

## Analysis and retrieval of events/actions and workflows in video streams

Anastasios D. Doulamis · Luc van Gool · Mark Nixon · Nikolaos D. Doulamis · Theodora A. Varvarigou

Published online: 1 April 2010  
© Springer Science+Business Media, LLC 2010

The contemporary trend in large scale enterprises, like public infrastructures or industrial plants, is to use architectures and platforms for *video surveillance* through installation of a network of cameras in critical sites and monitoring the captured video data. Video surveillance provides *quality assurance of services and/or products* (adherence to predefined procedures of a production), *traffic management control* (for high-dense urban areas), *security/safety* (prevention from actions that may lead to hazardous situations), *crisis management in public areas* (e.g., train stations, airports), or a series of other applications of high industrial/social impact. However, the current commercial video surveillance systems support mostly manual supervision, making them both inefficient and subjective.

The inefficiency stems from the fact that it is impossible for a human to continually concentrate on monitors that display different activities in different areas. The subjectivity arises from the fact that humans usually interpret the same visual information differently under different conditions. For this reason, methods, tools and algorithms that aim to detect and recognize high level concepts and their respective spatio-temporal and causal relations (to identify semantic video activities, actions and procedures) have been in the focus of the research community over the last years and many research efforts have been paid within the computer vision and machine learning communities.

The traditional approaches for event detection in videos assume well structured environments and they fail to operate in largely unsupervised way under adverse and uncertain conditions from those on which they have been trained. Another drawback of current methods, is the fact that they focus on narrow domains using specific concept detectors such as “human faces”, “cars”, “buildings”. This special issue seeks original high innovative research in the area of self configurable cognitive video supervision in several domains.

This Call for Papers was very well received, and we collected several high quality papers. A severe review process led to a selection of a number of very good papers, so that the Editor-in-Chief agreed to devote two Special Issues in order to include all the accepted papers.

This special issue consists of eleven (11) papers that cover most of the areas of computer vision research in towards events, actions and workflows analysis. The papers can be

---

A. D. Doulamis (✉) · L. van Gool · M. Nixon · N. D. Doulamis · T. A. Varvarigou  
Technical University of Crete, Chania, Greece  
e-mail: adoulam@cs.ntua.gr

classified in three generic categories, papers that apply low-level image analysis for events' detection, articles suitable for events' classification and works dealing with knowledge representation tools for understanding the visual events.

The first category articles apply algorithms for features extraction and object tracking, all appropriate for identifying events and actions in videos. More specifically, the paper of A. Kriechbaum et al. applies a framework for unsupervised segmentation of moving objects in image sequences that does not require any restriction on the video content. The approach extracts the moving objects using a mesh-based combination of colour segmentation and motion segmentation using a feature point tracking. D. Roth et al. present a visual object tracking method which is then assessed for events' detection. The tracker is able to “detect and monitor” multiple object classes in non-controlled visual environments using Bayesian per-pixel classification in real-time. Furthermore, a general new event metric is used to compare the proposed tracking scheme with the other tracking methods against ground truth of multiple public datasets. The next paper applies dynamic trackers' re-initialization schemes for improving their performance in case of complicated visual scenarios (A. Doulamis). Finally, a humans' gait recognition algorithm has been proposed by the work of I. Bouchrika et al.

The second category of articles includes schemes for analyzing and classifying the events. In particular, C. Simon proposes a method for recognizing visual events using a decision tree mechanism. On the other hand, a rule-based system that combines image/visual analysis for identifying events in metro stations is proposed by B. Krausz et al. Another specific event visual analysis system is presented by Wei-Ta Chu, suitable for tennis court events. The system combines knowledge of tennis rules with specialized image/video analysis algorithms. The work of N. Doulamis applies an innovative angle spectrum algorithm for detecting vehicles and then extracting information about their behaviors in the roads as far as the car safety is concerned. Finally, implicit human actions annotation on large multimedia database is the theme of the last paper in this class.

Finally, in the last category we include knowledge representation papers for visual events detection. In particular, Minh-Son Dao et al. proposes a new spatio-temporal method for adaptively detecting events based on Allen temporal algebra. The temporal information is captured by presenting events as the temporal sequences using a lexicon of non-ambiguous temporal patterns. Finally, Senem Velipasalar et al. presents a tool for spatiotemporal event detection that lets users specify semantically high-level and composite events, and then detects their occurrences automatically. Events can be defined on a single camera view or across multiple camera views.

**Guest Editors:**

**Anastasios Doulamis**  
**Luc van Gool**  
**Mark Nixon Professor**  
**Nikolaos D. Doulamis**  
**Prof. Theodora A. Varvarigou**

**Papers Order in the Special Issue**

Please find below the preferred presentation order for the accepted papers in this special issue.

---

No. order	Corresponding author	Paper title
1	Andreas Kriechbaum	A Framework for Unsupervised Mesh based Segmentation of Moving Objects
2	Daniel Roth	Multi-Object Tracking Evaluated on Sparse Events
3	Anastasios Doulamis	Dynamic Tracking Re-Adjustment: A Method for Automatic Tracking Recovery in Complex Visual Environments
4	Imed Bouchrika	Performance Analysis for Automated Gait Extraction and Recognition in Multi-Camera Surveillance
5	Cédric Simon	Visual Event Recognition using Decision Trees
6	Barbara Krausz	MetroSurv: Detecting Events in Subway Stations
7	Wei-Ta Chu	Modeling Spatiotemporal Relationships between Moving Objects for Event Tactics Analysis in Tennis Videos
8	Nikolaos Doulamis	Coupled Multi-Object Tracking and Labeling for Vehicle Trajectory Estimation and Matching
9	Klimis Ntalianis,	Human Action Annotation, Modeling and Analysis based on Implicit User Interaction
10	Minh-Son Dao	A New Spatio-Temporal Method for Event Detection and Personalized Retrieval of Sports Video
11	Senem Velipasalar	Detection of User-defined, Semantically High-level, Composite Events, and Retrieval of Event Queries

---



**Anastasios Doulamis** received the Diploma degree in Electrical and Computer Engineering from the National Technical University of Athens (NTUA) in 1995 with the highest honor. In 2000, he has received the PhD degree in electrical and computer engineering from the NTUA. From 1996–2000, he was with the Image, Video and Multimedia Lab of the NTUA as research assistant. From 2001 to 2002, he serves his mandatory duty in the Greek army in the computer center department of the Hellenic Air Force, while in 2002, he join the NTUA as senior researcher. His PhD thesis was supported by the Bodosakis Foundation Scholarship. In 2006, he is Assistant professor in the Technical University of Crete in the area of multimedia systems. Dr. Doulamis has received several awards and prizes during his studies, including the Best Greek Student in the field of engineering in national level in 1995, the Best Graduate Thesis Award in the area of electrical engineering and several prizes from the National Technical University of Athens, the National Scholarship Foundation and the Technical Chamber of Greece. In 1997, he was given the NTUA Medal as Best Young Engineer. He received the best PhD thesis award by the Thomaidion Foundation. He has served as Program Chair in the ACM Analysis and Retrieval of Events, Actions Workflows in Video Streams (AREA) 2008 conference held in Vancouver Canada in Conjunction with ACM Multimedia. In 2001, he served as technical program chairman of the VLBV'01. He has also served as program committee in several international conferences and workshops, such as ICPR, EUSIPCO, Robotics and Applications, Intelligent Computing, etc. He is reviewer of IEEE journals and conferences as well as and other leading international journals. He is author of more than 170 papers in the above areas, in leading international journals and conferences, most of them in IEEE archive. His research interests include, non-linear analysis, neural networks, multimedia content description, intelligent techniques for video processing.



**Luc Van Gool** is obtained the master and PhD degrees in Electrical Engineering at the Katholieke Universiteit Leuven, Belgium, resp.in 1982 and 1991. He is a full professor for Computer Vision at both K. U.Leuven and ETH Zurich. With his two research labs, he focuses on object recognition, tracking and gesture analysis, and 3D acquisition and modeling. Luc Van Gool is an associate editor of several journals,

incl. Int. J. on Computer Vision and IEEE Trans. on Pattern Analysis and Machine Intelligence. He is an editor-in-chief of Foundations and Trends in Computer Graphics and Vision. He also is a co-founder of the spin-off companies Eyetronics, GeoAutomation, kooaba, Procedural, and eSaturnus. He received several best paper awards, incl. at ICCV 1998, CVPR 2007, and ACCV 2007, and ICRA 2009.



**Mark S. Nixon** is the Professor in Computer Vision at the University of Southampton UK. His research interests are in image processing and computer vision. His team develops new techniques for static and moving shape extraction which have found application in automatic face and automatic gait recognition and in medical image analysis. His team were early workers in face recognition, later came to pioneer gait recognition and more recently joined the pioneers of ear biometrics. Amongst research contracts, he was Principal Investigator with John Carter on the DARPA supported project Automatic Gait Recognition for Human ID at a Distance. He has chaired or had major involvement in many conferences (BMVC, AVBPA, IEEE Face and Gesture, ICPR, ICB, IEEE BTAS) and given many invited talks. His vision textbook, co-written with Alberto Aguado, Feature Extraction and Image Processing (Academic Press) reached 2nd Edition in 2008. With Tieniu Tan and Rama Chellappa, their book Human ID based on Gait which is part of the new Springer Series on Biometrics was published in 2005. Dr. Nixon is a member of the IEEE and Fellow IET and FIAPR.



**Nikolaos D. Doulamis** received the Diploma degree in Electrical and Computer Engineering from the National Technical University of Athens (NTUA) in 1995 with the highest honor and the PhD degree in electrical and computer engineering from NTUA in 2000. He joined the Image, Video and Multimedia Lab of

NTUA in 1996 as research assistant. His PhD thesis was supported by the Bodosakis Foundation Scholarship. Since 2005, he is visiting professor at the National Technical University of Athens. Dr. Doulamis was awarded as the Best Greek Student in the field of engineering in national level by the Technical chamber of Greece in 1995. In 1996, he was received the Best Graduate Thesis Award in the area of electrical engineering. During his studies he has also received several prizes and awards from the National Technical University of Athens, the National Scholarship Foundation and the Technical Chamber of Greece. In 1997, he was given the NTUA Medal as Best Young Engineer. He has also served as program committee in several international conferences and workshops. He was given the Thomaidion Foundation best journal paper award. He is reviewer of IEEE journals and conferences as well as and other leading international journals. His research interest include video transmission, content-based image retrieval, summarization of video sequences and intelligent techniques for video processing.



**Prof. Theodora A. Varvarigou** received the B. Tech degree from the National Technical University of Athens, Athens, Greece in 1988, the MS degrees in Electrical Engineering (1989) and in Computer Science (1991) from Stanford University, Stanford, California in 1989 and the Ph.D. degree from Stanford University as well in 1991. She is now professor at the NTUA. Prof. Varvarigou has great experience in the area of information processing, semantic web technologies, media streaming and casting, scheduling over distributed platforms. In this area, she has published more than 150 papers in leading journals and conferences. She has participated and co-ordinated several EU funded projects. Her research interests include grid computing, knowledge technologies and multimedia data processing.