Public presentations and SUAB meeting at FAU

Felix Klein Building, Cauerstr. 11, 91058 Erlangen, Room K3-128

Thursday, December 1st, 2016

10:00 **Overview of the EARS project**
Prof. Dr.-Ing. Walter Kellermann
Friedrich-Alexander Universität Erlangen-Nürnberg

10:45 Coffee Break

11:00 **Presentation and live demo of EARS prototype system**
M.Sc.Rodolphe Gelin
SoftBank Robotics

12:00 Lunch Break

Invited Talks

13:00 **Personalized speech interfaces**
Dr. Martin Heckmann
Honda Research
(see abstract in appendix)

14:00 **Microphone blind directions and microphone arrangement in array processing**
Dr. Akihiko Sugiyama
NEC Information and Media Processing Labs
(see abstract in appendix)

15:00 Coffee Break

15:15 **SUAB Meeting**
SUAB and SB members

16:45 **Closing**

19:30 **Official EARS Dinner**
Restaurant "Basilikum" *Altstädter Kirchenplatz 2, 91054 Erlangen*

Abstracts of Invited Talks

Personalized speech interfaces

Dr. Martin Heckmann, Honda Research

In this presentation I will highlight recent results obtained at the Honda Research Institute Europe GmbH in the context of personalization of speech-based human-machine interfaces. I will first talk about the detection of word prominence. Thereby, I will discuss the performance of prominence detection from noisy audio signals, the contribution of additional visual information on the speaker's face and head movements as well as different strategies to fuse the two modalities. After that I will present a method to adapt the prominence detection to an individual speaker. The method is inspired by fMLLR, a well-known method in GMM/HMM-based speech recognition systems, and adapted to the SVM-based prominence detection. Next, I will talk about an advanced driver assistance systems (ADAS) which we currently develop to support the driver in inner-city driving and which is controlled via speech. This system will allow the driver to flexibly formulate his requests for assistance while the situation develops. In particular, when facing a left turn at an intersection the driver can delegate the task of observing the right side traffic to the system as he would do to a co-driver. The system will then inform him when there is an appropriate gap in the traffic to make the turn. Results of a user study we performed show that drivers largely prefer our proposed system to an alternative visual system or driving without any assistance. In this context I will show results on the estimation of the individual driver's left turning behavior. Based on these driver models the interaction with the driver can be personalized to further improve the usefulness of the system.

Microphone blind directions and microphone arrangement in array processing

Dr. Akihiko Sugiyama, NEC Information and Media Processing Labs.

This talk presents a blind direction problem in array processing and a solution by microphone arrangement. Multiple signal sources are discriminated by phase differences among signals received at multiple microphones. Signal sources located on the disk perpendicular to the array surface at the center of a linear array cannot be discriminated. It causes a problem in linear-array processing when the target and interfering signal sources are located on the disk. In case of a tablet PC which has a horizontal microphone placement in the landscape mode with signal sources located on the ground, the problem arises in the portrait mode which can be solved by a diagonal microphone placement. Evaluation results demonstrate that the diagonal microphone arrangement provides comparable correct speech recognition rates in both the landscape and the portrait mode. This solution is then extended to robot audition.