discussed during the workshop include:

- How do the different works on events and objects relate to each other?

- How can they be harmonized and brought together under the roof of a common model?

- Should events and objects be considered first class entities?

- Which different understandings of relations between objects and events exist?

4. WORKSHOP HISTORY

The workshop finds its roots in earlier activities by some of the organizing members. For example, there have been three successful "Events in Multimedia" workshops which 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 last year at ACM Multimedia in Barcelona, Spain 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].

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, *Proceedings of the 19th International Conference 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, *Proceedings of the 17th International Conference 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, *Proceedings of the 18th International Conference 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.

---

[3]`http://ebmip.disi.unitn.it`, last accessed: July 11, 2014.