



**FORMANT ANALYSIS AND SYNTHESIS OF (HINDI) VOWELS AND THEIR PERCEPTUAL LIMITS**

**M. Ganesan, S.S. Agrawal and K.D. Pavate**

**ABSTRACT**

In general formant frequencies are most significant parameters required for synthesis of speech sounds particularly vowels. The present study was conducted to estimate the first four formant frequencies of the (Hindi) Vowel sound and to synthesize vowel sounds based on the computed values. The minimum and maximum values of formants for accurate perception of these sounds were also determined. The speech samples consists of ten Hindi vowels spoken in a fixed context by a male speaker. These samples were digitized using a 12 bit A/D converter and a LSI-11/23 C.P.U. based mini-computer system. In order to reduce the computation, the end points of the utterance was detected by an end point algorithms which is based on energy calculation and zero crossing rate. The digitized signal was pre-emphasized for high frequency components. A 256 point Hamming Window was applied to emphasize the zero-frequency region of the spectrum. For the estimation of formant frequencies the cepstrally smoothed log-spectra is used. The peaks of this spectra corresponding to formant frequencies were computed for all the vowels under study.

The computed values of formant frequencies and their amplitudes and other parameters such as formant bandwidths duration etc. as fixed values were used to synthesize the vowels. A cascade/parallel formant synthesizer originally proposed by D.H. Klatt has been simulated for this purpose. The advantage of this synthesizer is that with minimum number of time varying parameters one can produce high quality synthetic speech. The synthesizer was suitably modified for the compatability with the dynamic range of the digital to analog converter used for this purpose. All the ten vowels were synthesized with the calculated parameters and subjected to listening tests. It was observed that the quality of synthetic vowels improves with the availability of time varying information related to pitch and intensity in addition to the data of formants.

For perceptual tests the values of formant frequency were increased and decreased in steps of 10Hz and 50Hz. The maximum and minimum values of each formant was determined to have correct recognition of each of the vowels. It has been found that the different formants can be varied to different limits without loosing high intelligibility.

Central Electronics Engineering Research Institute Centre, CSIR  
Complex, Hillside Road, New Delhi-12 INDIA  
Ph: (011) 581467, (011) 583172