

Preface

Human communication is naturally multimodal, involving the interaction of modalities such as speech, facial expressions, hand gestures and body posture. In order to have a better understanding of human-human communication and to improve human-computer interaction it is essential to identify, describe, formalise and model the interaction of the different modalities in human communication. The past two decades have witnessed numerous initiatives and research efforts to improve the state of the art, including collection and annotation of multimodal corpora, automatic recognition of the different modalities and modelling and generation of multimodal data. However, there are still many questions and problems concerning the annotation of multimodal data and the technology for capturing such data, not to mention the interpretation and reproduction of complex, natural multimodal behaviour.

The main aim of the workshop has been to provide a multidisciplinary forum to present results and discuss issues that concern research on human multimodal communication, its modelling and representation for computational systems. The workshop call mentioned a large range of relevant topics: cognitive aspects of multimodal communication; formal frameworks and descriptions of multimodal communication; representational issues, e.g. definition of annotation units, granularity of descriptions, spatio-temporal models of non-verbal modalities, definition of default values, representation of multimodal meaning and inclusion of world context; interaction of the different modalities; multimodality in intercultural communication; definition of communicative functions in multimodal communication; methodologies and tools to annotate, process and produce multimodal communication; multimodal signal processing and its integration with manual annotation.

Each submitted paper was blindly reviewed by three reviewers. Seven papers were submitted, one paper was withdrawn and five papers were accepted to be presented at the workshop. The accepted papers cover several of the aspects of multimodal communication listed in the workshop call, including annotation, representation, analysis and processing issues.

The invited speaker Nick Campbell in his “Technology for Processing Non-verbal Information in Speech” discusses the importance for researchers of being able to automatically collect and process non-verbal information in speech. This is a prerequisite for using and integrating this information in more and more advanced commercial applications such as machine interpretation, games, robotics and customer services.

In the paper “Gestures that precede and accompany speech – An analysis of their functions and use in the design of virtual agents in different activities”, Jens Allwood and Elisabeth Ahlsén discuss the behavioural and functional features of gestures produced before or simultaneously with speech and relate them to two activity types for Embodied Communicative Agents: front-end to database and educational training.

Christopher Habel and Cengiz Acarturk propose, in their paper “Eye-tracking evidence for multimodal language-graphics comprehension: The role of integrated conceptual representations”, a modular architecture for the comprehension of textual and graphical input in which the conceptual representations coming from the two modalities’ contribution are integrated. Experiments comparing eye-tracking movements caused by graphical input on the one hand and graphs and texts on the other are then described and discussed with respect to the proposed architecture.

Kristiina Jokinen and Minna Vanhasalo in “Stand-up Gestures – Annotation for Communication Management” discuss the form and functions in communication of so called stand-up gestures. The functions of these gestures comprise the regulation and coordination of communication, thus they are quite important in communication management. The authors also propose a way to annotate stand-up gestures via an extension of the MUMIN annotation scheme to include a meta-discursive context level.

“Coarticulation in sign and speech” by Stina Ojala, Tapio Salakoski and Olli Aaltonen presents a

study of coarticulation in Finnish sign language and speech with the main aim of discovering similarities between coarticulation in signing movements and speech. The study's results indicate that the alternations of deceleration and acceleration in signing movements can be compared to deceleration and acceleration patterns which occur at different levels in speech.

The paper "Integration and representation issues in the annotation of multimodal data" by Patrizia Paggio and Costanza Navarretta deals with issues related to the representation of gestures and speech in a multimodal sign in terms of feature structures in a unification-based grammar. The authors also discuss some of the complexities related to the interpretation of the multimodal sign, such as the interaction of gestures and speech at different conceptual levels and their multi-functionality.

The Organizing Committee

Costanza Navarretta, University of Copenhagen

Patrizia Paggio, University of Copenhagen

Jens Allwood, University of Gothenburg

Elisabeth Alsén, University of Gothenburg

Yasuhiro Katagiri, Future University, Hakodate

Acknowledgements

We want to thank the program committee for reviewing the workshop papers:

Nick Campbell, ATR, Osaka

Loredana Cerrato, Acapela Group Sweden

Dirk Heylen, University of Twente

Kristiina Jokinen, University of Helsinki and University of Tartu

Michael Kipp, DFKI Germany

Brian MacWhinney, Carnegie Mellon University

Jean-Claude Martin, CNRS-LIMSI France

Catherine Pelachaud, University of Paris 8

Isabella Poggi, Roma Tre University

Andrei Popescu-Belis, Idiap Research Institute

Rainer Stiefelhagen, Karlsruhe University

Johannes Wagner, University of Southern Denmark

Massimo Zancanaro, Bruno Kessler Foundation, Trento