An Investigation on the Mandarin Prosody of a Parallel Multi-Speaking Rate Speech Corpus

江振宇, 湯政璋, 余秀敏, 王逸如, 陳信宏
Humanities and Social Sciences
kuo@chu.edu.tw

Abstract

In this paper, the prosody of a parallel multispeaking rate Mandarin read speech corpus is investigated. The corpus contains four parallel speech datasets uttered by a female professional announcer with various speech rates (SRs) of 4.40 (fast), 3.82 (normal), 2.97 (median) and 2.45 (slow) syllables/second. By using the unsupervised joint prosody labeling and modeling (PLM) method proposed previously, the relationship between SR and various prosodic features, including pause duration, patterns of three high-level prosodic constituents, and the break labels, are investigated. The analyses reported in this study could be very informative in developing prosody generation mechanism for text-tospeech and prosody modeling for automatic speech recognition in various SRs.