



Heiga Zen. Toshiba Research Europe Ltd. (UK)

Fundamentals and recent advances in HMM-based speech synthesis

Abstract

Statistical parametric speech synthesis based on HMMs has grown in popularity over the last years. In this talk, its system architecture is outlined, and then basic techniques used in the system, including algorithms for speech parameter generation from HMM, are described with simple examples. Relation to the unit selection approach and recent improvements are summarized. Techniques developed for increasing the flexibility and improving the speech quality are also reviewed.

Speaker Bio

Heiga Zen received the Dr.Eng. degree in computer science and engineering from Nagoya Institute of Technology in 2006. He is currently a Research Engineer in the Speech Technology Group of Toshiba Research Europe Ltd. Cambridge Research Laboratory. He was an intern researcher at the ATR Spoken Language Translation Research Laboratories in 2003 and an intern/co-op researcher at the IBM T. J. Watson Research Center from 2004 to 2005. From April 2006 to July 2008, he was a postdoctoral research associate at the Nagoya Institute of Technology. He has been working on HMM-based speech synthesis for 9 years after joining Prof. Tokuda's research group in 2000. He was also the main developer and maintainer of HTS, one of the main developers of the Festival Speech Synthesis System, one of the main developers of SPTK, and one of the active contributors to the hidden Markov model toolkit (HTK). He published over 10 journal papers and over 40 conference papers, and received 5 paper awards.

