N-Best Decision for Thai Stressed Speech Recognition with Parallel Hidden Markov Model
Amomkul, P.; Kumhom, P.; Chanmongthai, K.

Summary: In integrating multi-isolated-word recognizers into a speech recognition for various stressed speeches, the best likelihood scopes as outputs of each recognizer are not guaranteed a correct recognition result. Since training sometimes does not cover all speakers, likelihood score of the correct recognition result is not the best and causes misrecognition. Moreover, the difference among recognizers also leads to mis-understanding. This paper proposes a decision-making method for Thai stressed speech recognition with parallel hidden Markov model. In this method, a voting scheme is applied on the words with the N-best likelihood score. Firstly, if the score margin between the first and the second-best is greater than a threshold, the voting is applied on the words with highest scores from each recognizer. If there is no clear winner, decided by considering the voting score, the next best score are included into the voting scheme. The process goes on until a winner is found or there is a tied score, its which case the average of the likelihood score of each tied word is used to decide the winner. The experiments were conducted with 4-stress speeches including angry, toward, loud, and neutral. It showed that the proposed method helped increase the recognition rate to 96.545% comparing with previous decision making techniques.

Digital Object Identifier 10.1109/ISPACS.2005.1595337