Sumedha A. Kshirsagar, A speech training aid for the hearing impaired, M. Tech. Thesis, Department of Electrical Engineering, Indian Institute of Technology Bombay, 1998.

Supervisor(s): Dr. P.C. Pandey

Abstract – Lack of auditory feedback in the hearing impaired results in failure to produce intelligible speech. Providing visual feedback of efforts involved in speech production process would help such a person in learning to speak. A system for analyzing speech for extracting pitch, intensity and vocal tract area, and displaying these in real time has been earlier developed at IIT Bombay. This system is based on TI/TMS320C25 (16-bit fixed point processor) DSP-board having on-board memory sharable with PC. The implementation of the algorithms for estimation of vocal tract area, pitch, and intensity has been done on the DSP-board and user interfacing is handled by the PC. The variation of these parameters with time can also be displayed for a duration of 1.28 second. "Areagram" is a display for variation of vocal tract area with time. The main limitation of the system is that, the vocal tract shape estimation cannot be made during stop closures as result of very low energy in the speech signal and these are inconsistencies in the shape display even during the vowel segment.

In this project, the areagram display software is modified to more realistic and consistent results. The modification is done for normalization of area values and their mapping to 16 grey levels to be displayed on the areagram. Care has been taken that erroneous peaks i9n area distribution do not affect this normalization. This gave more meaningful results when areagrams for individual vowels and vowel sequences were compared. An attempt was also made to improve this system for estimation over weak energy phonemes. For extraction of information during stop closures, based on the area variation just before and after the closure, off-line processing of areagram matrix was done. Spatial low-pass filtering for 2-D interpolation did not help in extracting the information regarding place of closure. Normalizing speech segments with respect to average magnitude improved the estimation over semi-vowels. The minimum signal strength required for meaningful estimation has also come down by a factor of 2. The estimation algorithm was also implemented using floating point arithmetic to investigate the improvement in estimation over weak energy signal, but this was not much better than that achieved by normalization of input segments and real-time analysis in fixed point arithmetic.