

Name of Scholar : Pankaj Sharma
Name of Supervisors : Prof. Khalil Ahmad and Dr. Z.A. Jaffery
Department : Mathematics
Title of the Thesis : On Some Aspects of Wavelets and Its
Applications in Speech Analysis

Abstract

From the ancient time, when human invented spoken languages, speech has been the most direct way for human to convey information to one another. Up to the present, the communication using spoken speech is still the most dominant and common service in the communication networks. The speech signal is now extended, through technological mediation such as telephone, movies, radio, television, and the Internet.

In this thesis, we study the applications of wavelet analysis and wavelet packet analysis for denoising speech signal. For speech signals we have recorded our own data as well as considered some existing data in Matlab software. The thesis as such has been classified into six chapters. Chapter-0 gives a brief survey of some important historical developments in the theory of wavelets and wavelet packets, which leads to the motivation of the study in the present thesis. The term digital signal processing, some basic definition of signals, problems in processing signals are mentioned in this chapter.

All the basic concepts, notations and essential tools to be used throughout the thesis have been listed in Chapter-1. The whole chapter is divided into four sections. The first section deals with the basic definition, characteristics of wavelets and wavelet analysis. The second section covers a short tour of wavelet transform and rest of the section deals with speech signal processing and reviews of wavelet based speech signal processing. In Chapter-2, we have studied a speech denoising method based on decomposition of speech signal in wavelet transform domain. Thresholding of noisy coefficients in speech signal is the main concern of this chapter. To choose optimal threshold value, we have computed the minimum error between detailed coefficients of noisy speech signal and the original noise free signal.

In Chapter-3, we have considered double filtering of noisy speech signal. Wavelet decomposition method has been considered to decompose noisy speech signal. After choosing optimal threshold value soft thresholding has been applied. Soft thresholding set all coefficients values equal to zero, whose absolute values is below optimal threshold value and reduces the magnitude of the remaining coefficients by the threshold value. This estimated speech signal is further passed thorough the wiener filter which makes estimated speech signal more robust in nature.

In Chapter-4, wavelet packets denoising methods have been considered. Wavelet packet transform is a simple but powerful extension of wavelets and multiresolution analysis, which yields basis functions with better frequency localization at the cost of a slightly more expensive transform. A very popular wavelet proposed by Daubechies (db4) is used as a mother wavelet. Since most of the noise lies at the first decomposition level, so a global threshold scheme has been used in this chapter.

To choose optimal decomposition level is one of the corner stone in speech signal denoising through wavelet packet transform. In Chapter-5 the concept of Shannon entropy has been discussed to achieve optimal decomposition level. Entropy is a common concept in many fields, mainly in signal processing and information theory. In this chapter, the non normalized Shannon entropy is used since the speech signal has imbalance, non-stationary, different frequency component and the different energy distribution. The assumption is that the signal spectrum is more organized in speech segments than in noise segments, the entropy will be maximal, if the signal is only noise, and minimal if it is a noiseless speech signal. Compute Shannon entropy for approximation coefficients Ea_j and detailed coefficients Ed_j at successive level of decomposition, where, j is the level of decomposition. The level at which $Ea_j < Ed_j$ has been considered as the optimal level of decomposition. Threshold value is estimated with wavelet packet coefficients selection rule, i.e. $thr = \sigma\sqrt{2(\log N)/N}$ using a penalization method proposed by Lucien Birge and Pascal Massart