Multimed Tools Appl (2011) 54:1–5
DOI 10.1007/s11042-010-0543-x

GUEST EDITORIAL

Management and interaction with multimodal information content

Richard Chbeir · Karin Coninx · Fernando Ferri · Patrizia Grifoni

Published online: 8 June 2010

1 Preface

Communication is the core for cultural, scientific, economic and technological evolution of the contemporary society, its progress and innovation. Nowadays, more and more Information and Communication Technologies are emphasizing the central role of all issues connected with communication. The numerous and different ways people are using to communicate with each other, the multiplicity of features to produce and exchange information, and the relevance and pervasiveness of multimodal, mobile devices and phenomena such as Internet, are all elements that enhance the complexity of communication processes and their management.

Devices supporting multimodal interaction become more and more widespread. When applied in an appropriate way, multimodal interaction provides users with a flexible, natural and robust interaction approach, allowing them to communicate in a synergistic manner using their five senses along with several communication channels. This also permits the organization to store, index, retrieve and more generally to manage a wide amount of multimodal data and information, thus, enabling people to use a multimodal dialog approach in order to access information and/or services.

In this context, we have decided to contribute to the scientific debate concerning these emerging questions defining this special issue. After a very selective peer review, nine papers were selected for publication. They deal with theories and techniques concerning multimodal information retrieval, indexing, query processing and extracting features from multimodal data, multimodal interaction issues, and applications.

This special issue is organized as follows.

R. Chbeir (✉)
Department of Computer Science, LE2I-CNRS, Bourgogne University, Dijon 21000, France
e-mail: Richard.Chbeir@u-bourgogne.fr

K. Coninx
Expertise Centre for Digital Media (EDM), Universiteit Hasselt, Campus Diepenbeek, Wetenschapspark 2, BE-3590 Diepenbeek, Belgium

F. Ferri · P. Grifoni
Institute of Research on Population and Social Policies, National Research Council, Rome, Italy

In the first paper titled “Scene Extraction System for Video Clips using Attached Comment Interval and Pointing Region”, S. Wakamiya et al. address the issue of information retrieval and querying a large amount of multimodal information and data. They describe a method enabling users of video sharing websites to easily retrieve video scenes according to their interests. Video analysis techniques, such as image processing, and speech recognition are useful for recognizing objects in a video clip.

However, these types of analysis are expensive. The proposed method enables users to view scenes and their attached annotations, both considering text information and non-text information, according to their specific interests.

In the second paper, titled “Bayesian Belief Network Based Broadcast Sports Video Indexing”, Maheshkumar H. Kolekar presents a method for automatic indexing of excitement clips of sports video sequences based on a probabilistic Bayesian belief network (BBN). The paper gives a general method to the automatic tagging of large scale multimedia content with rich semantics enabling browsing, searching and manipulating video documents. The proposed method has been validated in the sports domain by demonstrating successful indexing of soccer and cricket video excitement clips.

The third paper is called “MQSS: Multimodal Query Suggestion and Searching for Video Search” authored by L. Li and J. Li. Here, the authors aim at improving access for video information by describing a multimodal query suggestion method for video search. Suggested keywords and representative image examples are presented in an easy-to-use dropdown list in order to support users in specifying their query precisely and effortlessly. The effectiveness of the provided approach has been evaluated, and 96% of users estimated that the query suggestions of MQSS are useful, while 74% and 66% users considered the query suggestions of Google Video Search and Yahoo! Video Search useful, respectively.

Z. Wu et al. address in “Multimedia Selection Operation Placement” the complexity of queries involving multimedia operations. They describe their theory and algorithm for optimizing queries using expensive multimedia operations. Also, they try to establish the optimal placement of each multimedia operation in a query plan by considering selectivity and unit execution cost of each operation. The defined algorithm has a time complexity that is polynomial in the number of multimedia operations in the query plan.

Authored by P. Luigi Scala et al., the fifth paper is titled “TMS for Multimodal Information Processing”. It elaborates on a component-based system for modeling and automatizing the management of complex information-oriented working processes. In particular, working processes are considered that are able to search, acquire, describe and assemble computational agents.

In the sixth paper, J. Renny Octavia et al. provide “Adaptation in Virtual Environments Conceptual Framework and User Models” in which they propose a conceptual framework for an adaptive personalized interaction in virtual environments. Switching between interaction techniques in virtual environments according to the constructed user model has been implemented using the adaptation framework. An evaluation has demonstrated that users positively respond to the use of an adaptable system as their frustration usually decreases.

A. Cerekovic et al., in the seventh paper called “Multimodal Behavior Realization for Embodied Conversational Agents”, assume that decreasing frustration and increasing naturalness are not sufficient for an adequate interaction. In fact, interaction has to help overcoming the digital divide. In this scenario naturalness of the interaction processes is a particularly relevant factor. Thus, they describe how to create believable and expressive virtual characters in order to enhance the communication abilities of machines. They use Embodied Conversational Agents’ technologies to bring human-like abilities into machines

for verbal and nonverbal channels, discussing issues of multimodal behavior focusing on motion control.

In "On creating multimodal virtual humans—real time speech driven facial gesturing", G. Zoric et al. describe a multimodal interface, based on virtual humans, which uses speech as input and speech with facial gestures as output. They provide an original method for automatic audio to visual mapping to produce a wide set of facial gestures. The mapping only occurs based on the speech signal in the real time, using a hybrid data-driven and rule-based approach.

In the last paper concluding this special issue, "Improving Multimodal Web Accessibility for Deaf People: Sign Language Interpreter Module", M. Debevc et al. focus on the problem of web sites' Accessibility. The authors address the necessity of reducing the digital divide in the particular case of deaf and hard-of-hearing people. The paper underlines the importance of using sign language instead of information presented in an appropriate written form. The SLI software Module has been presented; on a contextual and technical basis the deaf and hard-of-hearing people can adjust the web access to their needs.

We hope this special issue motivates researchers to take the next step beyond building models to implement, evaluate, compare, and extend their proposed approaches. Making this special issue to become a reality has been a considerable effort for many persons. We would first like to gratefully acknowledge and sincerely thank all the reviewers for their timely and insightful valuable comments and evaluations of the manuscripts that greatly improved the quality of the final versions. Of course, thanks are due to the authors, who provided excellent articles and timely extended revisions. Finally, we are grateful to the editors of MTAP for their trust in us.

Dr. Richard Chbeir received his PhD in Computer Science from the University of INSA-FRANCE in 2001. The author became a member of IEEE since 1999. He is currently an Associate Professor in the Computer Science Department of the Bourgogne University, Dijon-France. His research interests are in the areas of distributed multimedia database management, XML similarity and rewriting, spatio-temporal applications, indexing methods, multimedia access control models, security and watermarking. Dr. CHBEIR has published (more than 40 peer-reviewed publications) in international journals and books (IEEE Transactions on SMC, Information Systems, Journal on Data Semantics, Journal of Systems Architecture, etc.), conferences (ER, WISE, SOFSEM, EDBT, ACM SAC, Visual, IEEE CIT, FLAIRS, PDCS, etc.), and has served on the program committees of several international conferences (ICDIM, IEEE SITIS, ACM SAC, IEEE ISSPIT, EuroPar,

SBBD, etc.). He has been organizing many international conferences and workshops (ICDIM, CSTST, SITIS, etc.). He is currently the Chair of the French Chapter ACM SIGAPP and the vice-chair of ACM SIGAPP.

Prof. Dr. Karin Coninx is full professor at Hasselt University and vice-dean of the faculty of sciences. As one of the initiators of the master on Human–Computer Interaction (HCI), she teaches several courses on computer science and specific HCI subjects.

Karin Coninx obtained a PhD in sciences, computer sciences after a study of HCI in immersive virtual environments. Her research interests include Context-Aware User Interfaces, Model-Based and User-Centered Software Engineering, Mobile Guides, Multimodal interaction and flexible development of Virtual Environments, Ubiquitous Computing, and (Collaborative) Interactive Workspaces.

As the HCI research group leader in the Expertise Centre for Digital Media (EDM) at Hasselt University, Karin Coninx advises about 25 researchers, among which about 10 PhD students. She was main applicant and otherwise involved in many research projects in the HCI domain, investigating the aforementioned topics. Karin Coninx has co-authored more than 200 international publications (among which about 70 journal articles and chapters in books), and co-organized several international workshops. An overview of publications can be found at http://www.edm.uhasselt.be/people/show/karin.coninx.

Fernando Ferri received the M.S. degree in electronics engineering in 1990 and the Ph.D. degree in medical informatics in 1993, both from the University of Rome “La Sapienza”. He is a senior researcher at the National Research Council of Italy since 2001. From 1996 to 2001 he was a researcher at the National Research Council of Italy. From 1993 to 2000 he was a professor of “Sistemi di Elaborazione” at the University of Macerata. He is the author of more than 140 papers in international journals, books and conferences. He was responsible for several projects funded by Italian Ministry of University and Research and European Commission. He co-organized several international workshops, conferences and he

participated at Program Committees for international Conferences and Journals. His main methodological areas of interest are: human–computer interaction, visual languages, visual interfaces, sketch-based interfaces, and multimodal interfaces and languages, data and knowledge bases, geographic information systems, Web technologies, social networks and risk governance.

Patrizia Grifoni received the M.S. degree in electronics engineering in 1990. From 1994 to 2000 she was a professor of Elaborazione digitale delle immagini at the University of Macerata. She is a researcher at the National Research Council of Italy since 1993. She is the author of more than 100 papers in international journals, books and conferences. She co-organized several international workshops, conferences and she participated at Program Committees for international Conferences and Journals. She was responsible for several projects funded by Italian and International Institutions. Her scientific interests have evolved from query languages for statistical and geographic databases to the focal topics related to human–computer interaction, multimodal interaction and languages, visual languages, visual interfaces, sketch-based interfaces, Web technologies, social networks and risk governance.