Title:
Tone recognition in continuous Mandarin Chinese

Author:
Jiang Wu, Stephen A. Zahorian, and Hongbing Hu (Department of Electrical and Computer Engineering, Binghamton University, Binghamton, NY, 13903)

Abstract:
Tones are important characteristics of Mandarin Chinese for conveying lexical meaning. Thus tone recognition, either explicit or implicit, is required for automatic recognition of Mandarin. Most literature on machine recognition of tones is based on syllables spoken in isolation or even machine-synthesized voices. This is likely due to the difficulty of recognizing tones from syllables extracted from conversational speech, even for native speakers of Mandarin. In this study, human and machine recognition of tones from continuous speech is evaluated and compared for four conditions: 1, vowel portions of syllables; 2, complete syllables; 3, syllable pairs; 4, groupings of three syllables. The syllables are extracted from the RASC-863 continuous Mandarin Chinese database. The human listeners are all native speakers of Mandarin. The automatic recognition is based on either Hidden Markov Models, or neural networks, and a combination of spectral/temporal, energy, and pitch features. When very little context is used (i.e., vowel segments only) the human and machine performance is comparable. However, as the context interval is increased, the human performance is better than the best machine performance.

Words in abstract: 176

Technical area: Speech Processing and Communication Systems

(PACS) Subject classification numbers(s): 43.72.Ar, 43.72.Ne

No preference for lecture versus poster