

# Exploration of Wavelets for Pre-processing of Speech Signals



Mohit Kumar Yadav, Vikrant Bhateja, and Monalisa Singh

**Abstract** Speech signal carries much more information than spoken words and speaker's information. Noise present in the speech signal is an obstacle in the path of analysing the speech signal precisely. Therefore, speech signal denoising has an important role in the speech signal analysis. In this paper, Wavelets are explored for pre-processing and suppression of noise from the noisy speech signals. The main aim of the proposed methodology involves, the pre-processing of the speech signals by using the combination of Discrete Wavelet Transform (DWT) and hard thresholding and to evaluate the result to find the optimal wavelet family with there optimal order followed by the suitable level of decomposition. This paper involves the exploration of the Wavelet Family and an improved Signal to Noise Ratio (SNR) values of the acquired speech signal which is shown in the result analysis.

**Keywords** Speech signals · Discrete wavelet transform (DWT) · Wavelet transform (WT) · Hard thresholding · Signal to noise ratio (SNR)

## 1 Introduction

Communication is the necessity of all human beings. Speech is one of the most important communication sources. It is the most natural, intuitive and fastest means of interaction among humans [1]. The speech signal contains the background noise and the distorted speech signal. The noise present in the speech signal causes the degradation in the quality of the speech. Over the last few decades, the pre-processing of speech signal which involves the removal of noise from speech has been an area of interest for the researchers [2]. Therefore, to remove a noise from the speech is very important as it may affect the further processing of the speech signal. So, filtering the noise from the acquired speech signal plays an important role in the speech signal analysis. As the frequency domain based signal processing can be designed easily

---

M. K. Yadav · V. Bhateja (✉) · M. Singh

Department of Electronics and Communication Engineering, Shri Ramswaroop Memorial Group of Professional Colleges (SRMGPC), Lucknow, Uttar Pradesh 226028, India

Dr. A.P.J. Abdul, Kalam Technical University, Lucknow, Uttar Pradesh, India

and most of the noise suppression methods require the use of Short Term Fourier Transform (STFT) but in today's scenario Wavelet Transform (WT) is gaining much importance as it is simple and efficient method for speech signal de-noising. Wavelet transform can analyze the signal and can select the information present in it that other signal de-noising techniques lack [2]. Therefore, lot of study has been done by the researchers for the suppression of noise in speech signals. Aggarwal et al. [2] introduced an approach of multi-resolution analysis using WT and found that the modified universal threshold gives better results of denoising. Bhowmick et al. [3] proposed a method of Voiced Speech Probability Detection Wavelet Decomposition (VSPDWD) and compared it with seven different techniques. It was found that the VSPDWD technique gave an improved SNR values at all levels of decomposition. Akkar et al. [4] in their study made a comparison between different thresholding techniques. It was found that the square wavelet thresholding gave better results than the traditional thresholding techniques. Babichev et al. [5] introduced a de-noising method that is the Emperical Mode and the Wavelet Decomposition techniques. It was found that there was a relative change in the values of the Shannon Entropy used for the quality criterion. This indicates that the technique used was effective. Hadi et al. [6] in their study made a comparison between different threshold selection rule. It was found that the sawtooth wavelet thresholding gave better results than the traditional thresholding techniques.

In this paper, a combination of Discrete Wavelet transform and Hard Thresholding technique for noise reduction has been proposed. Wavelet Transform provides a multi-resolution and is a better technique than Fourier Transform and STFT. This paper consists of following sections: Section 2 gives the description of the pre-processing of noisy speech signals using wavelets. Section 3 consists of experimental setup, Sect. 4 consists of exploration of suitable Wavelet Family for speech signal analysis and Sect. 5 consists of simulation results. Section 6 consists of the conclusion of the analysis done in the proposed work.

## 2 Pre-processing of Noisy Speech Signals Using Wavelets

### 2.1 Discrete Wavelet Transform

For suppressing the noise present in the speech signal, Discrete Wavelet Transform has been used. It involves the decomposition of the speech signal in the time frequency domain. The noise present in the speech signal cannot be easily removed by using the Kalman or Chebyshev filters. Therefore, it can be removed by applying the wavelet transform [7–9]. A wavelet is a wave like oscillation that begins at zero, increases and then again goes to zero [1, 10]. The scaled and shifted version of fundamental or mother wavelet  $\Psi$  is elucidated below [3]:

$$\Psi_{\tau,\beta}(t) = \beta^{-1/2}\Psi\left(\frac{t - \tau}{\beta}\right) \tag{1}$$

where  $\beta$  is the scaling parameter and  $\tau$  is the translation parameter. The noisy speech signal  $s(t)$  is decomposed into sub-bands through DWT into approximation and detailed coefficients. The detailed coefficient or the higher frequency component  $D(p, k)$  has been elucidated below [3]:

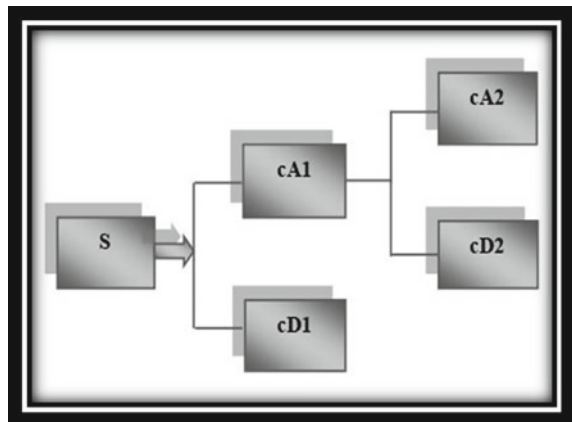
$$D(p, k) = 2^{-p/2} \sum_n s(n)\Psi^*(2^{-p}n - k) \tag{2}$$

where  $p, k$  and  $n$  are integers and  $\psi^*(t)$  is the complex conjugate of  $\psi(t)$ . The approximation coefficients or the lower frequency component  $A(p, k)$ , has been elucidated below [3]:

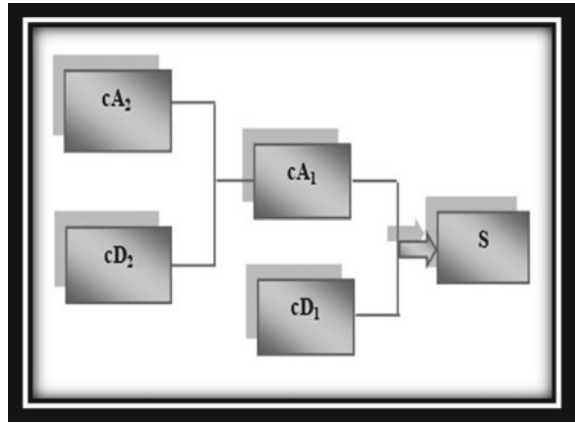
$$A(p, k) = 2^{-p/2} \sum_n s(n)\phi^*(2^{-p}n - k) \tag{3}$$

where  $\phi(n)$  is the complex conjugate of the scaling function  $\phi(n)$ . When DWT is applied to the noisy signal  $s(t)$  at different level then the speech signal decomposes to approximation and detailed coefficients [3, 11]. The detailed coefficients are obtained by filtering the high frequency component present in the noisy speech signals through high pass filter and the approximation coefficients are obtained by filtering the low frequency component present in the noisy speech signals through low pass filter. The reconstruction of the original speech signal is done by applying the Inverse Discrete Wavelet Transform (IDWT) to the filtered speech signal which is formed by combining the detailed and the approximation coefficients from the last level of decomposition to the first level. Figures 1 and 2, shows the two level wavelet decomposition and reconstruction process in which  $s$  is the noisy speech signal,  $cA1$  and  $cD1$  are the first level approximation and detailed coefficients and  $cA2$  and  $cD2$  are the second level approximation and detailed coefficients [2].

**Fig. 1** Wavelet decomposition [2]



**Fig. 2** Wavelet reconstruction [2]



**2.2 Universal Thresholding Based Filtering Method for Pre-processing of Speech Signals**

The noise present in the speech signal is a major issue in the speech signal analysis. In the proposed work, a universal thresholding based filtering technique using DWT has been proposed. The higher frequency components acquired through DWT is having a residual noise that cannot be removed by applying the simple filtering process [9, 12, 13].

**2.2.1 Threshold Selection**

The universal threshold value  $T$  can be evaluated by the equation elucidated below [3]:

$$T = \sigma \sqrt{2 \ln(L)} \tag{4}$$

where  $L$  is the noisy speech signal sample. The standard deviation  $\sigma$  can be evaluated as [5]:

$$\sigma = \frac{\text{MAD}(|D(n)|)}{0.6745} \tag{5}$$

where MAD is the Median Absolute Deviation and  $D[n]$  is the detailed coefficient of noisy speech signal.

### 2.2.2 Threshold Function

The universal threshold function, hard thresholding has been used in the proposed work. The calculation formula for hard ( $H_{m,n}$ ) threshold function is given below [1]:

$$H_{m,n} = \begin{cases} \omega_{m,n} & |\omega_{m,n}| \geq \mu \\ 0 & |\omega_{m,n}| < \mu \end{cases}, \quad (6)$$

where  $\omega_{m,n}$  is the wavelet decomposition coefficient of the noisy speech signal and  $\mu$  is the threshold value. The threshold value  $\mu$  is placed to zero and if the value of the coefficients is more than the threshold value then all the coefficients are threshold and this is known as Hard Thresholding.

## 3 Experimental Setup

The proposed method is performed and evaluated on NOIZEUS database [14]. The database accommodates 30 IEEE sentences contaminated with eight different noises at different SNRs. The noise is added to the sentences from the AURORA database [15] that includes train, babble, car, exhibition hall, restaurant, street, airport and train-station noise. In this experiment the noise from the noisy speech signal is removed using Discrete Wavelet Transform technique. In the proposed methodology the noisy speech signal is decomposed into Approximation and Detailed coefficients by using different types of wavelets like Daubechies, Coiflets, Symlets and Haar wavelet. The detail coefficients are difficult to remove through filters. Therefore, noise suppression in the noisy speech signals is done through hard thresholding [16–18]. The evaluation of the proposed work is done by calculating the SNR by using the mathematical expression as elucidated in Eq. (7) [3, 11]:

$$\text{SNR} = 10 \log 10 \left( \frac{\sum_{m=1}^L s^2(m)}{\sum_{m=1}^L |s(m) - \hat{s}(m)|^2} \right) \quad (7)$$

where  $L$  is the sample size for the filtered speech signal,  $s(m)$  is the noisy speech signal and  $\hat{s}(m)$  is the clean speech signal.

## 4 Exploration of Suitable Wavelet Family for Speech Signal Analysis

In the proposed work different Wavelet families has been explored for the suppression of noise from speech signals. From the different family used which are Daubechies,

**Table 1** Comparative analysis of SNR values for different wavelet families

Wavelet family	SNR (0 dB) Noisy signal = 2.7767 dB	SNR (5 dB) Noisy signal = 5.7318 dB	SNR (10 dB) Noisy signal = 10.0106 dB
Daubechies	3.7698	6.6212	11.4593
Symlets	3.7627	6.6698	11.3870
Coiflets	3.8136	6.6626	11.5248
Haar	3.0689	5.8379	9.2828

**Table 2** Comparative analysis of SNR values for different order of Coiflet wavelet

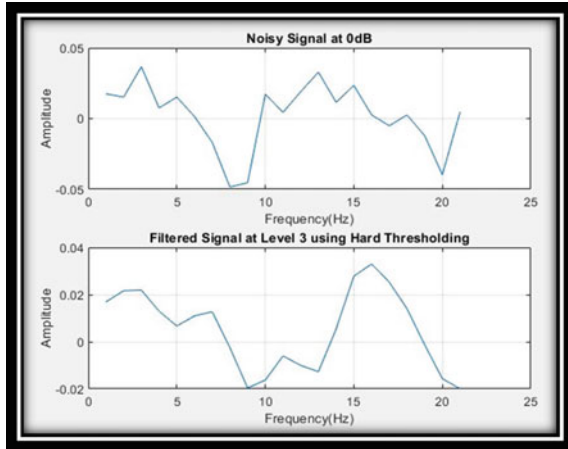
Order of Coiflet wavelet	SNR (0 dB) Noisy signal = 2.7767 dB	SNR (5 dB) Noisy signal = 5.7318 dB	SNR (10 dB) Noisy signal = 10.0106 dB
coif1	3.3330	6.1603	10.3155
coif2	3.5644	6.4282	10.9109
coif3	3.7176	6.4762	10.9817
coif4	3.7345	6.6274	11.3477
coif5	3.8136	6.6626	11.5248

Symlets, Coiflets and Haar wavelet, Coiflets tends to give the optimal SNR value. The comparative analysis of SNR values for different wavelet family is shown in Table 1. The noise in the signal decreases from 0 dB to 10 dB and the SNR value of the respective signal tends to increase. The order 5 of the Coiflet wavelet gives the optimal SNR value. The comparative analysis of SNR values for different order of Coiflet Wavelet are shown in Table 2. Here it is examined that as the order of the Wavelet family increases, the SNR value increases. The comparative analysis of SNR values for different level of decomposition of noisy speech signal is shown in Table 3. Here it is analyzed that the SNR improves to a certain level of decomposition and then it stops as the sample number decreases in lower sub-bands. So, based on the explored parameters the simulation of result has been done.

**Table 3** Comparative analysis of SNR values for different level of decomposition

Level of decomposition	SNR (0 dB) Noisy signal = 2.7767 dB	SNR (5 dB) Noisy signal = 5.7318 dB	SNR (10 dB) Noisy signal = 10.0106 dB
Level 1	3.6323	6.4968	11.4626
Level 2	3.7274	6.5847	11.5248
Level 3	3.8136	6.6626	11.3051
Level 4	3.7429	6.6096	10.8951
Level 5	3.7326	6.5798	10.8382

**Fig. 3** Noisy and the filtered speech signal. SNR (noisy signal) = 2.7767 dB. SNR (filtered signal) = 3.7698 dB



### 5 Simulation Results

The speech signals contain the noise which is important to remove as it causes the difficulty in the further processing of the signal. Figure 3 shows the noisy and filtered speech signal. The decomposition of the noisy speech signal is done through DWT at various levels along with a different wavelet family. The hard thresholding is applied by calculating the threshold value  $T$  to the coefficients obtained through the decomposition of the noisy speech signal [19, 20]. From the exploration of suitable wavelet family the result is simulated. Therefore, an improved result of pre-processing of speech signal by using the combination of DWT and Hard Thresholding has been obtained. The reconstruction of the noisy speech signal is done through IDWT. The important information present in the reconstructed speech signal is not lost [21, 22].

### 6 Conclusion

In the proposed work, the pre-processing of the noisy speech signal through the combination of DWT and hard thresholding has been done and the Wavelet Family has been explored to obtain the improved result. Comparative analysis for the best wavelet family, order of the wavelet and the best level of decomposition has been obtained and based on the explored parameters the simulation of the result for the suppression of noisy speech signal has been done.

## References

1. Zhong, X., Dai, Y., Dai, Y., Jin, T.: Study on processing of wavelet speech denoising in speech recognition system. *Int. J. Speech Technol.* **21**, 563–569 (2018)
2. Aggarwal, R., Singh, J.K., Gupta, V.K., Rathore, S., Tiwari, M., Khare, A.: Noise reduction of speech signal using wavelet transform with modified universal threshold. *Int. J. Comput. Appl.* **20**, 14–19 (2011)
3. Bhowmick, A., Chandra, M.: Speech enhancement using voiced speech probability based wavelet decomposition. *Comput. Electr. Eng.* **62**, 706–718 (2017)
4. Akkar, H.A.R., Hadi, W.A.H., Al-Dosari, I.H.: A squared-Chebyshev wavelet thresholding based 1D signal compression. *Defence Technol.* **15**, 426–431 (2019)
5. Babichev, S., Mikhalyov, O.: A hybrid model of 1-D signal adaptive filter based on the complex use of Huang transform and wavelet analysis. *Int. J. Intell. Syst. Appl.* 1–8 (2019)
6. Akkar, H.A.R., Hadi, W.A.H., Al-Dosari, I.H.: Implementation of sawtooth wavelet thresholding for noise cancellation in one dimensional signal. *Int. J. Nanoelectron. Mater.* **12**, 67–74 (2019)
7. Bhateja, V., Urooj, S., Verma, R., Mehrotra, R.: A novel approach for suppression of powerline interference and impulse noise in ECG signals. In: *Proc. of IMPACT-2013*. pp. 103–107. Aligarh, India (2013)
8. Chieng, T.M., Hau, Y.W., Bin Omar, Z., Lim, C.W.: Qualitative and quantitative performance comparison of ECG noise reduction and signal enhancement method based on various digital filter designs and discrete wavelet transform. *Int. J. Comput. Digital Syst.* **9**, 534–544 (2020)
9. Bhateja, V., Srivastava, A., Tiwari, D.K.: An approach for the preprocessing of EMG signals using canonical correlation analysis. *Smart Comput. Inform.* **78**, 201–208 (2017)
10. Jakati, J.S., Shridhar, S.K.: Efficient speech de-noising algorithm using multi-level discrete wavelet transform and thresholding. *Int. J. Emerg. Trends Eng. Res.* **8**, 2472–2480 (2020)
11. Taqee, A., Bhateja, V., Shankar, A., Srivastava, A.: Combination of wavelets and hard thresholding for analysis of cough signals. In: *2018 Second World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*. pp. 266–270. IEEE Press, London (2018)
12. Wang, K., Su, G., Liu, L., Wang, S.: Wavelet packet analysis for speaker-independent emotion recognition. *Neurocomputing* **398**, 257–264 (2020)
13. Venkateswarlu, S.C., Karthik, A., Kumar, D.N.: Performance on speech enhancement objective quality measures using hybrid wavelet thresholding. *Int. J. Eng. Adv. Technol. (IJEAT)* **8**, 3523–3533 (2019)
14. NOIZEUS: A Noisy Speech Corpus for Evaluation of Speech Enhancement Algorithms. <https://ecs.utdallas.edu/loizou/speech/noizeus/>
15. The AURORA Experimental Framework For The Performance Evaluation of Speech Recognition Systems Under Noisy Conditions. [https://www.isca-speech.org/archive\\_open/asr2000/asr0\\_181.html](https://www.isca-speech.org/archive_open/asr2000/asr0_181.html)
16. Verma, R., Mehrotra, R., Bhateja, V.: A new morphological filtering algorithm for preprocessing of electrocardiographic signals. In: *Proc. (SPRINGER) of the Fourth International Conference on Signal and Image Processing (ICSIP 2012)*. pp. 193–201. Coimbatore, India (2012)
17. Bhateja, V., Devi, S.: Combination of wavelet analysis and non-linear enhancement function for computer aided detection of microcalcifications. In: *Proc. of International Conference on Biomedical Engineering and Assistive Technologies (BEATs-2010)*, p. 44. Jalandhar, India (2010)
18. Alabbasi, H.A., Jalil, A.M., Hasan, F.S.: Adaptive wavelet thresholding with robust hybrid features for text-independent speaker identification system. *Int. J. Electr. Comput. Eng.* **10**, 2088–8708 (2020)
19. Mishra, A., Bhateja, V., Gupta, A., Mishra, A.: Noise removal in EEG signals using SWT–ICA combinational approach. *Smart Intell. Comput. Appl.* **105**, 217–224 (2018)
20. Srivastava, A., Bhateja, V., Shankar, A., Taqee, A.: On analysis of suitable wavelet family for processing of cough signals. *Front. Intell. Comput. Theory Appl.* **1014**, 194–200 (2019)



21. Bhateja, V., Taquee, A., Sharma D.K.: Pre-processing and classification of cough sounds in noisy environment using SVM. In: Proc. of 2019 4th International Conference on Information Systems and Computer Networks (ISCON), pp. 822–826. Mathura, India (2019)
22. Vani, H.Y., Anusuya, M.A.: Improving speech recognition using bionic wavelet features. AIMS Electron. Electr. Eng. **4**, 200–215 (2020)