Making speech tangible for a better understanding of human speech communication

HIDEKI KAWAHARA¹

¹Wakayama University, Wakayama, Wakayama, 640-8510, Japan
*e-mail: kawahara@sys.wakayama-u.ac.jp

Keywords: speech production, speech perception, speech analysis, speech synthesis

This talk presents the underlying concepts, technologies, applications and future of STRAIGHT, a framework for speech analysis, modification and resynthesis which was initially designed to facilitate speech perception research. Also introduced will be recent advances which may provide new opportunities in speech communication research. One concerns "Temporally variable multi-aspect morphing of arbitrarily many voices" and the other "SparkNG: Speech Production and Auditory perception Research Kernel: The Next Generation." The STRAIGHT and its applications are freely available to academics. The SparkNG package is open sourced. The future of STRAIGHT will start from the open source package, YANG vocoder.

Speech provides rich side information channels which modify/expand its linguistic contents. While the recent resurgence of machine learning technologies makes speech-based communication with smart machines practical and accessible, these rich side information channels are not well exploited. "Making speech tangible" by introducing tools which enable quantitative and precise as well as intuitive/direct manipulation of speech parameters will, I hope, lead to a better understanding of human speech communication.

Acknowledgement This research is currently supported by Grant-in-Aid for Scientific Research 15H02726 and 16K12464 of JSPS, Japan.

References
4) H. Kawahara, et al., APSIPA ASC 2013, 1-10 (2013)
5) H. Kawahara, Interspeech2016, Show & Tell, 1180-1181 (2016)