Santosh S. Pratapwar, Reduction of background noise in artificial larynx, M. Tech. Thesis, Department of Electrical Engineering, Indian Institute of Technology Bombay, 2004.

Supervisor(s): P. C. Pandey

Abstract -Transcervical electrolarynx is of great help in verbal communication to a large number of laryngectomy patients. The device is held against the neck, and the vibrations generated move up the vocal tract to produce useful speech. Its intelligibility suffers from the presence of self or background noise, caused by leakage of the acoustic energy from the vibrator. It has been shown earlier that the spectral subtraction technique, developed for enhancement of noisy speech, can be applied in a pitch-synchronous manner for reducing the leakage noise in electrolaryngeal speech. Implemented method had two modes, noise estimation mode and speech enhancement mode. Averaged magnitude spectrum of noise, obtained with the lips closed in the noise estimation mode, is subtracted from the magnitude spectrum of the noisy speech and the signal is reconstructed using the original phase spectrum. The noise spectrum is taken to be stationary over the entire duration of speech enhancement mode. But actually the background noise varies because of variations in the place of coupling of vibrator to the neck tissue and the amount of coupling. This results in variations in the effectiveness of noise enhancement over an extended period. The objective of the project is to investigate signal-processing techniques which will estimate noise continuously thereby improving the quality of the speech output. A single input technique for reducing the background noise in electrolaryngeal speech signal, using spectral subtraction in a pitch synchronous manner is implemented where updating of the noise magnitude spectrum is carried out using quantile based noise estimation, which does not require speech/non-speech detection. Investigations for selection of appropriate quantile values for estimation of noise spectrum from noisy speech, showed that a dynamic estimation of frequency dependent quantile values, based on signal strength, resulted in significant improvement in quality and intelligibility of speech and was effective during non-speech segments.