APRnews

Newsletter of the KIT Focus Anthropomatics and Robotics

Issue 1|2012

EDITORIAL

The capability of interaction, cooperation, coexistence and communication between humans and machines is of key importance for human-machine symbioses. How can humans and machines interact? How should the communication between these actors be designed? The answers to these questions must rely on a thorough understanding of the human communication processes to allow the development of intuitive multimodal interfaces as natural channels for human-machine communication.

Multimodal interaction and communication is one of the central scientific research pillars of the KIT Focus Anthropomatics and Robotics. Ultimate goal is the design and implementation of multilingual interactive and multimodal systems able to perceive humans, their activities, their intention, and their environments, by processing signals from multiple sensors to understand situations and react accordingly. Such a natural and intuitive communication will facilitate shaping the internal representation of the world for the robot and speeding up learning capabilities through human teachers.

This first issue of the APRnews provides a glimpse into some latest findings at KIT concerning human-machine multimodal interaction and communication. The Newsletter concludes by presenting interACT, an excellence joint center for international student exchange and cooperation.

Prof. Rüdiger Dillmann
Spokesperson KIT-Focus
Anthropomatics and Robotics

NEWS

University without Language Barriers
On June 11th, the Interactive System Labs presented the world's first automatic simultaneous translation service at a university. The lecture translator represents a cost effective way to assist foreign students, who can follow the lecture in their own language on their PC or mobile phone.

Opening of the FZI House of Living labs
On April 27th, the Forschungszentrum Informatik (FZI) celebrated the opening of the House of Living Labs (HoLL). The HoLL offers researchers and partners from academia and society an excellent environment for multidisciplinary research.

NAIST new interACT partner
The Nara Institute of Science and Technology, NAIST, joined the International Center for Advanced Communication Technologies, interACT, on April 9th.

CeBIT 2012
From March 6th to 10th, the KIT-Focus Anthropomatics and Robotics attracted many visitors to the KIT booth at CeBIT 2012, in Hannover. The humanoid robot ARMAR III demonstrated its capabilities in a kitchen environment. The FZI House of Living Labs presented the interactive HoLLiE service robot and solutions for intelligent energy management.

New Chair "Humanoid Robotic Systems"
In March, Dr. Tamim Asfour has been appointed to a new Professorship in "Humanoid Robotic Systems". The newly established Chair will strengthen KIT expertise in the anthropomatic research area.

First Karlsruhe Forum on Anthropomatics and Robotics
The First Karlsruhe Forum on Anthropomatics and Robotics was held on February 2nd, 2012, and brought together a multidisciplinary audience of academics and industry representatives to exchange achievements, solutions, visions, and foster technology transfer in the area of anthropomatics and robotics.

Prof. Rüdiger Dillmann elected as IEEE Fellow
In January, Professor Rüdiger Dillmann of the Department of Computer Science was honored with the prestigious title of IEEE Fellow.

KIT-Focus Anthropomatics and Robotics meets the citizens
The Focus Anthropomatics and Robotics presented itself on January 17th, in the Karlsruhe City Hall. The event ("KIT im Rathaus") attracted a large audience of Karlsruhe citizens that could gain insight into novel research activities, sharing the common goal of improving the quality of human life.

DFG grants new SFB / Transregio 125
The German Research Council (Deutsche Forschungsgemeinschaft, DFG) granted the new Transregional Collaborative Research Centre (SFB/TRR 125) "Cognition-Guided Surgery". The project starts in July 2012 and will last four years in the first phase. Partners in the project are the Heidelberg University Hospital, the Karlsruhe Institute of Technology and the German Cancer Research Center Heidelberg. The aim is to create a technical, cognitive system to support the surgeon.

Silent Speech Interfaces

The invention of mobile phones has changed our life by giving us the freedom to communicate with each other at any time from and to everywhere. As users we highly appreciate this convenience and sometimes wonder how we managed to get things done before. However, as bystanders we are often stressed out by the noise pollution resulting from phone rings and loud conversations carried out by our human fellows. While the mobile phone manufacturers were quick to invent the silent phone mode using vibration sensors, it is still an open question how to mute a speaker without limiting his or her communication needs! Text messaging is often used as a silent alternative, however text-based information exchange cannot compete with spoken communication in terms of efficiency, acquaintance, naturalness and richness, to name only a few aspects.

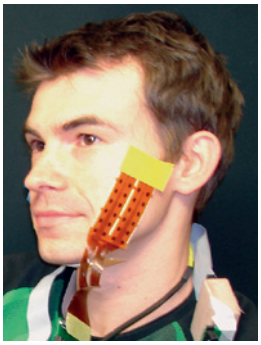
Since several years the Cognitive Systems Lab (CSL) team, part of the Institute for Anthropomatics and the KIT-focus on "Anthropomatics and Robotics", is developing "Silent Speech Interfaces", i.e. communication methods and devices which allow their users to communicate naturally via silent speech, meaning without making any sound. The technology is based on surface Electromyography (EMG) which captures electrical potentials that arise from muscle activity. Since speech is produced by the contraction of muscles which move our articulatory apparatus, the signals corresponding to electrical potentials generated by the articulatory muscles encode speech. The algorithmic processing and interpretation of these signals allows the reconstruction of what had been spoken silently. Since Electromyography captures the muscle activity rather than sound, speech can be recognized even when the speaker only mouths the words, i.e. utters silently without any vocal effort.

Electrodes capture the electromyographic signal at the surface of the facial skin

Photo: Deutsche Messe Hannover

Electrodes of the size of a one-cent coin capture the electromyographic signal at the surface of the facial skin. These signals are passed on to a mobile amplifier, from there are transmitted to a laptop and then preprocessed to filter out biophysiological, technical, and environmental artifacts, before relevant features are extracted. Based on these features the signals are fully automatically recognized and interpreted using algorithms based on pattern recognition and machine learning. At CSL, we are currently pursuing two different ways of signal interpretation. In the first case, dedicated speech recognition software is used to transform the features into a textual representation of what has been spoken. The text is then transferred to a text-to-speech generator to synthesize audible speech for the listener. The text may also be displayed on the screen, send as text message or used to control machines. In the second case, the recognition process is bypassed by directly mapping the myoelectric signals to a spectral representation of speech using voice transformation techniques. Subsequently, the spectral features are vocoded to transform them into audible speech.

There are more benefits to Silent Speech Interfaces than muting a speaker. Since communication can be carried out silently, bystanders will no longer be able to eavesdrop on private or confidential phone conversations. Consequently, the input of pins and passwords via speech comes within reach. Furthermore, Silent Speech Interfaces may give people a voice who have lost it due to illness or accidents, and may enhance voices of speakers who are too weak to speak aloud. Last but not least, EMG-based processing could be a solution to robust speech recognition in adverse environmental noise conditions. Since the speech signal is captured before it gets airborne, it is not prone to any surrounding noise.



Electrode arrays are envisioned to be integrated into digital devices and mobile phones

At CSL, the next stage of system development is currently under way. The single electrodes will be replaced by electrode arrays which are simpler and faster to handle, are able to compensate for electrode positioning shifts, and allow for advanced signal processing. These sensor arrays are envisioned to be integrated into digital devices and mobile phones in the near future. Technological advances may bring out new biosensors which – one day – might get implanted or simply injected under the skin. But this will be decided by future generations of mobile phone users.

- Bruce Denby, Tanja Schultz, Kiyoshi Honda, Thomas Hueber, Jim Gilbert, Jon Brumberg, Silent Speech Interfaces, Speech Communication Journal, Volume 52, Issue 4, April 2010.
- Michael Wand, Tanja Schultz, Speaker-Adaptive Speech Recognition Based on Surface Electromyography, Selected Papers from the International Joint Conference Biomedical Engineering Systems and Technologies, Communications in Computer and Information Science, Springer, Vol. 52, 2010.
- Matthias Janke, Michael Wand, Keigo Nakamura, Tanja Schultz, Further Investigations on EMG-to-Speech Conversion, IEEE International Conference on Acoustics, Speech, and Signal Processing, Kyoto, Japan, 2012.

Tanja Schultz

Further information:

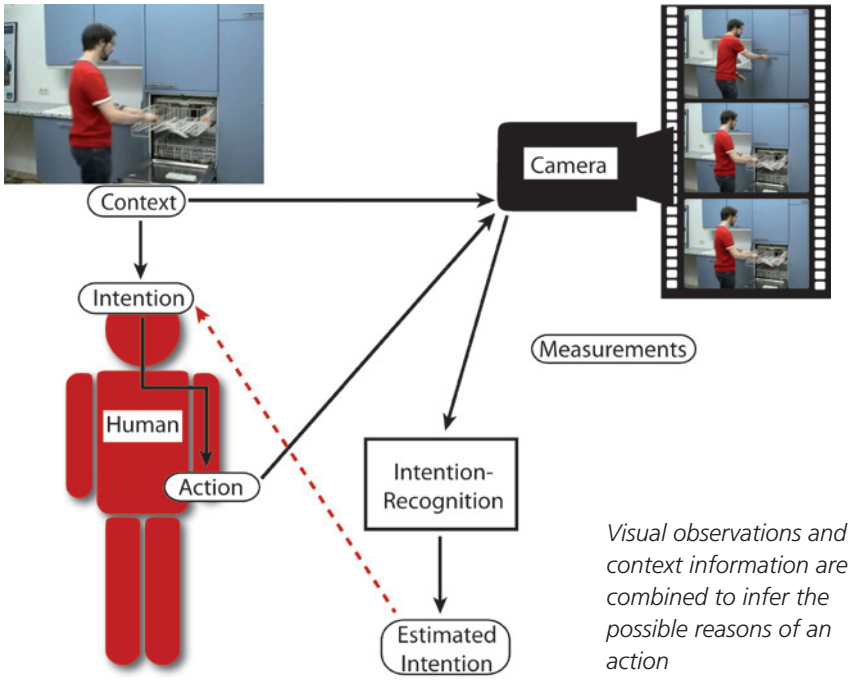
Prof. Dr.-Ing. Tanja Schultz
Cognitive Systems Laboratory (CSL)
http://csl.anthropomatik.kit.edu/

Intention Recognition

When communication is considered to be the exchange of information between two parties, then all communication can roughly be classified as either explicit or implicit. Explicit forms of communication are all actions with the goal of transferring information, such as talking to somebody. Implicit communication on the other hand occurs when information is obtained by observing an action not intended for communication. Watching a person opening a door will give an observer the information that the person may want to leave or let somebody in. There is no clear distinction between the two communication forms, as indeed body language and facial expressions provide important examples of mixture forms. They cannot be strictly explicit communication as they often occur without any recipients being present but neither are they are implicit as their only purpose is communication. This might also be dubbed involuntary or unconscious communication. A majority of human communication is of this form and its intercultural generality is well known, making it an important research topic in human-machine-interaction and humanoid robotics.

Implicit communication is often referred to as intention recognition as it is necessary to apply context knowledge to infer the possible reasons of an action to understand why it is performed. In other words, without knowing what a door is, it cannot be understood why it should be opened. Because there is no commonly applicable way to provide a machine with this context knowledge, it becomes extremely difficult to separate the desired information from irrelevant observations.

At the Intelligent Sensor-Actuator-Systems lab (ISAS), an intention recognition system is being developed based on visual observations of human actions. In experiments, a kitchen scenario is observed using video cameras and actions such as emptying a dishwasher are recognized. Because it is so difficult to give a machine insight into human motivations, machine learning techniques are commonly applied to derive probabilistic relationships between sequences of actions and possible intentions. In our case, hybrid dynamic Bayesian networks are used.



Visual observations and context information are combined to infer the possible reasons of an action

Hybrid dynamic Bayesian networks model the dependencies of random variables in a compact way by storing the conditional probabilities where causal dependencies exist. This allows a systematic fusion of different pieces of observed information, such as a person walking around, cupboards being opened and closed, and items picked up from the dishwasher to conclude that the dishwasher is being emptied in a robust way without confusing it with similar actions like clearing the table. Some of the inputs for this hybrid dynamic Bayesian network are created using computer vision techniques such as support vector machines or Hidden Markov Models provided by the Cognitive Systems lab and the cv:hci lab at KIT. These techniques detect simple motions such as walking or picking up objects, but the results suffer from a large amount of uncertainty. Other inputs are environmental factors including the time that makes it much less likely that breakfast will be eaten when it is late in the day.

Because the conditional probabilities necessary to define such a dynamic Bayesian network are not a priori known and it is not tractable to specify them all either, machine learning techniques are developed for these as well. The probability densities that define the network are derived using probability density estimation techniques developed at ISAS. Further improvements of such estimation techniques including approximations and optimizations for computational tractability are a subject of ongoing research.

- Dirk Gehrig, Peter Krauthausen, Lukas Rybok, Hildegard Kühne, Tanja Schultz, Uwe D. Hanebeck, Rainer Stiefelhagen, Combined Multi-Level Intention, Activity, and Motion Recognition for a Humanoid Robot, Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2011), San Francisco, California, USA, September, 2011.
- Peter Krauthausen, Uwe D. Hanebeck, A Model-Predictive Switching Approach To Efficient Intention Recognition, Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010),Taipei, Taiwan, October, 2010.

Uwe D. Hanebeck

Further information:
Prof. Dr.-Ing. Uwe D. Hanebeck
Intelligent Sensor-Actuator-Systems (ISAS)
http://isas.uka.de/en/Main_Page

interACT – a global partnership among the best

KIT is founding member of and key player in the international center for Advanced Communication Technologies.

The international center for Advanced Communication Technologies, interACT is a joint center between eight of the leading research institutions in the US, Europe and Asia:

- Carnegie Mellon University, Pittsburgh and Silicon Valley, CA, USA,
- Karlsruhe Institute of Technology, Karlsruhe, Germany,
- Hong Kong University of Science and Technology
- Waseda University, Tokyo, Japan,
- Italian Institute of Technology, Genova, Italy,
- National Institute of Information and Communications Technology, Tokyo, Japan
- University of Southern California, Los Angeles, USA
- Nara Institute of Science and Technology, Nara, Japan

Founded in 2004, the center is affiliated with the School of Computer Science at each institution. interACT offers international student exchange programs, seminars and academies and facilitates cross-national research projects.

Student Exchange
interACT is a network of international excellence in key areas of the world and is a model for international teamwork and cooperation. The center aims to exchange knowledge and to jointly train students in an environment of international immersion, where students get to know each other and then collaborate and participate in research projects. interACT's exchange program allows outstanding students to compile their final thesis in one of the partner-labs abroad.

interACT students are financially supported by the eight partners and by the "Baden-Württemberg-Stiftung", Germany. Up to now, more than 100 students were supported and via these exchanges more than 40 new co-operations between the partners were established.

Distinguished Lecture Series
interACT aims at sharing it's scientific know-how between the partners. Therefore a distinguished lecture series is organized in order to make the high competence available to a broader audience. Since the official founding of interACT in 2004 more than 25 distinguished lectures were given.

Summer Schools
Research workshops and academies with a cross-institutional and cross-national emphasis to achieve strong research collaboration and significant results are also part of interACT's mission. This ensures cross-cultural internationalization of the faculty, students and staff. Summer Schools are organised on an annual basis, each year hosted by another interACT-partner.

Joint Research and Projects
Joint research and joint projects are interACT's basis. Its partners lead and are involved in numerous international projects, they work closely together in many research projects.

Organization
The collaborative center follows local organizational principles at the eight institutions and operates in a light weight manner. The director of interACT is Prof. Alex Waibel. The center has an Advisory Board that reviews and oversees of the collaborative venture. All partners maintain interACT-offices. They serve as a contact point for researchers and students, wishing to do research in the other country or institution.

Prof. Alex Waibel, Center Director
Margit Rödder, Press and Communication

Further information:
http://interact.ira.uka.de/

Imprint
June 2012
Editorial Board: Tamim Asfour, Dagmar Gambichler, Sandra Tartarelli
Coordination: Sandra Tartarelli
Design, Layout: Wilfrid Schroeder, Bernd Königsamen
Print: Karl Elser Druck GmbH

This Newsletter is published half yearly by the KIT-Focus Anthropomatics and Robotics.
Office:
Adenauerring 2, 76131 Karlsruhe, Germany
Phone: + 49 721 608-44243
www.anthropomatics-robotics.kit.edu/