Speech Denoising and Speech Enhancement Using Wavelet Filter

Kavita Sharma(MTech*) Prateek Haksar (MTech*)
(Department of electronics and communication, MITS, Laxmangarh (UGC-1956)
(Department of electronics and communication, MEC, Bikaner(RTU)

ABSTRACT
Speech plays an important role in multimedia system. Speech enhancement is to remove noise from speech for multimedia systems. Noise is an unwanted interference in any form of communication generally degrades the quality of the information signal. During transmission and reception signals are often corrupted by noise which can cause severe problems for downstream processing and user perception. Therefore an automated means of removing the noise would be an invaluable first stage for many signal processing tasks. Denoising has long been a focus of research and yet there always remains room for improvement. There are so many techniques to improve the signal quality or to regenerate the signal. In this paper we present a method for speech denoising using wavelet filter. Firstly the speech is recorded using matlab and played, finally the recorded speech is processed using wavelet filter.

Keywords – Speech enhancement, Wavelet filter.

I. INTRODUCTION

The fundamental purpose of speech is communication, i.e., the transmission of messages. In speech production, as well as in many human-engineered electronic communication systems, the information to be transmitted is encoded in the form of a continuously varying (analog) waveform that can be transmitted, recorded, manipulated, and ultimately decoded by a human listener. Development of speaker identification systems began as early as the 1960s with exploration into voiceprint analysis, where characteristics of an individual's voice were thought to be able to characterize the uniqueness of an individual much like a fingerprint. The early systems had many flaws and research ensued to derive a more reliable method of predicting the correlation between two sets of speech utterances. Speaker identification research continues today under the realm of the field of digital signal processing where many advances have taken place in recent years.

Speech recognition is the process by which a computer (or other type of machine) identifies spoken words. Basically, it means talking to our computer, and having it correctly recognize what we are saying. The most challenging sources of variations in speech are speaker characteristics including accent, co-articulation and background noise. The filter bank in the front-end of a speech recognition system mimics the function of the basilar membrane. It is believed that the closer the band subdivision is to human perception, the better the recognition results. Filter constructed from estimation of clean speech and noise for speech enhancement in speech recognition systems. Experience has been gained in basic filtering of high frequency noise signals with the use of FIR, IIR and WAVELETS, FILTER BANKS.
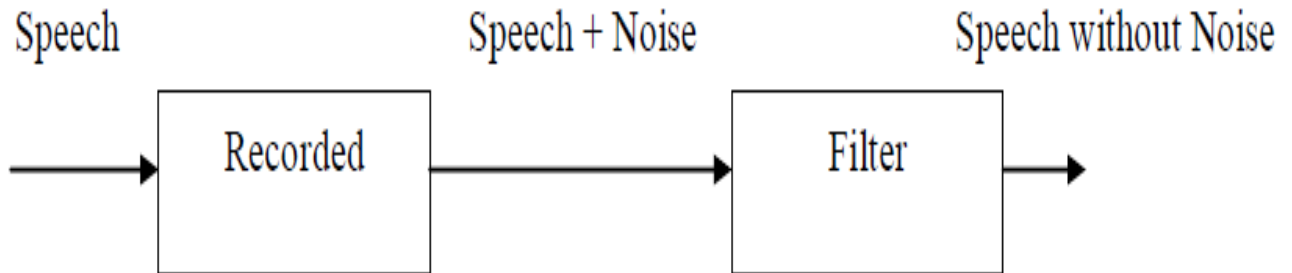
Figure 1 Block Diagram of Speech analysis

II. WAVELET FILTER

A wavelet is a term used to describe a short time series (typically less than 100 samples) which can be used to represent, for example, the source function. As previously shown, the wavelet can be studied as a time series in the time domain or in the frequency domain as an amplitude or phase spectrum. For any amplitude spectrum there are an infinite number of time domain wavelets which can be constructed by varying the phase spectrum. There are two special types of phase spectra of specific interest.

The **minimum phase** wavelet has a short time duration and a concentration of energy at the start of the wavelet. It is zero before time zero (causal). An ideal seismic source would be a spike (maximum amplitude at every frequency), but the best practical one would be minimum phase. It is quite common to convert a given wavelet source wavelet into it's minimum phase equivalent since several processing stages (e.g. predictive deconvolution) work best by assuming that the input data is minimum phase.

The **maximum phase** wavelet is the time reverse of the minimum phase and at every point the phase is greater for the maximum than the minimum. All other causal wavelets are strictly speaking **mixed-phase** and will be of longer time duration. The convolution of two minimum phase wavelets is minimum phase. The **zero-phase** wavelet is of shorter duration than the minimum phase equivalent. The wavelet is symmetrical with a maximum at time zero (non-causal). The fact that energy arrives before time zero is not physically realisable but the wavelet is useful for increased resolving power and ease of picking reflection events (peak or trough). The convolution of a zero-phase and minimum phase wavelet is mixed phase (because the phase spectrum of the original minimum phase wavelet is not the unique minimum phase spectrum for the new modified wavelet) and should be avoided.
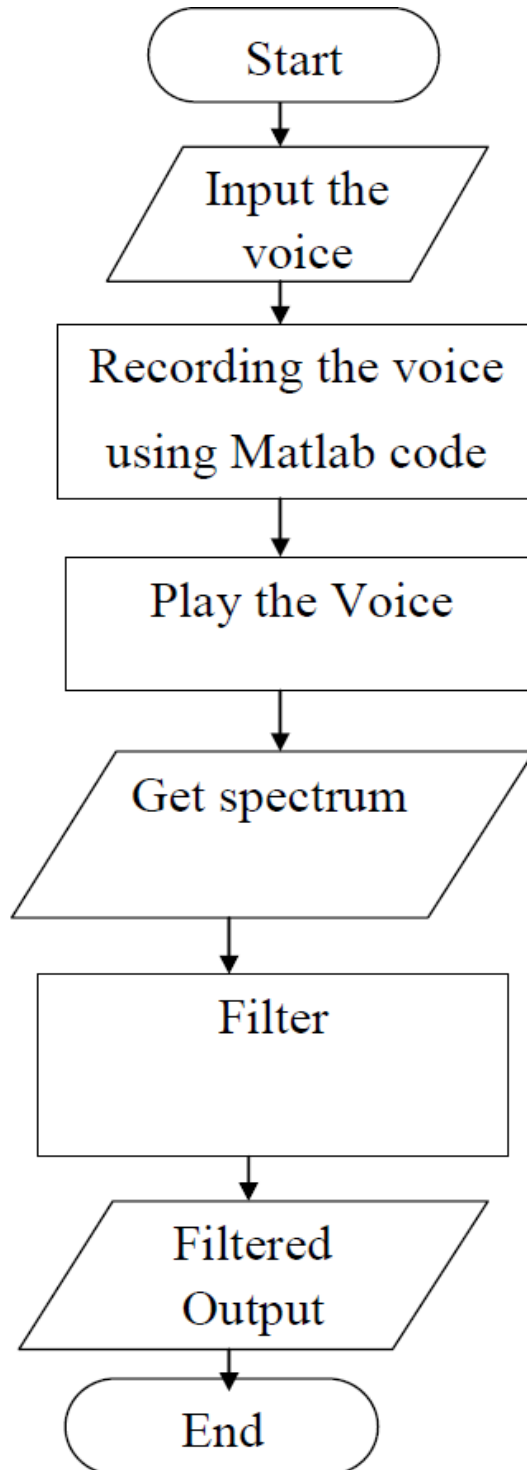
III.   FLOW CHART

Figure 2 Flow Chart of Procedure Followed

IV. CONCLUSION

The work that I have done involves filtering of recorded voice that includes noise .The future scope of speech recognition is implementation of different filters on recorded voice to remove noise.
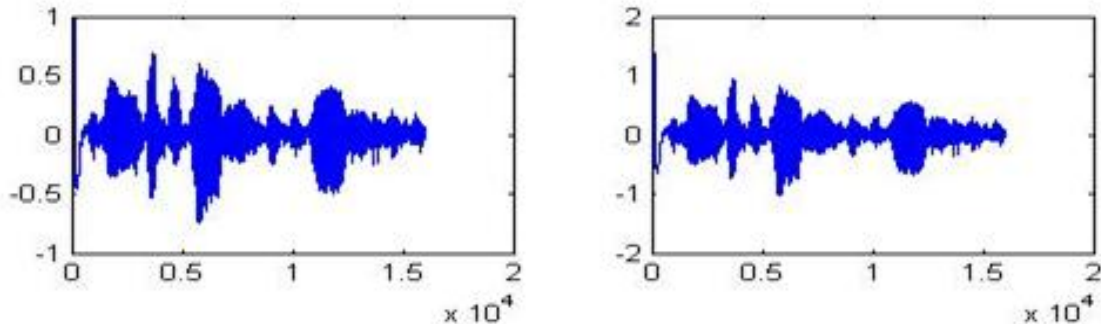

Figure 3 Recorded speech signal,Filtered speech signal

REFERENCES

[1] S. Ayat. Enhanced Human-Computer Speech Interface Using Wavelet Computing. *IEEE VECIMS*. 2008.

[2] S. G. Chang, B. Yu, M. Vetterli. Adaptive Wavelet thresholding for Image Denoising and Compression. *IEEE Trans On Image processing.* 2000, vol. 9, no. 9.

[3] D. L, Donoho, I.M. Johnstone. Ideal Spatial Adaptation via Wavelet Shrinkage. *Biometrik.* 1995, vol. 81, no.3, pp.425-455.

[4] D. L. Donoho. Denoising by Soft thresholding. *IEEE Trans on Information Theory*. 1995, vol. 41, no. 3, pp. 613-627.

[5] G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Processing, vol. 9, no. 9, pp. 1532–1546, 2000.*

[6] Diego Santa Cruz and Touradj Ebrahimi, "The JPEG 2000 image coding standard," *Dr. Dobb's Journal, vol. 26, no. 4, pp. 46–54, Apr. 2001..*

[7] M. Bahoura and J. Rouat, "Wavelet speech enhancement based on the teager energy operator," *IEEE Signal Processing Letters, vol. 8, no. 1, pp. 10–12, 2001.*

[8] S. Chen and J. Wang, "Speech enhancement using perceptual wavelet packet decomposition and teager energy operator," *Journal of VLSI Signal Processing, vol. 32, pp. 125–139, 2004.*