

Internship

Audio-Visual Prosody

We offer the possibility to excellent students of Electrical Engineering and Information Technology to do an internship at the Honda Research Institute Europe. During their project, the students have the opportunity to contribute to cutting edge research projects in an exciting and international environment.



Goal:

Speech plays an important role for the Human-Machine-Interface. Despite tremendous progress over the past years, speech interfaces are still perceived as little intuitive and not very robust. One reason for this is that conventional speech interfaces focus only on the meaning of the words. Yet speech is not only the words we say but also how we say them. This is reflected by prosody, the rhythm, stress, and intonation of speech. Such prosodic cues play a very important role for human-human communication, e.g. they structure the utterance in phrases (elements of a clause), emphasize the novel information in an utterance, and differentiate between questions or statements.

We at the Honda Research Institute Europe GmbH in Offenbach develop a system which integrates prosodic cues into the human machine interface. Thereby we do not only use the acoustic signal but also evaluate the movements of the head of the speaker and its facial expressions. This enables the system to infer the meaning of an utterance not only based on the words.

Tasks:

The topic of the internship is the extraction of prosodic information as well from the speaker's speech signal as from the movements of his face and the subsequent integration into a speech understanding system. A first system for the prosodic analysis using acoustical and visual cues is already available. The results of this internship will be integrated into this system.

The extraction of the prosodic information can be based on existing algorithms for speech signal analysis and the OpenCV, an open source library of programming functions for real time computer vision, other published algorithms, and also algorithms developed during the project. New algorithms can initially be developed in Matlab and are later to be ported to our ToolBOS framework, a component based integration and programming environment.

For the evaluation of the algorithms a previously recorded database with different speakers uttering sentences with varying prosody can be used. After successful evaluation the integration into a real-time system is targeted.

Contact:

Dr. Martin Heckmann
Honda Research Institute Europe
Carl-Legien-Strasse 30
D-63073 Offenbach/Main, Germany
Phone +49 (0)69 890 11-755
martin.heckmann@honda-ri.de