

P. K. Lehana and P. C. Pandey, Speech synthesis with pitch modification using harmonic plus noise model, J. Acoust. Soc. Am., vol. 114, pp. 2395, 2003

Contact: Prof. P. C. Pandey

Department of Electrical Engineering,

Indian Institute of Technology Bombay, Powai, Mumbai.

mailto: pcpandey@ee.iitb.ac.in

***Abstract* - In harmonic plus noise model (HNM) based speech synthesis, the input signal is modeled as two parts: the harmonic part using amplitudes and phases of the harmonics of the fundamental and the noise part using an all-pole filter excited by random white gaussian noise. This method requires relatively less number of parameters and computations, provides good quality output, and permits pitch and time scaling without explicit estimation of vocal tract parameters. Pitch scaling to synthesize the speech with interpolated original amplitudes and phases at the multiples of the scaled pitch frequency results in an unnatural quality. Our investigation for obtaining natural quality output showed that the frequency scale of the amplitudes and phases of the harmonics of the original signal needed to be modified by a speaker dependent warping function. The function was obtained by studying the relationship between pitch frequency and formant frequencies for the three cardinal vowels naturally occurring with different pitches in a passage with intonation. Listening tests showed that good quality speech was obtained by linear frequency scaling of the amplitude and phase spectra, by the same factor as the pitch scaling.**