

Speech Enhancement in a Noisy Environment Using Sub-Band Processing

K. Sravanthi¹, M. S. Anuradha²

¹(M.Tech 2/2, ECE dept, Andhra University, INDIA)

²(Sr. Asst. professor, ECE dept, Andhra University, INDIA)

Abstract: One of the methods for recovering a signal corrupted by additive noise using Weiner filtering approach. This paper mainly deals with speech enhancement/recovering using Weiner filtering technique. It uses the Weighted Overlap Add filter bank implemented using digital signal processing techniques. The spectral enhancement is achieved by Voice Activity Detector in conjunction with Weiner filtering, and the results were discussed for different noise floor levels for a stationary recorded input signal.

Keywords: speech enhancement, WOLA filter bank, VAD, weiner filtering.

I. Introduction

Speech enhancement is most important in speech signal processing. Several techniques have been proposed for this purpose. Weiner filtering is one of the speech processing methods which involve linear estimation of a desired signal sequence from previous sequence [5]. The main approach is to minimize the mean square error signal that is defined as the difference between desired response and the actual filter output [2]. WOLA (Weighted Overlap and Add) filter bank approach is utilized to obtain the objective which is shown in the below fig1. It decomposes the original noisy speech signal into different frequency sub-bands [3]. Filter bank processing can be considered as a divide-and-conquer approach within signal processing, since larger problems are sub-divided into many smaller problems. As the original signal is decomposed into sub bands it achieves faster convergence rate and low mean square error. In speech processing systems usually the voice activity detector (VAD) is used, which identifies the time intervals with speech signal and those with only background noise. The outcome from the VAD allows discarding the silent frames from further processing, as they carry no speech information, therefore improving the performance of the rest of the system [4].

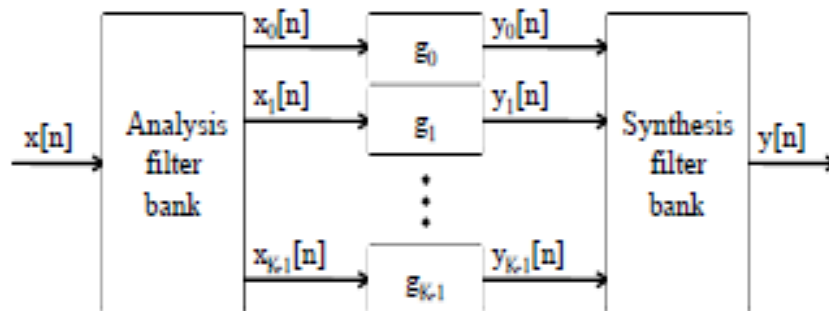


FIG: 1 simplified block diagram for speech enhancement.

II. Theoretical Aspect

The WOLA (Weighted Overlap and Add) filter bank is an efficient method used to implement a uniformly distributed multi-channel filter bank. It decomposes the input signal into a series of separate frequency bands by this it minimizes the overlap between adjacent bands. This results in a pure tone input to the analysis filter bank. For WOLA filter bank with $K/2$ frequency bands and band outputs are decimated by K/OS ($OS=2$ was used above). [1].

The following three stages are represented in the following fig: 2

There are three main stages:

1. analysis filter bank
2. Weiner filtering
3. synthesis filter bank

ANALYSIS FILTER BANK

The operations are:

- Read D input block samples
- Shift the block samples into the input FIFO buffer $u[n]$
- Applying window and time-fold to K samples
- Applying circular shift by $K/2$ samples in order to produce a zero-phase signal for the FFT.
- Take K-point FFT (even or odd)
- Applying channel gains to (complex) frequency data

III. Wiener Filtering

The Wiener filtering is optimal in terms of the mean square error. It minimizes the overall mean square error in the process. They operate on the past and present of the input signal. A Wiener filter used is a linear filter. The linear filter assumes that the inputs are stationary. The goal of the Wiener filter is to filter out noise that has corrupted a signal. The received sub band signals comprise a mixture of clean speech plus degrading noise, i.e.,

$$x_k[n] = s_k[n] + v_k[n], \quad (1)$$

Where $s_k[n]$ denotes clean sub band speech and $v_k[n]$ denote sub band noise.

Here, the filter bank analysis stage produces a set of sub-band signals $s_k[n]$ from a time signal $s[n]$. The sub-band signals are subject to sub-band processors, denoted as g_k , to compute sub-band output signals $y_k[n] = g_k\{x_k[n]\}$.

The sub-band processing implemented in this paper is Adaptive Wiener Filtering. The Wiener filter provides an optimal estimate of a signal embedded in additive noise by minimizing the Mean Squares Error (MSE),

$$E\{|e_k[n]|^2\} = E\{|s_k[n] - g_k[n] * x_k[n]|^2\} \quad (2)$$

Where $E\{\cdot\}$ is the expectation operator. The Wiener filtering subtracts the noise signal spectrum estimate from the input signal spectrum to acquire an approximate of the original speech signal spectrum. The speech signal which is approximated using wiener filtering is obtained by multiplying the input signal with the obtained gain in wiener filtering as shown below,

$$Y=G*X \quad (3)$$

IV. Synthesis Filter Bank

The Inverse WOLA (IWOLA) performs the opposite function to the WOLA filter bank. It takes the K channels, and re-synthesizes these to form a single time-domain sequence.

The operations are

- Take inverse FFT(IFFT) of (complex) input (even or odd)
- Applying circular shift by $K/2$ samples, to counter-act the circular shift used in the analysis stage
- Applying synthesis window
- Accumulate into output FIFO, shift out L samples

Where:

- D is the input block size
- L is the input window size
- K is the FFT size
- DF is the decimation factor
- OS is the over sampling factor.

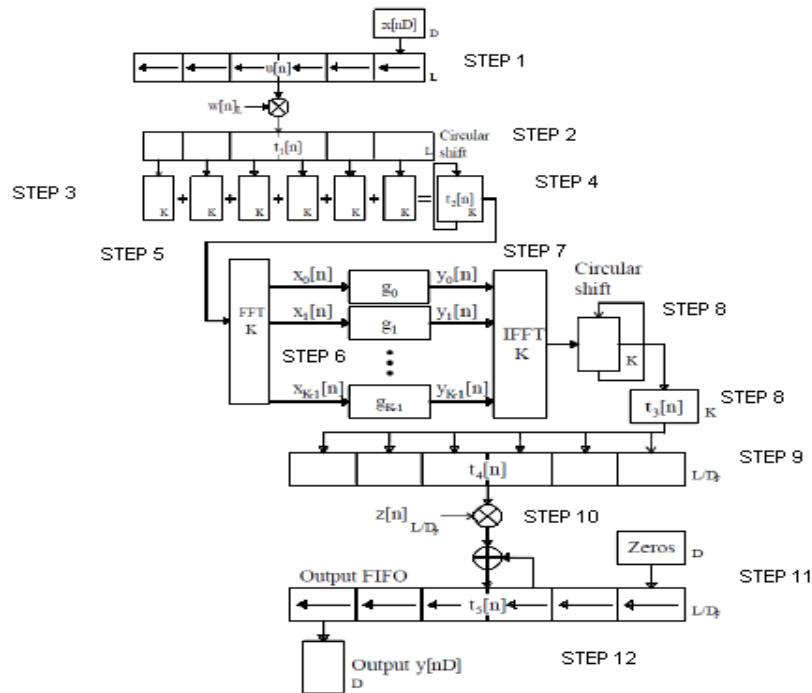


Fig: 2 block representation of WOLA filter bank

III. Matlab Implementation

The input signal which is used for the speech processing in this work is recorded through a mobile phone in the .amr format and this is converted into .wav file. The duration of the signal considered is 10seconds with 8 kHz sampling frequency with a message containing the information “Hello..... Hello”. The VAD data is calculated with this recorded signal. VAD=1 indicates when there is a speech signal and VAD decision =0 indicates where there is a noise is represented in fig1. The SNR threshold is chosen to be 1 for VAD decision. After Weiner filtering the noisy speech signal and the output signal are presented in fig4.

Where:

- the input block size $D = 16$
- the input window size $L = 128$
- the FFT size $K = 32$
- the decimation factor $DF = 1$
- the over sampling factor $OS = 2$

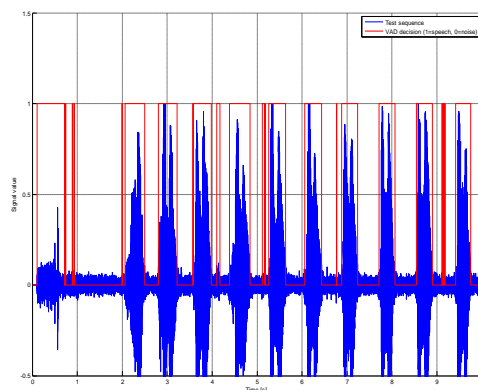


Fig: 3 input noisy speech signal and the VAD data.

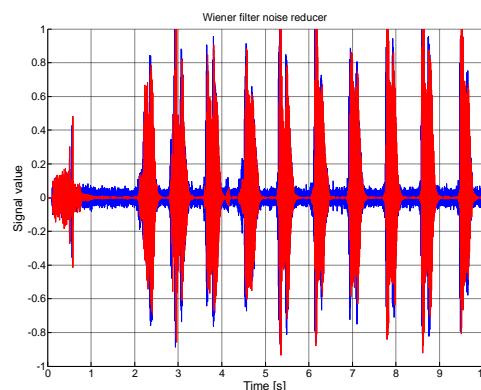


Fig: 4 speech signal after Wiener filtering without noise.

The additional noise is added to the original recorded input signal with different noise floors and the results are presented in the following figures.

Case-1:

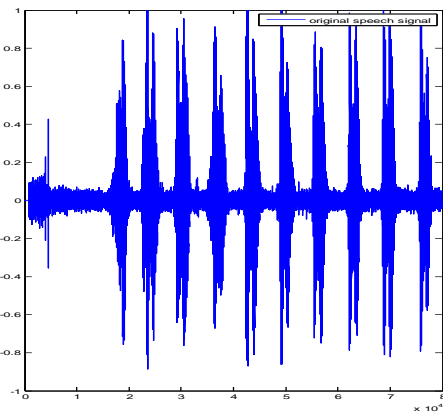


Fig 5: original speech signal

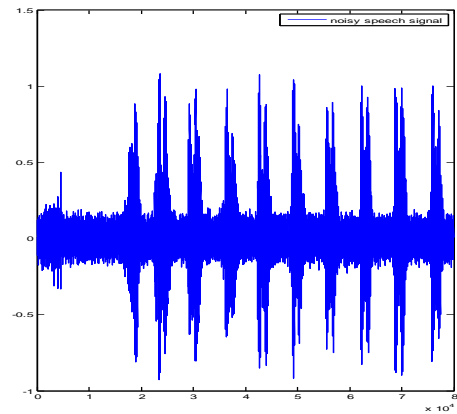


fig:6 noisy speech signal

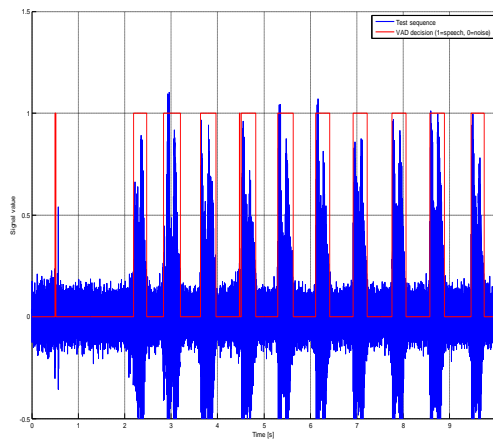


Fig: 7 input noisy speech signal and the VAD data

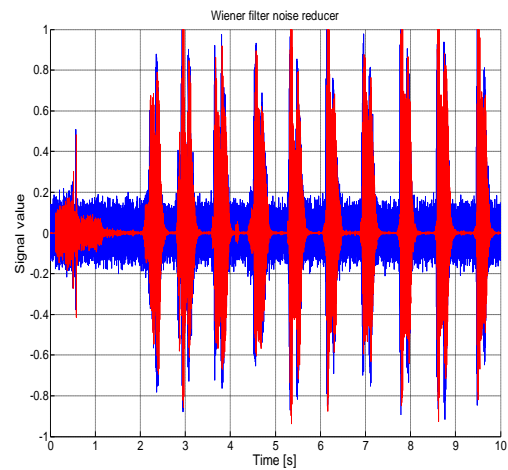


fig: 8 speech signal after Wiener filtering without noise

Case 2:

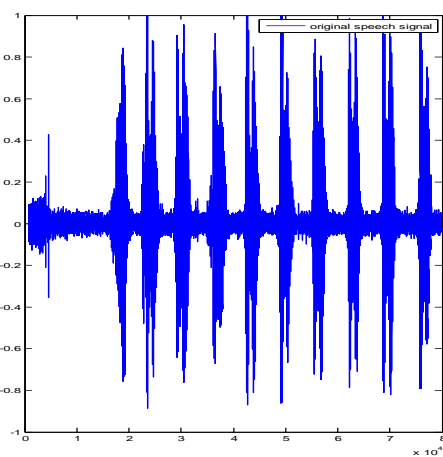


Fig 9: original speech signal

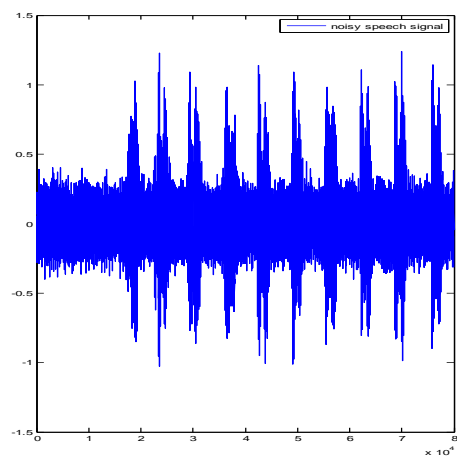


fig: 10 noisy speech signal

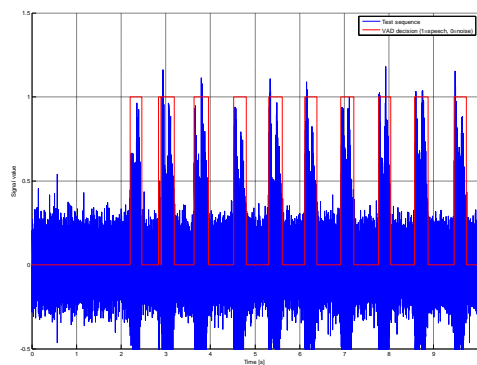


Fig: 11 input noisy speech signal and the VAD data

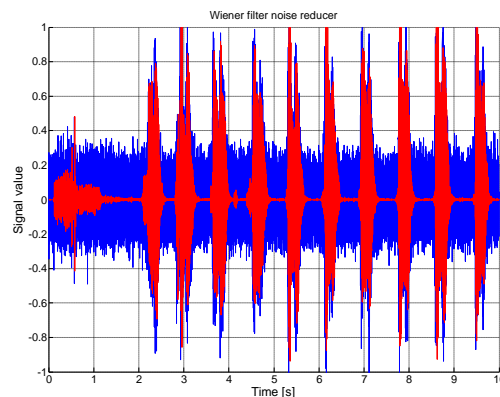


fig: 12 speech signal after Wiener filtering without noise

Case 3:

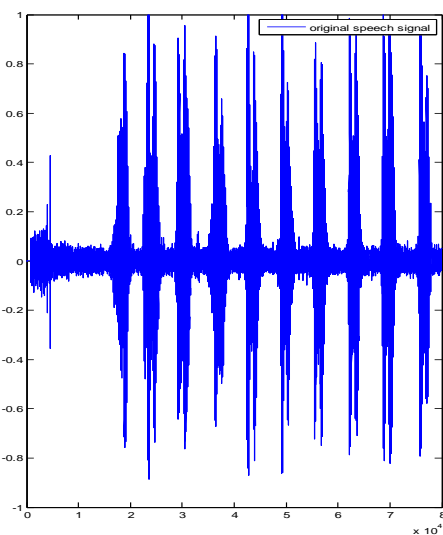


Fig: 13 original speech signal

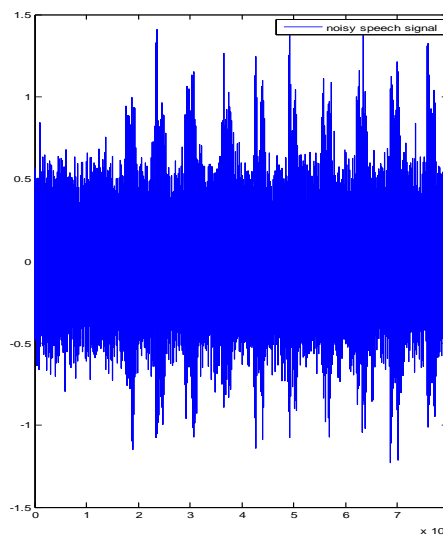


fig: 14 noisy speech signal

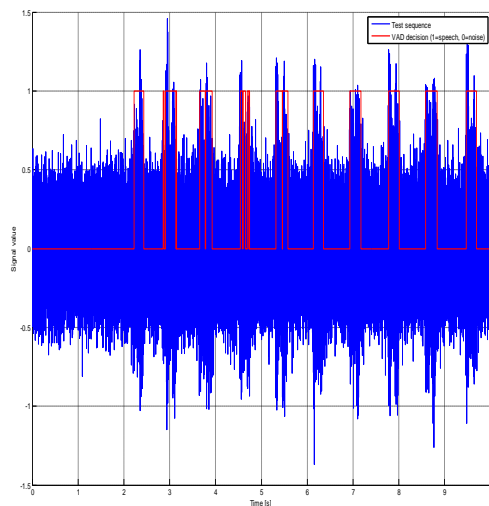


Fig: 15 input noisy speech signal and the VAD data

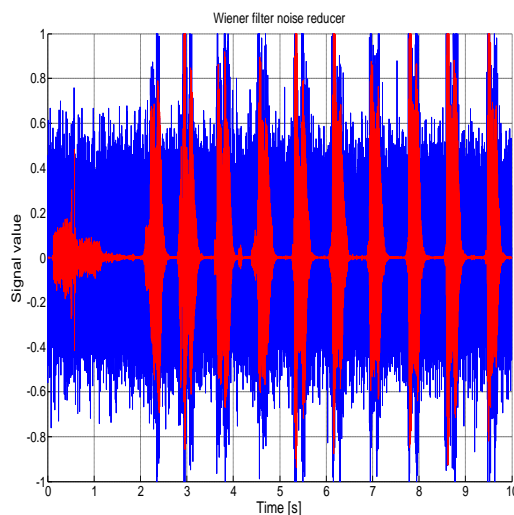


fig: 16 speech signal after Wiener filtering without noise

IV. Conclusion

This paper discusses the Noise suppression in speech signal using WOLA filter bank to divide the signal into sub-bands. WOLA filter bank is implemented using Hamming windowing technique and the sub-band processing improves better efficiency under different noise conditions. Noise suppression is achieved through Voice Activity Detector and Weiner filter. This paper is limited to the stationary signals. For real time applications adaptive Weiner filtering approach may be applied and the efficiency can be improved with increased sampling rates and also by selecting different windowing techniques for sub-band processing.

Reference

- [1] A flexible filter bank structure for extensive signal manipulations in digital hearing aids by Robert Brennan Todd Schneider dspFactory, Waterloo Ontario, Canada, N2J 4V1
- [2] Technion { Israel Institute of Technology, Department of Electrical Engineering Estimation and Identification in Dynamical Systems (048825) Lecture Notes, Fall 2009, Prof. N. Shimkin
- [3] Wideband High-Resolution Filterbank Firmware: Acceptance Testing and Excision Filtering K. L. Harman and C. Potter
- [4] Modified LTSE-VAD algorithm for applications requiring reduced silence frame misclassification by Iker Luengo, Eva Navas, Igor Odriozola, Ibon Saratxaga, Inmaculada Hernaez, Iñaki Sainz, Daniel Erro
- [5] Speech Signal Processing School of Electronic Information, Wuhan University.

BOOK

- [6] Adaptive Filtering Primer with MATLAB - Poularikas and Ramadan.