Multimed Tools Appl (2013) 63:281–286
DOI 10.1007/s11042-012-1083-3

GUEST EDITORIAL

Editorial for automated media analysis and production for novel TV services

Gerald Friedland · Alberto Messina · Robbie De Sutter · Masanori Sano

Published online: 22 May 2012

Welcome to this special issue on Automated Media Analysis and Production for Novel TV Services (AIEMPro).

Modern media is characterised by a significant increase in the amount of digital video content (digital TV channels in particular), and the diversification of broadcast possibilities and storage and consumption devices. This phenomenon has recently given rise to the emergence of many new services and novel TV programmes consumption schemes and usage trends (TV-on-demand, interactive TV, Personal Video Recorders, Catch-Up TV), aimed at making broadcast content available to consumers according to their needs, i.e. without any constraint on location and/or time and with the possibility to cherry pick and navigate according to the viewers desire. These services have been already proven to be very effective in significantly increasing the content audience and open new niches for profitable use of broadcast content.

However, this new operational context naturally brings a side effect: many more digital media items need to be produced by the same or reduced staff, and in a faster and more cost-efficient way.

In addition, the new media distribution and consumption paradigms and the corresponding new production workflows based on digital computer-based tools require an immediate revision of the traditional ways of making business in media industry, by substituting the traditional one-to-many broadcasting model with a novel framework in which social media and interactivity play a major role.

G. Friedland (✉)
International Computer Science Institute, 1947 Center Street, Suite 600, CA 94704 Berkeley, USA
e-mail: fractor@gmail.com

A. Messina
Centre for Research and Technological Innovation, RAI Radiotelevisione Italiana, Corso Giambone 68, 10135 Turin, Italy

R. De Sutter
VRT-medialab, Ghent, Belgium

M. Sano
NHK Science and Technology Research Laboratories, Tokyo, Japan

Automation of processes is foreseen as one of the key for success in this scenario, because of the potential costs reduction introduced by it. As a consequence, the adoption of tools for intelligent analysis and synthesis of multimedia data are seen as substantial enabling factors in making interactive, multi-channel and multi-purpose productions value-returning.

In order to build innovative services, crucial and highly time-consuming pre-processing steps of TV content are required. As an industrial environment, however, the media sector is characterized by challenging requirements to teams who aim at creating suitable information extraction and indexing tools, e.g. high accuracy and real-time constraints during the production of live events. Therefore, the ability to efficiently index multimedia streams during real-time acquisition/generation of multimedia content is a general requirement for any technology in this domain.

The aim and scope of the AIEMPro workshop series, with its four-years history of activity, and specifically with this special issue of Multimedia Tools and Applications, is to answer some of the key questions that arise by the application of these techniques in the professional media sector, including:

- how to repurpose material originally produced for the one-to-many broadcasting paradigm on new interactive distribution channels?
- how to protect intellectual property from potential abuses in a useful timeframe?
- how to efficiently index existing multimedia archives with regards to new production paradigms?
- how to organize and run integrated cross-media production platforms in an efficient way?

Unlike controlled experimental conditions, in real-life scenarios the variability of productions and assortment of content genres are elements to be studied scientifically in order to evaluate theoretical and practical limitations of existing or to be developed techniques. This implies the existence of several connected problems, which are also the target of AIEMPro:

- what are the best practices in multimedia object annotation to maximise the efficiency of search and retrieval systems for very large repositories?
- which new multimedia description models have to be adopted to represent the complexity introduced by interactivity with large communities of users and by cross-media interaction?
- how to deal with description/annotation systems when the observation time of a live repository of multimedia objects becomes large (e.g. decades)?
- how to capture the semantic evolution of categorization concepts over time and over large repositories of objects?
- how to assess systems for future-proof quality?
- how to robustly evaluate the validity of experimental results on real-life scenarios?
- how to develop techniques to automate metadata translation and interoperability?
- which are the best ways to index extremely large amount of data obtained from diverse sources, such as modern high quality production and old digitized b/w films?

Being positioned in this highly challenging context, the primary significance of AIEMPro is fostering exchange of ideas and of practices between leading experts in research and leading actors in the media community, in order to catalyze the migration towards new ways of producing media content, aided by large scale introduction of tools for automated multimedia analysis and understanding.

This special issue reflects these objectives by presenting a set of carefully selected articles. The articles belong to roughly three categories: The first set of papers deals with

media archival, retrieval, and navigation, the second set of articles regards automatic editing and production of content and the third class of articles describes basic enabler technologies like audio segmentation and visual object recognition. We summarise the contributions as follows:

The reading start with a system description by Jan Nandzik et al. who present their CONTENTUS approach towards an automated media processing chain for cultural heritage organisations and content holders. Their proposed workflow allows for unattended processing from media ingest to availability through a search and retrieval interface. The engine combines Semantic Web technologies for representing relations between the media and entities such as persons, locations and organizations with a full-text approach for searching within transcribed information gathered through an elaborate set of multimedia content analysis steps. The special issue continues with the Mieke Haesen et al., who present their interactive video explorer. The video exploration tool provides visual representations of automatically detected concepts to facilitate individual and collaborative video search tasks. The approach is evaluated through a user study with professional video searchers. The special issue continues seamlessly with an article by Ilaria Bartolini et al. who present SHIATSU, a system for video retrieval which is based on the (semi-)automatic hierarchical semantic annotation of videos. Combining visual content analysis and user contributions, videos can be searched by means of attached tags and/or visual features. The final article in the block of retrieval and archival systems is the article by Friedland et al. who describe the Joke-O-Mat system. Originally an ACM Multimedia Grand Challenge winner, the next version of their audio-cue-based navigation system allows to browse a sitcom by scene, punchline, and dialog segment, and to filter these themes by actor and by keyword. The article also presents an evaluation of the overall user satisfaction as determined by a human-subject study.

The second set of articles is lead in by Erik Mannens et al. describing an automatic news recommendation system. They discuss how personalised recommendation and distribution of news events, encoded using an RDF/OWL representation of the NewsML-G2 standard, can be enabled by automatically categorising and enriching news events metadata via indexing in combination with linked open datasets available on the web of data. The resulting recommendations are based on a global, aggregated profile, which also takes into account the (dis)likings of peer friends and are then finally fed to the user via a personalised RSS feed. The second article in this block is by Alberto Messina et al. The contribution presents the Hyper Media News (HMNews) system, an approach that automatically aggregates information streams from both digital television and the Internet. TV newscasts are automatically segmented, annotated, and indexed and then integrated with content available from the Internet, such as blogs, newspapers and press agency publications. Facing the explosive growth of near-duplicate videos in the Internet, Junge Shen et al. then present their video archeology approach: The system determines derived videos according to the occured manipulation operations and then creates a video migration map showing the pair-wise relationships in a set of near-duplicate videos.

The block of basic enabling technologies is started with the article by Mohammad A. Haque and Jong-Myon Kim on audio segmentation, an enabling technology that plays a key-role in many content analysis tasks. The methods presented by the authors allows efficient segmentation even when the signal contains audio effects such as fade-in, fade-out, and cross-fade. The issue continues with the work by Tobias Schwarze et al. on role-based identity recognition for TV. Their paper presents a system that implements an approach to identify actors by using robust face detection first to localise actors. Then, by clustering similar face instances the relative frequency of their appearance within a sequence is determined. In combination with a coarse textual annotation manually created by the broadcast station's archivist the roles and

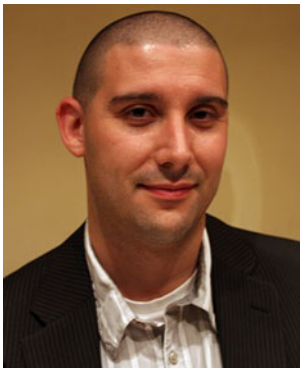
consequently the identities can be assigned and labeled in the video. Stefano Messelodi and Carla Maria Modena, then propose to use scene text recognition and tracking to identify people in videos. Their work, however, is based on sports videos and they identify athletes. The determination of the genre of a broadcast is then discussed by Hazım Ekenel and Tomas Semela. Their automatic video genre classification system utilizes several low level audio-visual features as well as cognitive and structural information, and in case of web videos tag-based features, to classify the types of TV programs and YouTube videos. Last but not least, Varin Chouvatut et al. present an article on 3D face and motion estimation. As 3D television becomes reality, it needs to be embraced more and more by the multimedia content analysis community.

We hope you enjoy reading our special issue as much as we enjoyed composing it. We hope that this sample of contributions from the AIEMPro community will promote the understanding of the field and the technical challenges therein. Please feel free to contribute papers to the ongoing workshop series and also to contact us for any questions and suggestions.

Gerald Friedland, Alberto Messina, Robbie De Sutter, and Masanori Sano
Guest Editors, AIEMPro Special Issue

Gerald Friedland is a senior research scientist at the International Computer Science Institute, a private lab affiliated with the University of California, Berkeley, where he leads multimedia content analysis research, mostly focusing on ("non-speech, non-music") acoustic techniques as an aid for video analysis. He is currently heading a group of multimedia researchers supported by government and industry grants. Gerald has published more than 100 peer-reviewed articles in conferences, journals, and books and is currently authoring a new textbook on multimedia computing together with Dr. Ramesh Jain. Gerald co-founded the IEEE International Conference on Semantic Computing and is a proud founder and program director of the IEEE International Summer School on Semantic Computing at UC Berkeley. He is associate editor for ACM Transactions on Multimedia Computing, Communications, and Applications, is in the organization committee of ACM Multimedia 2011, 2012, and 2014. He is also serves as TPC Co-Chair of IEEE ICME 2012. He is the recipient of several research and industry recognitions, among them the European Academic Software Award and the Multimedia Entrepreneur Award by the German Federal Department of Economics. Most recently, he lead the team that won the ACM Multimedia Grand Challenge in 2009. Gerald received his doctorate (summa cum laude) and master's degree in computer science from Freie Universität Berlin, Germany, in 2002 and 2006, respectively.

Alberto Messina is from RAI Centre for Research and Technological Innovation, where he leads the research area concerning automated analysis and management of multimedia information, a field in which he counts more than 70 publications. He is involved in several research projects in the field of digital archiving, automated documentation, and automated production. Former chairman of the EBU Expert Community on Metadata, he now chairs the EBU Strategic Programme on Media Information Management (SP/MIM). As part of his professional activity he has been working in several EC projects and he is now involved in FP7 TOSCA-MP, VISION Cloud and PrestoPRIME. He regularly serves in the Programme Committee of scientific conferences and workshops, and he's General Co-Chair of the International Workshop on Automated Media Analysis and Production for Novel TV Services (AIEMPro). He participates in International Standardisation bodies, specifically MPEG, where most notably recently contributed to amend MPEG-7 Part 9 (AVDP - AudioVisual Description Profile). He has been recently appointed Contract Professor at Politecnico di Torino for a course entitled Multimedia Archival Techniques.



Robbie De Sutter received the M.S. degree in computer science in 1999 and his Ph.D. degree in 2006 from Ghent University, Belgium. After a brief period as full time professor at the University College of West Flanders, Belgium, he joined VRT-medialab as researcher. He was in charge of the information management research group, which initially focused on the creation, adaptation, integration, and exploitation of audiovisual descriptive metadata and also investigated automatic content analysis tools and improved indexing and search algorithms for audiovisual materials. He was vice-chair of the EBU Expert Community on Metadata and was involved in the ECM/SCAIE workgroup which studies and evaluates automatic information extraction analysis tools for audiovisual production and archived content. In 2011, his group started to investigate how to realize a CRM methodology in the audiovisual media (TV and radio) by taking advantage of the proliferation of mobile internet connected devices like tablets and smartphones in the home. These so-called "second screen devices" gives broadcasters, and their advertisers, the possibility to (re)establish a direct

relation with their customers. Beginning 2012, VRT-medialab evolved into IBBT' Media Innovation Centre where Robbie De Sutter is project leader of innovative projects beneficiary for the entire media Flemish sector.

Masanori Sano he received B.E, M.E. in electronic engineering from Waseda Univeristy, and Ph.D degrees in informatics from The Graduate University for Advanced Studies. He joined NHK in 1994 and worked at the Sendai Broadcasting Station. Since 1997 he has worked for NHK Science and Technology Research Laboratories. His research interests include information extraction from media, information integration, and metadata production system. He has involved with standardization, MPEG, EBU, ABU and ARIB. He is a member of IEEE, ACM, IEICE and ITE.