Denoising technique for speech signal using orthogonal wavelets

Ala chenchaiah[1], G.Umamaheswara Reddy[2].
[1]*M.Tech, Sri Venkateswara University College of Engineering, Tirupati-517502.*
[2]*Professor, Sri Venkateswara University College of Engineering, Tirupati-517502.*
Email:[1]chenchaiah.ala@gmail.com, [2] umaskit@gmail.com

Abstract—This paper introduces an algorithm for denoising the noisy speech signal for the better quality of perceptual information. Due to the limitation of window size and spectral leakages in voice not getting the optimum results with the Short Time Fourier Transform. For avoiding this problem denoise the speech signal by using wavelet transform. The Symlet wavelet and the Daubechies wavelet are suitable for Speech Signal Processing. In this paper, the family of Symlet wavelet of order Sym1, Sym2, Sym3 and Sym4 and the family of Daubechies wavelet of order Db1, Db2, Db3 and Db4 have been used to denoise the noisy Speech signal. The test speech quality measure PESQ (perceptual speech quality parameter) and PSNR values had shown that the Symlet wavelet of order 4 with the hard threshold technique can obtain better performance.

Keywords— *PESQ, Daubechies wavelet, Symlet wavelet, Short time Fourier transform, detailed coefficients.*

I. INTRODUCTION

Speech processing has many applications in speech recognition systems, teleconferencing systems, mobile communication systems, speech intelligibility systems and etc. During the speech signal propagation, it may affect with different noises like babble noise, train noise, white noise and etc[1]. Babble noise and train noise are non-stationary which means slow changes will happen in the signal. White noise is a type of noise which contains all frequencies with constant amplitude. Removing the noise from the noisy speech signal is termed as denoising the signal [2]. Short Time Fourier Transform (STFT) is one of the techniques for removing the noise in noisy speech signal, but there are two main problems with STFT that can limit the performance [3]. (i) Selecting an optimal window length of data segments that contain several features may not possible. (ii) Shorting the data segment could also result in the loss of low frequencies that are no longer fully included in the data segment. Hence, if the window is made smaller to improve the time resolution, then the frequency resolution is degraded and vice versa [4]. To overcome these problems wavelet denoising threshold techniques have been used.

II. METHODOLOGY

Depending upon particular application has been chosen different wavelets. For Speech signal propagation the Symlet wavelets and Daubechies wavelets have been used. The Sym wavelet family has sym1, sym2, sym3, and sym4 and the Daubechies wavelet family has Db1, Db2, Db3, and Db4. In this work, the performance of the Symlet wavelet is compared with the Daubechies wavelet. For most signal processing applications, discrete wavelet transforms (DWT) based analysis is best described in terms of filter banks [5]. In the proposed method in order to remove the noise from noisy speech signal Symlet wavelet and Daubechies wavelet have been used in filter banks independently [6]. The filter bank that decomposes the original signal is usually termed as analysis filters while the filter bank that reconstructs the signal is called synthesis filters [7].The speech signal affected by different noises like babble noise, train noise, and white noise is given to Symlet wavelet and Daubechies wavelet independently. By using the group of filters to divide up a signal into various spectral components is termed as decomposition or sideband coding.

The rule which is used to set a threshold value is known as threshold selection rule. In this work the thresholding selection rules are (i) Sure shrink (ii) minimax (iii) Haur sure. The obtained threshold value is compared with the wavelet coefficients using a threshold technique. In this paper two types of thresholding techniques (i) Soft thresholding (ii) Hard thresholding have been used [8]. In both soft threshold technique and Hard threshold technique, the wavelet coefficients which are less in magnitude than the threshold tend to zero. In the soft threshold method, the coefficients which are higher in magnitude than the threshold are compressed towards zero by subtracting the threshold value with coefficient value. In hard threshold, the coefficients which are higher than the threshold are unchanged. In this paper hard threshold with sure shrink gives better results.

In the proposed method the number of levels is four. In each stage, the output of the Symlet wavelet is down-sampled and passed to the next level and so on. A similar process will be performed for Daubechies wavelet also. At each section of the analysis filter got output as detailed coefficients and have known that the proposed method has four levels, got four detailed coefficients.

As the number of levels increases, it is complicated to get back the original signal due to redundancy increases in the samples. The last stage output of analysis filter is given to the synthesis filter. The output of the Symlet wavelet in the synthesis filter is up-sampled at each level and is given to the next stage and so on, finally get back the estimated Speech signal $\hat{s}(n)$ by either soft thresholding or hard thresholding using the threshold selection rule (the sure shrink or the minimax). This process continued for Daubechies wavelets also.

III. PROCEDURE

The algorithm part contains (i) The noisy speech signal (ii) Decomposition of the signal using wavelet transform (iii) Thresholding and (iv) Reconstructing the Speech signal. The block diagram for the denoising of the noisy speech signal is as shown in the Fig (1) below.

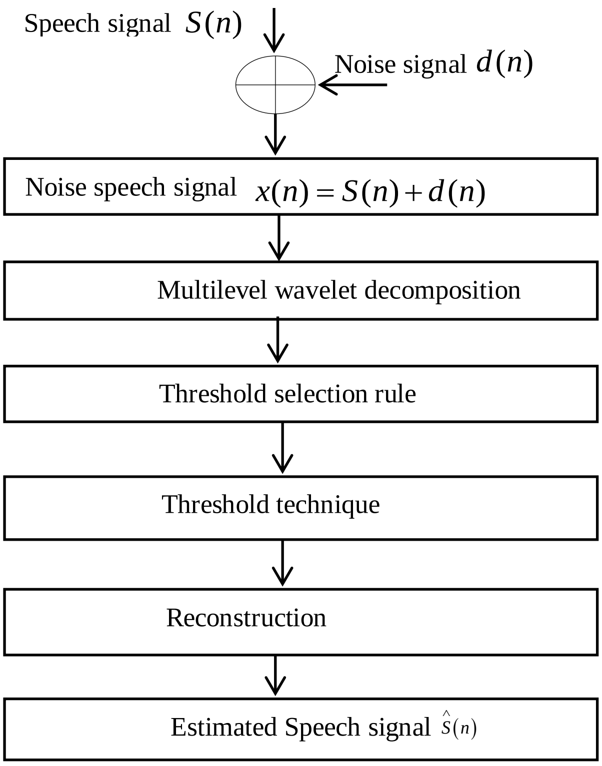

Fig (1): The Block diagram for Denoise of noisy speech signal

The noise removing has been carried out on the core test set of the NOIZEUS speech corpus which was composed of 30 phonetically-balanced sentences belonging to speakers (three males and three females). There were 8 kinds of non-stationary noises at different SNRs except noisy speech with white noises, during an evaluation process, we have generated a noisy speech set corrupted by white noise, babble noise, train noise taken from the NOISEX-92 noise database at four SNR levels:0dB,5dB,10dB, and 15dB. Then the corrupted files were denoised by the proposed method. For evaluation purpose, we employed an objective speech evaluation measures that are the Perceptual Estimation of Speech Quality (PESQ) and PSNR. PESQ is a denoised perceptual quality measurement for voice quality in telecommunications [9]. Its score is between 1.0 and 4.5, where 1.0 corresponds to bad and 4.5 corresponds to distortion less. The parameters PESQ and PSNR obtained from the original speech signal and estimated speech signal [8]. In our evaluations, we computed the mean PESQ scores for each method and each noise case. In our experiment, we had family of symlet wavelet sym1, sym2, sym3 and sym4 and Daubecheies wavelets Db1, Db2, Db3, and Db4 [10]. We got good results for symlet wavelet of order 4 with hard thresholding using sure shrink threshold selection rule when compared with the existing method that is STFT in terms of PESQ and PSNR parameters. The output waveforms for the STFT are shown below Fig (2)

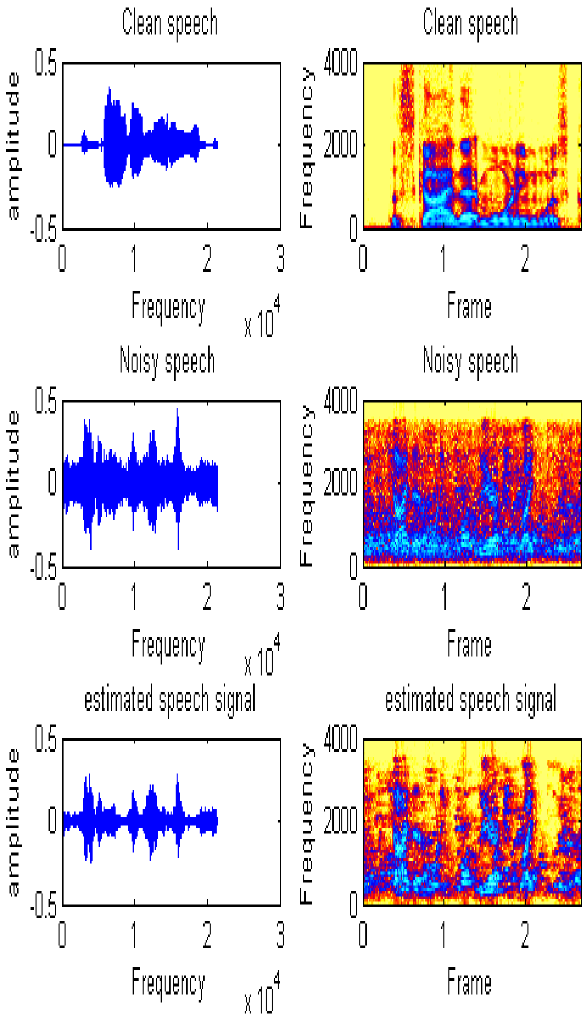

Fig (2): Output waveforms for the existing method.

The output waveforms for the Symlet wavelet of order 4 with hard thresholding using sure shrink threshold selection rule have given as shown in Fig 3(a), 3(b), 3(c), 3(d) below

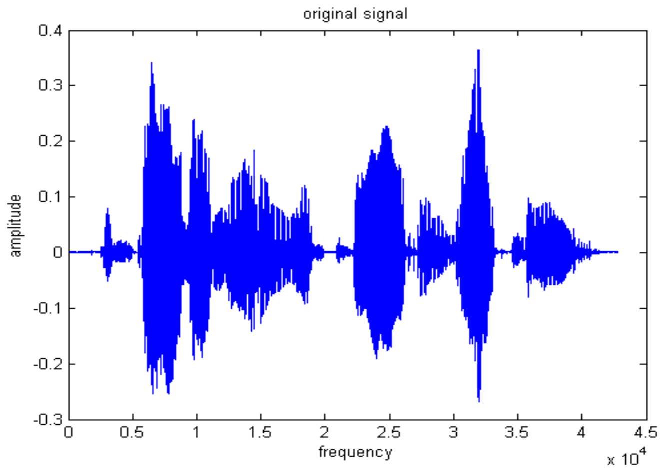

Fig 3 (a): Speech signal $S(n)$

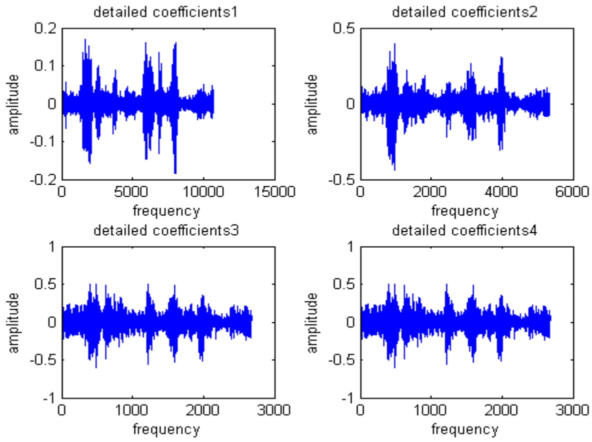

Fig3(c): detailed coefficients for level 1, level 2, level 3 and level 4

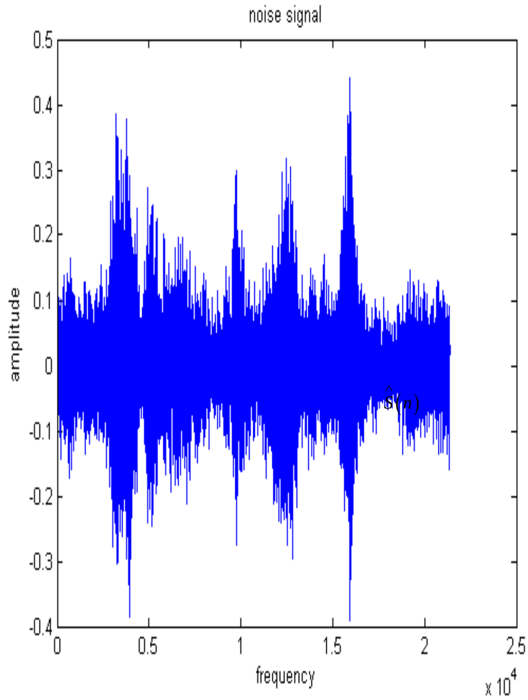

Fig 3(b): noise signal $d(n)$

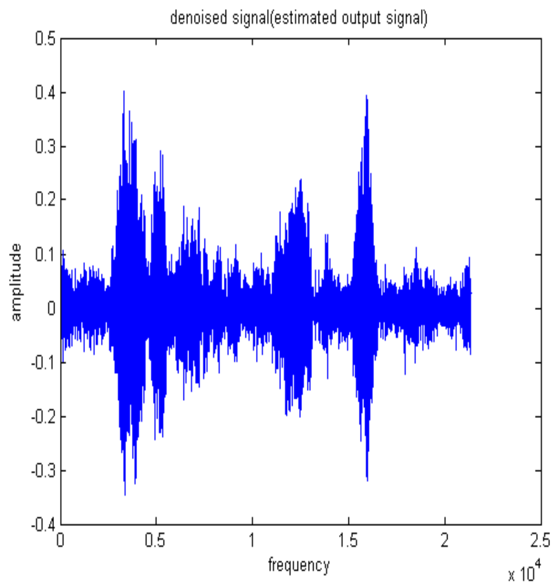

Fig 3(d): Estimated output signal $\hat{S}(n)$

IV. RESULTS

In this paper the threshold selection rules sure shrink, minimax and Haur sure are applied to the sp10.wave speech signal with different noises. Each threshold selection rule is applied to the family of Daubechies wavelet and the family of Symlet wavelet using soft thresholding technique and hard threshold technique independently. Table 1 had shown that sure shrink got good PESQ scores with Symlet wavelet of order 4 using hard threshold technique than other threshold selection rules and the existing method.

TABLE 1: Comparison of PESQ parameter values of Sym 4 wavelet using Hard thresholding technique

Type of noise	Existing method	Proposed method		
		Sym4+hard	*Sym4+hard*	*Sym4+hard*
	STFT	*minimax*	*Haur sure*	*Sure shrink*
	PESQ	*PESQ*	*PESQ*	*PESQ*
Babble _0db	**1.5**	2.5	1.8	**2.8**
Babble _5db	**1.0**	2.3	1.95	**2.4**
Babble _10db	**1.2**	2.1	1.4	**2.2**
Babble _15db	**1.6**	2.4	1.6	**2.8**
Train_0db	**1.3**	1.9	1.7	**2.5**
Train_5db	**1.4**	2.5	1.86	**2.7**
Train_10db	**1.0**	2.3	1.9	**2.6**
White noise_10db	**2.1**	2.5	2.1	**2.8**

The graph for the Table1 has been drawn, PESQ along the Y-axis and different noises in db along X-axis, shown in below figure (4). From the graph, the sure shrink has better PESQ scores across all noises.

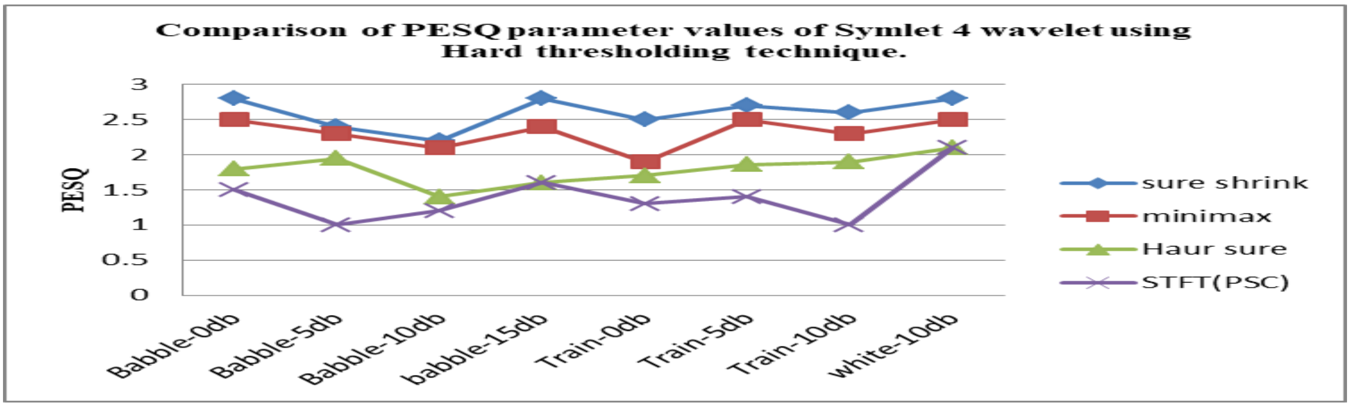

Fig (4): comparison of PESQ parameter values of Sym 4 wavelet.

The results of the Denoised Speech signal, in terms PESQ scores and PSNR, are shown in the Table 2 and Table3. The Table 2 and The Table 3 show that compared to the previous method, sym4 wavelet with a hard thresholding technique using sure shrink selection rule, performed well, it would obtain consistent improvements across all PESQs and PSNR values. In Table 2 the PESQ values of the symlet wavelet of order 4 with hard thresholding using sure shrink threshold selection rule got better results than the family of Daubechies wavelets. The PESQ values of sym 4 wavelet compared to sym1, sym 2 and sym 3 got better results. In Table 3 the PSNR values in db of the symlet wavelet of order 4 have better results than the family of Daubechies wavelets. The graphs for Table 2 and Table 3 have been drawn.

Table 2: Comparison of PESQ parameter values using sure shrink threshold selection rule with the existing method

Type of noise	Existing method	Proposed method						
	STFT	*Db2+ soft*	*Db1+ hard*	*Db4+ hard*	*Sym1+ hard*	*Sym2+ hard*	*Sym3+ hard*	*Sym4+hard*
	PESQ	*PESQ*	*PESQ*	*PESQ*	*PESQ*	*PESQ*	*PESQ*	*PESQ*
Babble _0db	**1.5**	2.4	2.5	1.9	2.23	2.4	2.5	**2.8**
Babble _5db	**1.0**	2.0	2.3	2.0	2.0	2.1	2.3	**2.4**
Babble _10db	**1.2**	1.8	2.0	1.4	1.5	1.7	1.8	**2.2**
Babble _15db	**1.6**	2.2	2.5	2.0	1.8	2.0	2.0	**2.8**
Train_0db	**1.3**	1.7	2.4	1.6	1.7	1.9	2.1	**2.5**
Train_5db	**1.4**	1.8	2.5	1.7	2.4	2.5	2.4	**2.7**
Train_10db	**1.0**	2.0	2.0	1.8	2.1	2.2	2.3	**2.6**
White noise_10db	**2.1**	2.6	2.7	2.4	2.3	2.5	2.6	**2.8**

.

The graph for the Table 2 has been drawn, PESQ along the Y-axis and different noises in db along X-axis, shown in below figure (5). From the graph, the sym 4 with a hard thresholding technique using sure shrink threshold selection rule has better PESQ scores across all noises.

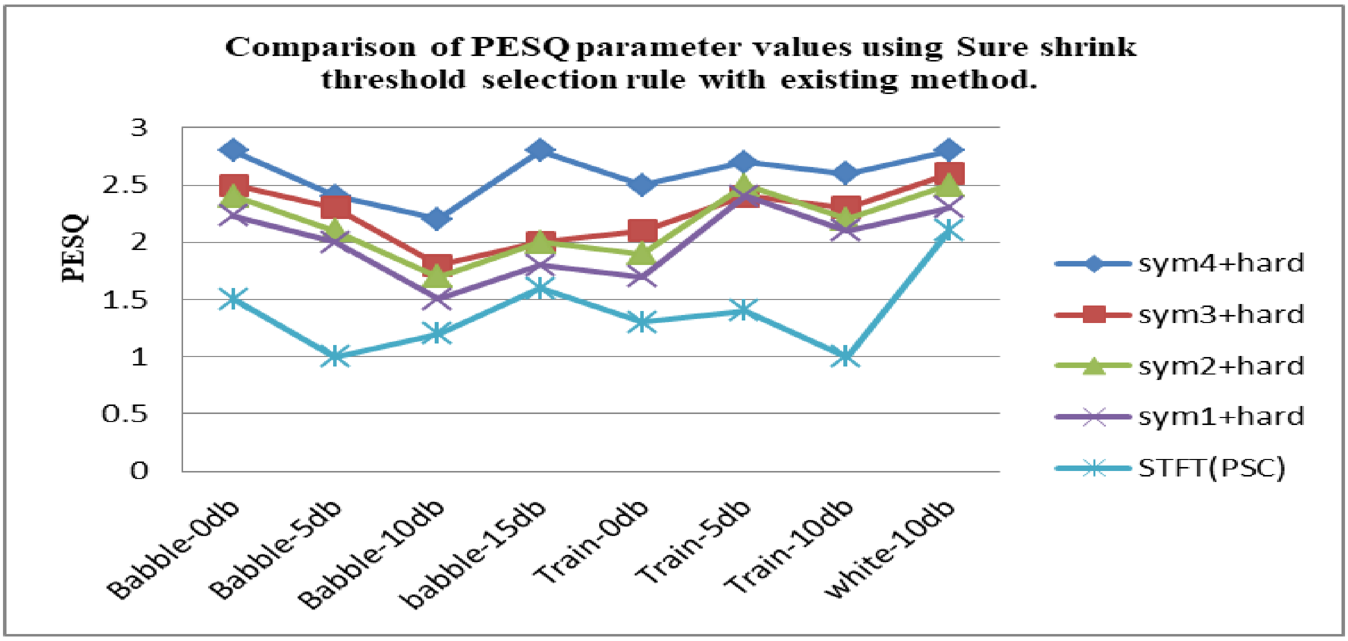

Fig (5): comparison of PESQ parameter values using sure shrink threshold selection rule.

TABLE 3: Comparison of PSNR values of the symlet wavelet of order 4 using sure shrink threshold selection rule

Type of noise	Existing method	Proposed method						
	STFT	*Db2+ soft*	*Db1+ hard*	*Db4+ hard*	*Sym1+ hard*	*Sym2+ hard*	*Sym3+ hard*	*Sym4+hard*
	PSNR	*PSNR*	*PSNR*	*PSNR*	*PSNR*	*PSNR*	*PSNR*	*PSNR*
Babble _0db	**22**	72	72	70	72	72	73	**75**
Babble _5db	**24**	73	73	72	73	73	72	**74**
Babble _10db	**26**	74	74	73	70	72	75	**77**
Babble _15db	**28**	75	74	72	72.9	73	74	**78**
Train_0db	**30**	74	75	74	72	73	73.4	**75**
Train_5db	**40**	74	74	71	71	73	74	**74**
Train_10db	**50**	73	76	72	72	70	75	**76**
White noise_10db	**60**	88	87	80	86	87	88	**89**

The graph for the Table 3 has been drawn, PSNR in db along the Y-axis and different noises in db along X-axis, shown in figure (6). From the graph the sym 4 with a hard thresholding technique using sure shrink threshold selection rule has better PESQ scores across all noises.

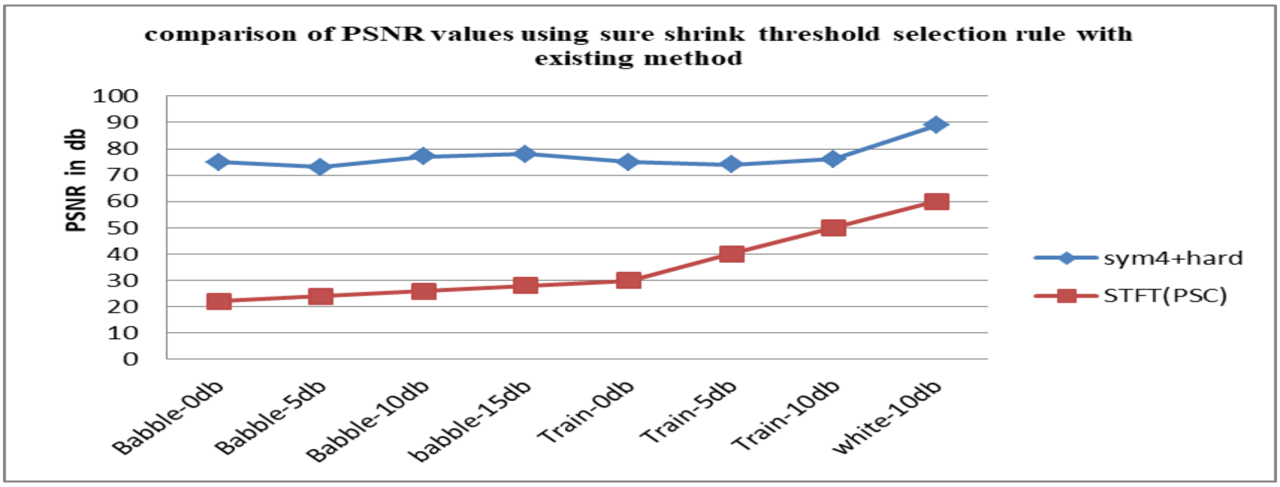

Fig (6): comparison of PSNR values with existing method STFT

Fig (5) and Fig (6) shows a single utterance sp10.wave corrupted by babble noise, train noise, white noise at 0dB, 5dB, 10dB, 15dB. As we can see, our proposed method symlet wavelet of order 4 with the hard threshold technique using sure shrink threshold selection rule had shown better performance, and exhibited better noise cancellation than the existing speech technique. Further subjective tests are needed to verify effectiveness of the algorithm on improving both subjective quality and speech intelligibility. If it works better, it is worth mentioning that our method, symlet wavelet of order 4 with Hard threshold, can be used in the systems which may want speech as clean as possible with very less distortion.

V. CONCLUSION

In this paper, we have proposed speech denoising method using sure shrink based symlet wavelet. Experimental results of the PSNR and objective speech quality measure PESQ had shown that the proposed method, symlet wavelet of order 4 with a hard thresholding technique, achieve better speech quality than the conventional speech denoising technique. The technique can be used in the systems which need to cancel the background noises such as speech recognition, speech communication, etc., and it can further improve the speech quality and intelligibility.

REFERENCES

[1]. Zhen Li, Wenjin Wu,Qin Zhang, Hui Ren, Shilei Bai, “Speech enhancement using magnitude and phase spectrum compensation,” IEEE/ACIS 15th International conference on computer and information science (ICIS)., PP.1-4 ,,jun. 2016.

[2]. P.Loizou, speech enhancement: theory and practice.Boca Raton, FL: CRC, 2007.

[3]. BOLL S F.Suppression of acoustic noise in speech using Spectral subtraction[J]. IEEE TRrans.Acoustics ,speech signal processing , 1979,27(2):113-120.

[4]. Ephraim Y,Malah D. Speech enhancement using a minimum mean square error short time spectral amplitude estimator. IEEE Transactions on Acoustic, speech, signal processing, 1984, 32(6):1109-1121.

[5]. N.Wiener, The Extrapolation, Ineterpolation, and Smoothing of Stationary Time Series with Engineering Applications.New York: Wiley, 1949.

[6]. Biosignal and Biomedical Image Processing MATLAB-Based Applications: JOHN L.SEMMLOW.

[7]. P.C.Hansen and S.H.Jensen,”FIR filter representation of Reduced rank noise reduction,”IEEE Tran.Signal Process., vol.46, no.6, pp.1737—1741, Jun.1998.

[8]. S Grace chang, Bin Yu and Martin vetterli,”Adaptive wavelet thresholding for image denoising and compression”, IEEE transactions on image Processing 9(9), 1532-1546, 2000.

[9]. Y.Ephraim and H.V.Trees,”A Signal subspace approach for speech enhancement,”IEEE Trans.Speech Audio Process., vol.3, no.4, pp.251-266, Jul.1995.

[10]. Hojjat Adeli, Ziqin Zhou and Nahid Dadehr,”Analysis of EEG records in an epileptic patient using wavlet Trans-Form”, Journal of neuroscience methods 123(1), 69-87, 2003.