

## Special issue on multimodal interfaces in cognitive infocommunication systems

Peter Baranyi · Adam Csapo

Published online: 18 May 2014  
© OpenInterface Association 2014

Humans and the infocommunications network surrounding them are merging together at various levels, ranging from the level of interactions with personal devices to the highest level of sensing collective behaviors such as mass movements, mass habits etc. Consequently, humans and infocommunications will soon coexist as an entangled web, resulting in an augmentation and merging of both natural and artificial cognitive capabilities. This process of merging is occurring today, and is expected to gain further impact in the near future.

Cognitive infocommunications (CogInfoCom) is an interdisciplinary field that aims to reflect on this process of merging by investigating links between the research areas of infocommunications and the cognitive sciences, as well as the various engineering applications that have emerged as the synergic combination of these sciences.

Given the interdisciplinary background of CogInfoCom, many aspects of the synergies behind it can be considered. The papers in this special issue all focus on some aspect of multimodal interfaces in the context of infocommunication systems that enable enhanced cognitive capabilities through the merging of the natural and artificial. Papers range from the more theoretical to those which introduce practical applications.

Several papers address multi-modal aspects of human speech. The paper by Á. Abuczki focuses on a multi-modal corpus based approach to modeling cognitive, information and interactional states behind speech interactions, based on a set of discourse markers and hand-based gestures. The papers

by C. Navarretta and I. Szekrényes *et al.* also describe theoretically important benefits of multi-modal speech corpora. The former focuses on the possibility to enhance the prediction of the function of various facial expressions based on emotional labels as well as head movements and turn taking, while the latter introduces the HuComTech corpus and its use in modeling and retrieving multi-dimensional (interpreted) aspects of prosodic structures. Finally, the paper by D.S. Cheng *et al.* presents results in using features based on prosody, voice quality and gestural behaviors in predicting the user ratings of on-line lectures.

Two of the papers in the special issue, authored by J.R. Blum *et al.* and J. Balata *et al.* highlight the strong value in enhancing the capabilities of artificial systems based on the cognitive capabilities of a potentially large, distributed group of human collaborators. The former paper introduces a real-time emergency response system that involves a layered architecture of disaster recovery management and volunteers, and a platform for targeted video stream sharing. The paper by J. Balata *et al.* discusses the value of supporting the navigation of the visually impaired through instructions given by the visually impaired, and provides an analysis of the requirements for navigation assistance application that is to be developed based on this idea.

A further two papers in the special issue focus on novel methodologies for the design of 3D augmented/virtual collaboration. The paper by T. Duval *et al.* introduces several kinds of approaches that support collaborative interaction based on two key directions: by adding elements of the user's environment to the virtual environment, and by using visual feedback elements to enhance the user's awareness of specific processes of interaction. The paper by C. Pontonnier *et al.* addresses multi-modal aspects of digital mockup design which will enable costly assembly scenarios to be reproduced and tested/evaluated in virtual environments.

---

P. Baranyi · A. Csapo (✉)  
MTA SZTAKI/BME, Budapest, Hungary  
e-mail: csapo.adam@tmit.bme.hu

P. Baranyi  
e-mail: baranyi.peter@sztaki.mta.hu

Somewhat related to these augmented aspects of interaction, the paper by *Á. Török et al.* investigates the effects of the intensities and co-located/contradicting nature of audio-visual stimuli in terms of multi-sensory integration and response time. The results of the paper can benefit the design of e.g. intelligent warning systems in a wide range of application environments.

Finally, the paper by *Macik et al.* describes an approach to context sensitive user interface generation based on a context model that addresses a variety of parameters, including—among others—the user’s cognitive capabilities relative to both the device and the application domain, the device para-

meters, and the complexity of the information that is to be displayed. The paper uses code inspection and aspect-oriented programming techniques, and includes two use-case examples that illustrate the approach.

In summary, the papers of the special issue cover a wide range of topics that may contribute to a deeper appreciation of the many ways in which the modern development of ICT facilitates the merging of various levels of human and artificial cognitive capabilities.

January 2014