SPEECH INTELLIGIBILITY ENHANCEMENT USING SPEECH TRANSIENTS EXTRACTED BY A WAVELET PACKET-BASED REAL-TIME ALGORITHM

Daniel Rasetshwane
Department of Electrical and Computer Engineering
University of Pittsburgh

Abstract

Studies have shown that transient speech, which is associated with consonants, transitions between consonants and vowels, and transitions within some vowels, is an important cue for identifying and discriminating speech sounds. Emphasis of transient speech can improve the intelligibility of speech in background noise, but methods to demonstrate this improvement have either identified transient speech manually or proposed algorithms that cannot be implemented to run in real-time. The goal of this study is to develop an algorithm to automatically extract transient speech in real-time. The algorithm involves the use of a function, which we term the transitivity function, to characterize the rate of change of wavelet coefficients of a wavelet packet transform representation of a speech signal. The transitivity function is large and positive when a signal is changing rapidly and small when a signal is in steady state. The extracted transient speech signal is used to create modified speech, for speech intelligibility enhancement, by combining the amplified transient speech with original speech and adjusting the modified speech so that its energy is equal to that of the original speech. An index to quantify the transient nature of speech (specifically, the extent to which the onsets and offsets of formants are emphasized compared to steady segments) is described and applied to several versions of transient speech and to processed speech that have been described in literature. The transient extraction algorithm includes parameters which, when varied, influence the intelligibility of the extracted transient speech. The use of psycho-acoustic testing to select the best values for these parameters and to evaluate the improvements in speech intelligibility over original speech provided by modified speech is described.

Speaker's Biographical Sketch

Daniel Rasetshwane is a PhD candidate in the Department of Electrical and Computer Engineering at the University of Pittsburgh. He received his B.S. and M.S. in Electrical Engineering from the University of Pittsburgh in 2002 and 2005, respectively. He is a student member of the IEEE and the IEEE Signal Processing Society. His research interests include speech intelligibility enhancement, speech enhancement, natural language processing and digital signal processing.

DATE: Wednesday, February 11, 2009
TIME: 12:00 pm
LOCATION: 424 Benedum Hall