Modeling Cognitive Processes from Multimodal Signals

Felix Putze
University of Bremen
Bremen, Germany
felix.putze@uni-bremen.de

Jutta Hild
Fraunhofer IOSB
Karlsruhe, Germany
jutta.hild@iosb.fraunhofer.de

Akane Sano
Rice University
Houston, TX, USA
akane.sano@rice.edu

Enkelejda Kasneci
University of Tübingen
Tübingen, Germany
enkelejda.kasneci@
uni-tuebingen.de

Erin Solovey
Worcester Polytechnic Institute
Worcester, MA, USA
esolovey@wpi.edu

Tanja Schultz
University of Bremen
Bremen, Germany
tanja.schultz@uni-bremen.de

WORKSHOP SCOPE

Multimodal signals allow us to gain insights into internal cognitive processes of a person, for example: speech and gesture analysis yields cues about hesitations, knowledgeability, or alertness, eye tracking yields information about a person's focus of attention, task, or cognitive state, EEG yields information about a person's cognitive load or information appraisal. Capturing cognitive processes is an important research tool to understand human behavior as well as a crucial part of a user model to an adaptive interactive system such as a robot or a tutoring system. As cognitive processes are often multifaceted, a comprehensive model requires the combination of multiple complementary signals. In this workshop at the ACM International Conference on Multimodal Interfaces (ICMI) conference in Boulder, Colorado, USA, we discussed the state-of-the-art in monitoring and modeling cognitive processes from multi-modal signals.

OBSERVATIONS

The presented research covers a large range of different topics centered on the workshop theme. The cognitive states and processes, which are modeled in the submissions, range from engagement, over episodic memory to problem solving. The two most prominent aspects of cognition at the workshop are attention and cognitive load. Both topics are tackled in multiple contributions from different perspectives. This clustering is logical as both concepts are central to understanding human behavior and have an immediate impact on human-computer interaction.

The models employed in the various submissions use many different modalities in their models: For example, eye tracking (from standard video or dedicated eye tracker devices) is one of the central modalities to capture visual attention. Another modality that is used as a correlate of cognitive processes is brain activity (e.g. captured through EEG or fNIRS) as it carries relevant information even if the outward appearance gives no indication of internal processes. Of course, features from audio and video data also play an important role in the analysis of cognitive states, especially when studying interactions. In the spirit of a conference on multimodality, many contributions investigate the combination of different modalities to create models that are more robust or a broader coverage of concepts. In the submissions to the workshop, we see that many researchers now routinely select from a "toolbox" of available modalities to find the most plausible model for their cognitive processes of interest.

The domains researchers investigate cover many different areas and range from controlled basic research to more applied research in the wild. Some topics that are touched are game playing, life logging, or medical professionals. Some applications are set in a human-computer (or human-robot) interaction scenario while others investigate the interaction of multiple humans with each other. This variety shows that modeling of cognitive processes is relevant to a wide range of scenarios and not only limited to research in cognitive science but makes an important impact on the development of intelligent and adaptive technology.

Finally, the discussed research does not only deal with the detection and modeling of cognitive states but also with strategies for using the resulting models to provide improved human-computer interfaces, for example through automatic adaptation of system behavior.

CONCLUSIONS

The quality and range of the submissions to the workshop shows that research on modeling cognitive processes has progressed in a fast and impactful way. Next steps in this field should look for a convergence of terms, features, and benchmark datasets for multimodal modeling of cognitive processes as the area of affective computing has done for the modeling of emotions.

ICMI '18, October 16–20, 2018, Boulder, CO, USA

ACM ISBN 978-1-4503-5692-3/18/10.
https://doi.org/10.1145/3242969.3265861