

Excitation Modeling for Hidden Markov Model-Based Speech Synthesis Based on Wavelet Analysis

Authors : M. Kiran Reddy, K. Sreenivasa Rao

Abstract : The conventional Hidden Markov Model (HMM)-based speech synthesis system (HTS) uses only a pulse excitation model, which significantly differs from natural excitation signal. Hence, buzziness can be perceived in the speech generated using HTS. This paper proposes an efficient excitation modeling method that can significantly reduce the buzziness, and improve the quality of HMM-based speech synthesis. The proposed approach models the pitch-synchronous residual frames extracted from the residual excitation signal. Each pitch synchronous residual frame is parameterized using 30 wavelet coefficients. These 30 wavelet coefficients are found to accurately capture the perceptually important information present in the residual waveform. In synthesis phase, the residual frames are reconstructed from the generated wavelet coefficients and are pitch-synchronously overlap-added to generate the excitation signal. The proposed excitation modeling method is integrated into HMM-based speech synthesis system. Evaluation results indicate that the speech synthesized by the proposed excitation model is significantly better than the speech generated using state-of-the-art excitation modeling methods.

Keywords : excitation modeling, hidden Markov models, pitch-synchronous frames, speech synthesis, wavelet coefficients

Conference Title : ICCESP 2019 : International Conference on Computer Engineering and Signal Processing

Conference Location : Bangkok, Thailand

Conference Dates : January 17-18, 2019