ACOUSTICS2008/2702
Study on voice quality parameters for anger degree estimation

Yoshiko Arimoto[a], Sumio Ohno[b] and Hitoshi Iida[b]
[a]Tokyo University of Technology, Graduate School of Bionics, Information and Media Science, Faculty of MS, Iida labolatory, 1404-1 Katakura, Hachioji, 192-0982 Tokyo, Japan
[b]Tokyo University of Technology, 1404-1 Katakura, 192-0982 Hachioji, Japan

With great advance of automatic speech recognition (ASR) systems and a voice command system are demanded to be more sensitive to user's intention or emotion. These systems currently process linguistic information, but not process nonlinguistic information or paralinguistic information which users expressed during dialogs. For that reason, computers can obtain less information about a user through a dialog than human listeners can. If computers will recognize user's emotions conveyed by acoustic information, more appropriate response can be made toward users. For realization of emotion recognition, we have continued our study on anger degree estimation by both prosodic features and segmental features with anger utterances which were recorded during two kinds of pseudo-dialogs. This report focuses on only segmental features related to voice quality and examines them for capabilities to estimate anger degree. The first cepstral coefficient of anger utterances has been analyzed to obtain acoustic parameters related to spectral tilt in our previous works. In addition to the previous voice quality parameters, features related to amplitude of formants and harmonics were extracted from the same anger utterances as that in the previous works. Comparisons were made with the previous parameters to evaluate estimation accuracy of each voice quality parameters.