

# 4<sup>th</sup> ACM/IEEE ARTEMIS 2013

## International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams

Anastasios Doulamis  
Technical University of Crete  
Chania Crete, Greece  
Polytechnioupolis Campus  
Tel: + 30 28210 37430  
adoulam@cs.ntua.gr

Nikolaos Doulamis  
National Technical  
University of Athens, Greece  
9, Heroon Polytechniou Street  
Tel: + 30 210 772 2678  
ndoulam@cs.ntua.gr

Marco Bertini  
University of Florence  
Media Integration and  
Communication Center, Italy  
+39 055 4796 468  
bertini@dsi.unifi.it

Jordi Gonzalez  
Universitat Autònoma de Barcelona, Spain  
08193 - Bellaterra, Barcelona  
Fax: +34 935811670  
poal@cvc.uab.es

Thomas Moeslund  
University of Aalborg,  
Faculty of Engineering and Science  
Sofieendalsvej 11, Aalborg, Denmark  
Tel: 45 9940 8787  
tbm@create.aau.dk

### ABSTRACT

In this paper, we give a short summary of the papers proposed in ACMARTEMIS 2013 which is held in Barcelona Spain in conjunction with ACM Multimedia. The workshop handles the areas of features analysis both at low and high level for efficient events detection, retrieval of multimedia events and objects and video synchronization issues and also events and behavior recognition from visual data. All papers were classified into three sessions of a single track workshop. The first session named "Video Features and Scene Analysis" includes articles that handle low level and high level visual analysis appropriate for event detection. The second session entitled "Retrieval of Multimedia Objects/Events" applies schemes for media data retrieval and video synchronization. Finally the third session "Analysis of Visual Events" describes algorithms for detecting actions, behaviors and events in complex visual scenes.

### Categories and Subject Descriptors

I.4.0 [Image Processing and Computer Vision: General]-*image processing software*; H.3.0 [Information Storage and Retrieval: General];

### Keywords

Analysis of events, retrieval of actions, tracking, video analysis.

### 1. INTRODUCTION

Recently, it can be argued that the intelligence behind many pattern recognition and computer vision systems is mainly

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author(s).  
MM '13, Oct 21-25 2013, Barcelona, Spain  
ACM 978-1-4503-2404-5/13/10.  
<http://dx.doi.org/10.1145/2502081.2503825>

focused on two main approaches; (i) extraction of smart features able to efficiently represent the rich visual content and (ii) adoption of non-linear and adaptable (semi-supervised) learning strategies able to fill the gap between the extracted low level features and the high level concepts, humans use to perceive the content. The feature extraction is a data dimensionality reduction strategy that addresses the difficulty that learning complexity grows exponentially upon a linear increase in the dimensionality of data. It is also clear that extraction of representational features is a challenging and application-dependent process. Non-representative features significantly affect the recognition accuracy, especially for complex and dynamic environments even though they are processed by highly non-linear feature transformation models. Thus, the main goal of this workshop is to seek original articles in the area of multimedia research in the direction of detecting and recognizing high level concepts and their respective spatio-temporal and causal relations in order to identify semantic video activities.

The research topics covered in this workshop are very important for the multimedia community since automatic extraction of events in video data is still one of the main challenging and open research issues in the specific area. This is verified by the high success of all the previous ARTEMIS workshops held in Florence in 2010 in conjunction with ACM Multimedia, Barcelona in 2011 in conjunction with ICCV and in Florence again in 2012 in conjunction with ECCV. In these workshops, we need also to stress the success story of the relevant AREA 2008 workshop held in Vancouver Canada in conjunction with ACM Multimedia 2008.

The success of this workshop was so great that the Editor in Chief of the Multimedia Tools and Applications Journal (Springer) has asked twice the workshop organizers to prepare a special issue in this journal. The issue has been successfully completed with high quality papers and finally two volumes have been dedicated to this issue by the editor. The importance is also verified by other related events in high quality journals such as the recent special issues of the IEEE Transactions on Multimedia with title of "Semantic Image and Video Indexing in Broad Domains"

and the recent call for papers of the IEEE Transaction on Circuits & Systems for Video Technology with title of “Event Analysis in Videos [1], [2], [3].

## 2. Previous ARTEMIS Workshops

The goal of this proposal is to continuation the success story of ARTEMIS workshops launched from 2010 and then in conjunction with ACM Multimedia, IEEE International Conference of Computer Vision (ICCV) and European Conference of Computer Vision (ECCV). All these conferences organized as full day workshops in which 12-16 research papers were presented, carefully selected after a peer review process, and with an average success rate of about 50%. We should mention that three of the previous workshops were also related with journal papers publications at Computer Vision Image Understanding Journal and Multimedia Tools and Applications.

ARTEMIS 2010: In total, 27 papers from 17 different countries were submitted. To ensure the high quality, all submissions were refereed strictly by at least two experts in the related fields based on the criteria of originality, significance, quality and clarity. As a result, 14 papers (52%) were accepted. Out of these 14 papers, 9 were oral presentations (33%) and 5 were posters and 2 demos.

ARTEMIS 2011: In total, 24 papers from 9 different countries were submitted. All submissions were reviewed by three experts in related fields, and reviews were than assessed by the organizing chairs. As a result, 12 papers (50%) were accepted, all oral.

ARTEMIS 2012: In total, 32 papers from 10 different countries were submitted. All submissions were reviewed by three experts in related fields, and reviews were than assessed by the organizing chairs. As a result, 12 papers (50%) were accepted as oral and 4 as posters. They were presented as oral in 4 sessions. We should mention that ARTEMIS was the second most active ECCV workshop in terms of submissions.

## 3. SUMMARY OF ARTEMIS 2013 WORKSHOP

We have received works from all over the world. All papers were reviewed by at least three (3) Reviewers. However, most of the papers were reviewed by four or even five Reviewers. Only papers that are above the borderline were accepted for publication. We have organized the papers into three sessions according to the thematic areas covered.

### 3.1 Session I “Video Features and Scene Analysis”

The first session is dedicated to “Video Features and Scene Analysis” and includes algorithms, techniques and methods for video scene analysis. In particular, the first article deals with the application of a new algorithm that modifies optical flow on the use of textural features. Then, we introduce methods for hand gesture recognition on exploiting depth sensors from Kinect™ cameras. The third paper describes a foreground detection method for traffic scenes within complex real-world urban environments. Finally, the fourth article of this session discusses the introduction of transfer learning for re-identifying persons from multiple cameras.

### 3.2 Session II “Retrieval of Multimedia Objects/Events”

The second session deals with “Retrieval of Multimedia Objects/Events”. This session includes two papers. The first proposes a new non-parametric method for clustering image content with respect to the semantics. The clustering exploits visual features like the SIFT transform while cultural heritage objects are surveyed. The second and last paper of this session describes a methodology for synchronizing two video sources which capture the same scene from different views.

### 3.3 Session III “Analysis of Visual Events”

The third session describes events detection and abnormal behavior recognition from complex visual data. The title of this session is “Analysis of Visual Events”. We start with an article which applies methods for abnormal behavior recognition from complex visual scenes. Then, we propose techniques for recognizing complex behaviors based on a human constrained descriptor and adaptable neural networks. The third article applies computer vision tools for maritime detection in outdoor environments of sea ports. Finally, the last not least paper of the workshop describes a methodology for cross-domain traffic scene understanding by the use of motion models.

## 4. ACKNOWLEDGMENTS

This event is supported by the European funded projects, *EXPERIMEDIA*, “Experiments in live social and networked MEDIA experiences”, (<http://www.experimedia.eu/>), *4D-CH-World*, “Four Dimensional Cultural Heritage World”, [www.4d-ch-world.eu](http://www.4d-ch-world.eu), and the Greek-Cypriot Collaborative project *POSEIDON* “Development of an Intelligent System for Coast Monitoring using Camera Arrays and Sensor Networks” (<http://www.poseidonproject.eu/>) and *e-Park*, “Exploitation of New Technological Trends in Management and Payment of Public Parkings,” [www.e-park.eu](http://www.e-park.eu). In addition the Greek National Projects *Viopolis* ([www.viopolis.gr](http://www.viopolis.gr)) and *i-Promotion* support this event.

## 5. REFERENCES

- [1] Shih-Fu Chang, Jiebo Luo, Steve Maybank, Dan Schonfeld, Dong Xu, “Event Analysis in Videos,” *IEEE Trans. On Circuits and Systems for Video Technology*, Nov. 2008.
- [2] A.D., Doulamis, L. Van Gool, M. Nixon, N.D. Doulamis, T.A. Varvarigou, “Guest Editorial: Analysis and retrieval of events/actions and workflows in video streams,” *Multimedia Tools and Applications* 50 (1), pp. 1-6, 2010.
- [3] A. Doulamis, L. van Gool, M. Nixon, T. Varvarigou, Doulamis, “First ACM international workshop on analysis and retrieval of events, actions and workflows in video streams,” *Proceedings of the 2008 ACM International Conference on Multimedia*, 2008.