

## Introduction to the special issue on “semantic multimedia”

**Yannis Avrithis · Noel E. O’Connor ·  
Steffen Staab · Raphael Troncy**

Published online: 18 June 2008  
© Springer Science + Business Media, LLC 2008

It is the target of this special issue to collect and report on recent work that aims at narrowing the large disparity between the low-level descriptors that can be computed automatically from multimedia content and the richness and subjectivity of semantics in user queries and human interpretations of audiovisual media—the so-called Semantic Gap.

Research in this area is important because the amount of information available as multimedia for the purpose of entertainment, security, teaching or technical documentation is overwhelming but the understanding of the semantics of such data sources is very limited. This means that the ways in which it can be accessed by users is also severely limited and so the full social or economic potential of this content cannot be realised.

Addressing the grand challenge posed by the Semantic Gap required a multi-disciplinary approach and this is reflected in recent research in this area. Thus, by its very nature, this special issue is targeting an interdisciplinary community. For example, multimedia researchers who deal with signal processing, computer vision, pattern recognition, multimedia analysis, indexing, retrieval and management of ‘raw’ multimedia data are increasingly leveraging methods and tools from the Semantic Web field by considering how

---

Y. Avrithis (✉)  
NTUA, Athens, Greece  
e-mail: [iavr@image.ntua.gr](mailto:iavr@image.ntua.gr)  
URL: <http://www.image.ntua.gr/~iavr/>

N. E. O’Connor  
Dublin City University, Dublin, Ireland  
e-mail: [oconnorn@eeng.dcu.ie](mailto:oconnorn@eeng.dcu.ie)  
URL: <http://www.eeng.dcu.ie/~oconnorn/>

S. Staab  
University of Koblenz, Koblenz, Germany  
e-mail: [staab@uni-koblenz.de](mailto:staab@uni-koblenz.de)  
URL: <http://isweb.uni-koblenz.de>

R. Troncy  
CWI, Amsterdam, The Netherlands  
e-mail: [Raphael.Troncy@cwi.nl](mailto:Raphael.Troncy@cwi.nl)  
URL: <http://www.cwi.nl/~troncy/>

to enrich their methods with explicit semantics. Conversely, Semantic Web researchers consider multimedia as an extremely fruitful area of application for their methods and technologies and are actively investigating how to enhance their techniques with results from the multimedia analysis community. A growing community of researchers is now pursuing both approaches in various high-profile projects across the globe. However, it remains difficult for both sides of the divide to communicate and learn from each other. It is our hope that this special issue will go some way towards easing this by presenting recent state-of-the-art results from both communities.

To provide an outlet and to improve communication and collaboration between the two fields this special issue on “Semantic Multimedia” is actually *a double special issue* brought forward by Elsevier’s *Journal of Web Semantics* as well as by Springer’s *Multimedia Tools and Applications*. Contributions were sought from both communities and for both journals in a joint call. The 37 papers submitted to the joint special issue demonstrate the large interest in semantic multimedia. Only nine of these papers met the rigorous standards of the two journals and were accepted. We thank all authors for their contributions, as they have exerted great enthusiasm for the research topic and have exposed interesting research—sometimes at early stages of development—that will bear plenty of fruit over the next few years. We also thank all reviewers, from both communities, who helped ensure the highest possible quality in all accepted papers.

In this issue of *Multimedia Tools and Applications* one may find six contributions. They mainly target the use of explicit semantics, ontologies and semantic models to advance multimedia-related techniques in a variety of problems including video analysis, image classification, image/object/music retrieval and user relevance feedback.

*Content understanding and structuring* The first paper “*Play Segmentation for the Play-Break based Sports Video Using a Local Adaptive Model*” by Jinguik Jeong presents a framework for providing access to semantic segments in sports videos that could be subsequently used in a variety of advanced content-based applications. A generic framework is presented that can be adapted to specific sports, and its performance in the case of detecting play segments in baseball is illustrated on a test corpus of over 75 h.

*Semantic classification of visual content* The second paper “*Semantic Image Classification using Statistical Local Spatial Relations Model*” by Dongfeng Han, Wenhui Li and Zongcheng Li introduces a novel learning model employing spatial and statistical information for semantic image classification. Inspired by probabilistic Latent Semantic Analysis for text mining, this approach introduces a new factor modeling spatial relation information between image regions.

*Leveraging the power of ontologies in multimedia search* The third paper “*Image Retrieval ++—Web Image Retrieval with An Enhanced Multi-Modality Ontology*” by Huan Wang, Liang-Tien Chia and Song Liu investigates the design and use of an ontology for multimodal similarity matching and ranking for multimedia retrieval on the Web. A new ranking mechanism is proposed to measure conceptual similarity in the ontology, integrating both textual and image features.

*Bridging the gap in 3D object retrieval* The fourth paper “*Semantic Force Relevance Feedback, Content-Free 3D Object Retrieval and Annotation Propagation: Bridging the Gap and Beyond*” by Efstathios Onasoglou and Petros Daras introduces user feedback into the content-based retrieval problem and extends existing approaches to relevance feedback

to include both low-level and semantic features. They demonstrate how their approach can be used to improve retrieval performance for databases of 3D objects.

*Using semantics in music retrieval* The fifth paper “*Relevance Feedback for Category Search in Music Retrieval Based on Semantic Concept Learning*” by Man-Kwan Shan, Meng-Fen Chiang and Fang-Fei Kuo also considers relevance feedback, but targets audio rather than visual data, and specifically semantic retrieval of music segments. The approach uses both local and global features to identify significant audio segments, that are in turn mapped to semantic concepts.

*Considering the user and what she really wants* The sixth paper “*Media objects for user-centered similarity matching*” by Jean Martinet, Shin'ichi Satoh, Yves Chiaramella and Philippe Mulhem presents a user-centric definition of importance for use in similarity matching of media objects. Various criteria and hypotheses for quantifying importance are presented, leading to weighting scheme that is validated in the case of image-based retrieval.

In the special issue of the Journal of Web Semantics, the reader may find three contributions. In this case, they rather target the management of multimedia based on semantic metadata for varying purposes such as generation and delivery on the sides of producers, mediators and consumers of multimedia content.

*From semantic annotation to multimedia generation* The first paper “*Automatic generation of matter-of-opinion video documentaries*” by Stefano Bocconi, Frank Nack and Lynda Hardman describes a model for personalizing video documentaries. Taking into account user choices and rhetorical presentation patterns, the model, implemented in the system Vox Populi, generates documentaries in a bottom-up fashion from given data exploiting semantic annotations.

*Semantics for producers and mediators of multimedia news content* The second paper “*The Semantic Web as a Newspaper Media Convergence Facilitator*” by Roberto Garcia, Ferran Perdrix, Rosa Gil and Marta Oliva considers the changing processes of news production and delivery. It integrates metadata of multimedia news content (press, radio and television) and provides an environment with advanced browsing and searching functionalities for this content together with the audio transcript and background knowledge.

*A semantic system for your kid—and not only for her* The third paper sides with the consumer—rather than with the producers and mediators of multimedia knowledge. The system paper “*ZemPod: A Semantic Approach to Podcasting*” by Oscar Celma and Yves Raimond describes a running system that combines audio analysis with semantic capabilities of subscribing to audio resources for easier access to such multimedia data.

In conclusion, we hope that this selection of papers goes some way towards illuminating some recent exciting results in the field of semantic multimedia. From these papers it is clear that there is significant effort being invested by both the Semantic Web and multimedia analysis research communities in this area. We believe that a key objective of both communities should be to continue and broaden interdisciplinary efforts in this field with a view to extending the significant progress made to date.



**Yannis Avrithis** received the Diploma degree in Electrical and Computer Engineering from the National Technical University of Athens in 1993, the M.Sc. in Communications and Signal Processing (with Distinction) from the Department of Electrical and Electronic Engineering of the Imperial College of Science, Technology and Medicine, University of London, UK in 1994, and the PhD from NTUA on digital image processing in 2001. He is currently a senior researcher at the Image, Video and Multimedia Systems Laboratory (IVML) of NTUA, coordinating R&D activities in Greek and EU projects, and lecturing in NTUA. His research interests include image/video segmentation and interpretation, knowledge-assisted multimedia analysis, annotation, content-based and semantic indexing and retrieval, video summarization and personalization. He has published 1 book, 13 articles in international journals, 15 book chapters and 62 in conferences and workshops. He is an IEEE member, and a member of ACM and EURASIP.



**Noel E. O'Connor** is a Senior Lecturer in the School of Electronic Engineering of Dublin City University and a Principal Investigator in the Science Foundation Ireland funded CLARITY Centre for Science and Engineering Technology. Since 1999 he has published over 130 peer-reviewed publications, filed 5 patents and spun off a campus company. He has acted as PC Chair for 3 international conferences. He has guest edited 5 special issues of different journals. His research interests are in the broad field of multi-modal content analysis for knowledge extraction, including work on audio and visual analysis (including stereo and multi-spectral sources), content-based information retrieval, user interface design and hardware for media processing on mobile devices. He is a member of the IEEE, Engineers Ireland and the Institution of Engineering and Technology.



**Steffen Staab** is professor for databases and information systems at the University of Koblenz-Landau, leading the research group on Information Systems and Semantic Web (ISWeb). His interests lie in researching core technology for ontologies and semantic web as well as in applied research for exploiting these technologies for knowledge management, multimedia and software technology. He has participated in numerous national, European and intercontinental research projects on these different subjects and his research has led to over 100 refereed contributions in journals and conferences. Dr. Staab held positions as researcher, project leader and lecturer at the University of Freiburg, the University of Stuttgart/Fraunhofer Institute IAO, and the University of Karlsruhe and he is a co-founder of Ontoprise GmbH. For more information see: <http://isweb.uni-koblenz.de/> and <http://www.uni-koblenz.de/~staab/>.



**Dr. Raphaël Troncy** obtained with honors his Master's thesis in Computer Science at the University Joseph Fourier of Grenoble, France, after one year spent in the University of Montreal, Canada. He benefited from a PhD fellowship at the National Audio-Visual Institute (INA) of Paris and received with honors his PhD from INRIA Grenoble in 2004. He was then an awarded ERCIM Post-Doctorate Research Associate in the National Research Council (CNR) in Pisa, Italy in 2005, and in the National Research Institute for Mathematics and Computer Science (CWI) in Amsterdam, the Netherlands in 2006 where he is currently employed. He was co-chair of the W3C Incubator Group on Multimedia Semantics, and is an active participant in the K-Space EU Network of Excellence. His research interest include Semantic Web and Multimedia Technologies, Knowledge Representation, Ontology Modeling and Alignment. Raphaël Troncy is an expert in audio visual metadata and in combining existing metadata standards (such as MPEG-7) with current Semantic Web technologies. He works also closely with the IPTC standardization body and the relationship between the NewsML language family and the Semantic Web technologies.