TECHNISCHE FAKULTÄT

Department Elektrotechnik-Elektronik-Informationstechnik

EEI-KOLLOQUIUM

Unified Speech/Audio Coding

Dr. Bernhard Grill
Fraunhofer Institut IIS A, Tennenlohe

Donnerstag, der 10.06.2010, 17^{15} Uhr
Cauerstraße 7/9, Hörsaal H5

Diskussionsleitung: Prof. Dr.-Ing. H. Gerhäuser

Up to a few years ago the low bit rate coding of speech signal on the one hand and general audio signals on the other hand could have rightfully been considered to be two separate branches in the art of digital signal processing. There were two scientific communities publishing and meeting at different conferences, and international standards were issued by clearly separated committees.

There was some reason to do so as the technical principles employed were sufficiently different and so were the fields of application. While speech coding was almost exclusively employed in communication applications such as cell phones and digital fixed line telephony, audio coding became popular in the form of mp3 players and the sound channels of digital TV or DVD discs. However, this is changing now with companies on the one hand trying to upgrade voice services to higher qualities and on the other hand TV sets receiving secondary functionalities that allow to employ them as communication devices.

Consequently, the two communities have been more and more peering into the other party's territory, trying to learn from the other side. Interestingly, for general audio codecs like mp3 or AAC when employed at very low bit rates, speech signals are among the most difficult to handle. The talk will try to explain what makes speech signals so special for these codecs and what possibilities exist to alleviate the difficulties they create - without compromising the quality for music signals. With this newly gained knowledge it was possible to design a new codec, which will be presented. It extends the usability of current audio codecs down to bit rates so far exclusively achievable with highly specialized voice codecs, which, however, are unable to handle any non-speech material.