Short Time Fourier Transform – A Mathematical Tool to Speech Processing System

Sanghamitra Mohanty
sangham1@rediffmail.com

Suman Bhattacharya
suman1_bh@yahoo.com

RC-ILTS
Dept. of Comp. Sc. & Appln.
Utkal University, Bhubaneswar
Orissa, India

ABSTRACT

In several areas of applied sciences and engineering, the representation of signals or other functions by sums of sinusoids or complex exponentials leads to convenient solutions to problems and often to greater insight into physical phenomenon than is available by other means. The model for the production for a steady state speech sound such as a vowel or fricative simply consists of a linear system excited by a source, which is either periodically or randomly varying with time. The spectrum of the output of such a model would be the product of the frequency response of the vocal tract system and the spectrum of the excitation. On the contrary the speech waveform are generally much more complicated than simply a sustained vowel or fricative sound. Thus the standard Fourier representations those are appropriate for periodic, transient or stationary random signals are not directly applicable to the representation of speech signals whose properties change markedly as a function for time. The short time analysis principle is a valid approach to speech processing. The spectral properties of speech can be assumed to change relatively slowly with time. The Short Time Fourier Transform (STFT) decomposes a speech signal into frequency components around a particular point in time. It is a function of both time and frequency. It is often used in the analysis of speech, since speech signals are non-stationary. The window functions that we normally use are Hann, Hamming, Barlett and Kaiser. We can't determine time and frequency precisely because of uncertainty principle. The windowed speech signal can be computed employing a FFT algorithm.