1479: CCI-MOBILE: AUTO-LSP BASED SPEECH ENHANCEMENT WITH COCHLEAR IMPLANT LISTENERS USING CONVOLUTIONAL NEURAL NETWORK CONSTRAINT MAPPING

Nursadul Mamun, John H.L. Hansen
University of Texas at Dallas, Dallas, TX, USA

Speech perception in the presence of real-world distortion, such as competing for babble noise, is difficult for cochlear implant (CI) users due to the limited frequency resolution of speech produced with CI devices. One approach to address this issue is to suppress real-world noise without introducing processing artifacts into the speech signal that are perceived by CI listeners. Traditional speech enhancement algorithms such as MMSE have been shown to be effective in stationary noise, however, their application has limited success for non-stationary noise such as babble noise. Earlier unsupervised speech enhancement methods such as Auto-LSP (Hansen, Clements, 1991) and ACE-1: Auditory Constrained Enhancement (Nandkumar, Hansen, 1995) have been formulated which apply speech production and auditory based time-frequency spectral constraints to achieve improved overall speech quality. Recently, machine learning strategies have been shown to be successful in addressing such complex noisy environments. This study proposes to advance a previous Auto-LSP-based speech enhancement technique using a new strategy to apply constraints with a fully convolutional neural network (CNN) for cochlear implant users. The proposed algorithm calculates linear predictive coefficients (LPC) and line-spectral pair (LSP) parameters from each input signal frame over time. Inter- and Intra-frame constraints are then applied to line-spectral pair parameters to ensure that vocal tract characteristics do not vary widely from frame-to-frame overtime when speech is present. A fully connected convolutional neural network (CNN) is then deployed to predict the enhanced LSP parameters from their noisy versions. The proposed algorithm contains a limited number of parameters compared to traditional fully connected neural networks, which suggests a greater opportunity to transition to real-world CI/HA platforms. A large size speech training set is used to train the CNN network which ensures a powerful model capable of estimating the nonlinear mapping between noisy and clean speech via supervised learning. Finally, a constrained Wiener filter, present in the original Auto-LSP solution, is incorporated to predict the clean speech signal from the enhanced time-frequency speech constrained LPC parameters. Evaluation using objective measures suggests that the proposed algorithm achieves measurable improvement versus existing baseline systems, and therefore represents a viable option for implementation and field evaluation on the UTDallas CCi-Mobile research platform as a preprocessor.

Hansen, J.H.L., Clements, M., "Constrained Iterative Speech Enhancement with Application to Speech Recognition," IEEE Transactions on Signal Processing, vol. 39, no. 4, pp. 795-805, April 1991.

Nandkumar, S., Hansen, J.H.L., "Dual-Channel Iterative Speech Enhancement with Constraints Based on an Auditory Spectrum," IEEE Transactions on Speech & Audio Processing, vol. 3, no. 1, pp. 22-34, January 1995.

This work was supported by Grant No. R01 DC016839-02 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.