GUEST EDITORIAL

Semantic multimedia

**Werner Bailer · Bernard Mérialdo ·
Yiannis Kompatsiaris · Tat-Seng Chua**

Growing amounts of multimedia data of various modalities (video, audio, 3D objects, etc.) make management, distribution of and access to multimedia material ever harder, both for lay and professional users. Novel approaches, bridging the large disparity between descriptors computed automatically from multimedia content, and the subjectivity and context in user interpretation and interaction, are required. There is still a need for further exploration on how technologies from different research areas can be used to enhance the value of multimedia content, for example, to support multimedia representation, analysis and annotation. Towards this direction, a number of research approaches are focusing on combining multimedia signal and text processing with semantics and knowledge based methods in order to achieve higher-level understanding and annotation of multimedia content. This research direction, often called *semantic multimedia*, combines techniques such as low-level multimedia feature extraction, common semantic representation schemes for features and concepts, machine learning, reasoning and visualization. These

W. Bailer (✉)
Institute for Information and Communication Technologies,
JOANNEUM RESEARCH, Graz, Austria
e-mail: werner.bailer@joanneum.at

B. Mérialdo
EURECOM, Sophia Antipolis, France
e-mail: Bernard.Merialdo@eurecom.fr

Y. Kompatsiaris
Informatics and Telematics Institute, Centre for Research and Technology Hellas,
Thermi-Thessaloniki, Greece
e-mail: ikom@iti.gr

T.-S. Chua
Department of Computer Science, School of Computing,
National University of Singapore, Singapore, Singapore
e-mail: chuats@comp.nus.edu.sg

techniques are applied to exploit information from all available modalities including 3D content and social media tags.

This special issue includes contributions addressing related theoretical and practical aspects of semantic multimedia. Some of the contributions are extended and updated papers based on materials presented at the 4th International Conference on Semantic and Digital Media Technologies (SAMT 2009).

Not surprisingly, models for semantic annotation of different kinds of multimedia data are a core issue in several contributions. Damiano and Lombardo discuss the *Semantic annotation of narrative media objects*. Based on a narratological and computational background, they introduce a model and a schema for the annotation of the narrative features, as well as a software annotation tool. The paper also discusses the application of low-level signal analysis to narrative media objects and illustrates a few projects elaborated with the proposed tools that provide an empirical validation of the annotation process.

Pammer et al. study the characteristics of the problem of identifying descriptive tags of a picture in their contribution titled *Tag-based Algorithms Can Predict Human Ratings of Which Objects a Picture Shows*. Given the theoretical feasibility of a well-performing tag-based algorithm, which they show via an optimal algorithm, they describe the implementation and evaluation of a WordNet-based algorithm as a proof-of-concept. The paper discusses the difficulty of deciding whether a tag is descriptive and distinguishes between different types of disagreement based on a qualitative analysis.

Being more specific in terms of the media type, two contributions deal with the description of semantics of 3D scenes. In *Extending MPEG-7 For Efficient Annotation of Complex Web 3D Scenes* Malamos et al. present an annotation scheme based on MPEG-7 for describing 3D scenes encoded in X3D. The proposed annotation scheme considers geometrical and appearance characteristics of the 3D content as well as animations and interactivity issues. The new descriptors are proposed as extensions of MPEG-7 Visual and MDS, together with the corresponding schema extensions.

Pina et al. propose in *Semantic Visualization of 3D Urban Environments* a scene graph structure for visualizing geometric data with semantic meaning while the user is navigating inside the 3D city model. BqR-Tree, an improved R-Tree data structure, is defined for this purpose and allows for the inclusion of mobile and semantic elements in a natural way. The usefulness of the 3D scene graph has been tested with low-structured data, including dynamic elements in the data structure.

Finally, two contributions address the problems of using Semantic Web technologies to integrate semantic information about media objects from different sources. In *Linked Data and Multimedia: The State of Affairs*, Schandl et al. survey the application of the concepts of Linked Data in the context of multimedia data. While Linked Data has been widely applied in many areas in recent years, the relation between Linked Data and multimedia has not been well studied. This paper discusses the application of Linked Data in two multimedia-related applications and outlines open research issues.

Sentic Computing, a new paradigm for the affective analysis of natural language text, is used by Cambria et al. in *Sentic Computing for Social Media Marketing* for mining opinions about commercial products from a number of sources on the web.

Different ontologies are used to represent the gathered information in a semantic aware format and make it accessible through a multi-faceted browsing web site.

The works appearing in this issue are among the latest developments in semantic multimedia research. We thank the authors and the reviewers for their extensive efforts in preparing this issue. We are confident that the ideas presented here will help to bring about strong improvements in our capabilities to efficiently manage and use the huge amounts of multimedia information we are faced with.