

Unit selection for Malay text-to-speech system using segmental context and simulated annealing

Abstract:

Unit selection method has become the main approach in speech synthesis. The increasing size of recorded speech has resulted in better synthesis speech quality but at the same time also resulted in more expensive computational effort. Therefore, this paper proposes a combination of segmental context matching procedure and Simulated Annealing (SA) in unit selection to improve the quality of synthetic speech and reduce the computational time. The process of unit selection is based on minimization of two costs: target cost and join cost. The segmental context (target cost), the first stage of unit selection matching procedure used to narrow down the search space, followed by an optimization method which is SA to find the units sequence with minimum join cost. Result shows that the synthesis words produced by the proposed system are 15.48% better compared to previous version of corpus-based Malay Text-to-Speech system. Future works may focus on combining SA with other heuristic methods to further enhancing the performance of unit selection.