Journal of Engineering and Applied Sciences 13 (Special Issue 14): 10923-10931, 2018
ISSN: 1816-949X

A New Speech Denoising Technique Based on Wavelet Thresholding and Hybrid Algorithm

Haider J. Abd, Hadeel Abdallah and Ali Shabban
Department of Electrical Engineering, College of Engineering, University of Babylon, Hillah, Iraq

Abstract: Speech signals play an important role in digital signal processing. When this signal passes through the medium, it interacts with the noise, so, noise must be removed without affecting the original signal. Denoising methods are a compromise between removing as much noise as possible and maintaining signal integrity. In this work, a Hybrid Bacterial Foraging Particle Swarm Optimization Model (HBFPSO) was proposed to estimate the threshold value without any prior information on signal and noise distribution to measure kurtosis function of remaining noise to locate optimal value of threshold when kurtosis value is maximized. It is noted that the suggested denoising technique showed an excellent performance over single models (PSO and BFO) at the same conditions. Furthermore, denoising results showed that the proposed HBFPSO algorithm provided the least MSE which resulted in an improvement of (0.004) compared with (BFO, PSO) algorithms.

Key words: Speech signal de-noising, PSO, BFO, HBFPSO, kurtosis, MSE

INTRODUCTION

The denoising of speech signal plays a significant role in speech enhancement applications. The main denoising aim is to eliminate the background noise from the corrupted signal (Liu *et al*., 2017). Speech signal corrupted with many types of noise such as white noise, pink noise, babble noise which are a major factor affecting the accuracy of the results in speech processing. Therefore, it is necessary to remove the unwanted noise signal (Arora *et al*., 2017; Arora and Oktar, 2017).

In the speech signal, noise is usually concentrated in the signal part which is represented by high-frequency components. Thus, the need to convert Wavelets (WT) has become very necessary for the purpose of signal analysis to its various frequency components and thresholding its to remove the noise (Shaban *et al*., 2017).

Thresholding principle should be elected accurately in denoising technique as it is still the key challenge. If threshold value is too large, this lead to filter out the significant signal features, producing deviation. If too small, the denoised signal keeps the noisy data (Wu, 2014). For the success of wavelet-based denoising, the value of the threshold estimations is very vital (Sumithra and Thanushkodi, 2009).

A few denoising ways are studied to improve speech signal by decreasing the corrupted noise. Wang and Shi (2009), Chen *et al*, (2015) used an adaptive denoising method depend on adaptive LMS to decrease the relationship between error and vector differences when the voice signal is influential and showed that the extra MSE can increase around (20 decibel) in various noise environments under the same condition. Santosh *et al*. (2012) introduced various adaptive filters to remove noise from speech signals such as RLS, LMS, Block LMS, NLMS and adaptive filters. The results showed that the NLMS filter offered high quality in CPU time consuming and PSNR of the output Block LMS was more convergence from others. Chen *et al*, (2015) used a method of adaptation threshold denoising depend on wavelet entropy. It combined the adaptive threshold with wavelet entropy to calculate the threshold value for high frequency coefficients combined the adaptive threshold with wavelet entropy to calculate the threshold value for high frequency coefficients. Cengiz and Aroiz (2016) used the Discrete Wavelet Transform (DWT) for denoising speech signal by hard and soft thresholding. Speech signal was corrupted by AWGN noise. The denoising results approved that the DWT is a suitable method for speech signal denoising when the parameters are elected correctly. Munegowda (2016) applied the wavelet transform technique for audio signal denoising from realistic noise, DWT was characterized to denoise of the one-dimensional signal from group of realistic noise. Shaban *et al*. (2017) employed Invasive Weed Optimization (IWO) of speech denoising method. They proved that the results of the IWO algorithm are better than PSO for one and five DWT levels.

Corresponding Author: Haider J. Abd, Department of Electrical Engineering, College of Engineering, University of Babylon, Hillah, Iraq

In this research, Hybrid Bacterial Foraging Particle Swarm Optimization algorithm (HBFPSO) is proposed to optimize thesolution by estimating the optimal threshold. It relies on kurtosis criteria determining of the estimated signal from detail coefficients by the inverse soft threshold. The hybrid algorithm assumes that there is a single threshold value which is named optimal threshold. The proposed technique offered good performance compared with BFO and PSO optimized algorithms.

MATERIALS AND METHODS

Effect of discrete wavelet transform on speech signal: DWT decomposes speech signal into two components which are Approximate coefficients (cA) and Detailed coefficients (cD) of lower and higher frequencies, respectively. The signals decomposition is obtained by passing the time domain through low pass and high pass filters (Arora and Bansal, 2012; Saikia *et al*., 2015). AWGN is the most common noise model which can be expressed in Eq. 1 (Wu, 2014).

$$\text{nSig} = \text{Sig}+\text{Noise} \tag{1}$$

where, nSig represents a noised signal, sSig is an unknown original signal, noise indicates to AWGN with zero mean and finite variance σ^2. The major aim of denoising technique is to obtain an estimated (eSig) with lower MSE as seen in Eq. 2 (Kumari and Devarakonda, 2013).

$$\text{MSE} = \frac{1}{N}\sum_{i}^{N}(\text{Sig}_i\text{-eSig}_i)^2 \tag{2}$$

In the practical engineering applications, the suitable signal usually can represent by low frequency part which shows its features while the noise signal states in high-frequency signal component, thus allowing us to decompose the noised signal with wavelet bases (Wu, 2014).

Different denoising techniques were proposed and the most efficient one was a wavelet transform method which is a famous tool for signal analysis into different frequency components as shown in Fig. 1.

To perform that, compared signal with wavelet basis function, then search their resemblances in frequency (Gao and Yan, 2010). The wavelet coefficients are explained in Eq. 3:

$$\begin{aligned} W\times \text{nSig} &= Wx_{\text{Sig}}+W\times \text{Noise} \\ W_{\text{nSig}} &= W_{\text{Sig}}+W_{\text{Noise}} \end{aligned} \tag{3}$$

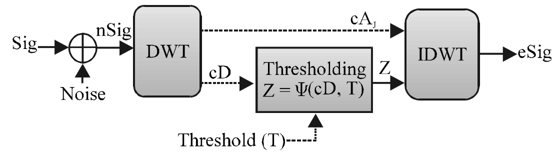

Fig. 1: Traditional denoising technique

where, W denotes to orthogonal DWT, W_{noise} is the noise (Chang *et al*., 2000). The Approximate coefficient (cA) and soft threshold (Z) reconstruct the estimated signal by taking the inverse discrete wavelet transform as in Eq. 4:

$$\text{eSig} = W^{-1}\times[cA_J, Z] \tag{4}$$

where, W^{-1} is the IDWT operator (Thyagarajan, 2005). The main threshold functions can be divided into two types:

Soft threshold function: It takes Detail coefficients (cD) and shrinks these coefficients toward 0 by Threshold T which is expressed in Eq. 5.

$$Z = \psi(cD,T) = \text{Sign}(cD).\times\max\{(|cD|\text{-}T), 0\} \tag{5}$$

Hard threshold function which can be stated as:

$$Z = \psi(cD, T) = cD.\times L(|cD| \geq T) \tag{6}$$

where, (cD, T) is the logic function which procures 1 if coefficient value higher than T, else, it yields 0 (Chang *et al*., 2000).

Proposed system description: Figure 2 describes the proposed denoising technology. The conventional denoising methods sophisticated to predict the threshold value relying on the statistical calculations of the noisy signal when some of the original signal and noise distributions are known. This is not compatible with practical applications where only the corrupted signal is determined. Therefore, a proposed denoising technique of speech signal is developed without any prior knowledge about signal and noise distributions. The proposed technique utilizes kurtosis function of remaining noise to achieve the best threshold value when kurtosis value is maximal and then employs an HBFOPSO algorithm to reach to the ideal value after a number of iterations. Kurtosis function is expressed in Eq. 7 (Graf *et al*., 2016):

$$\text{Kurt}(x) = \frac{E((x\text{-}m_x)^4)}{E((x\text{-}m_x)^2)^2}\text{-}3 \tag{7}$$

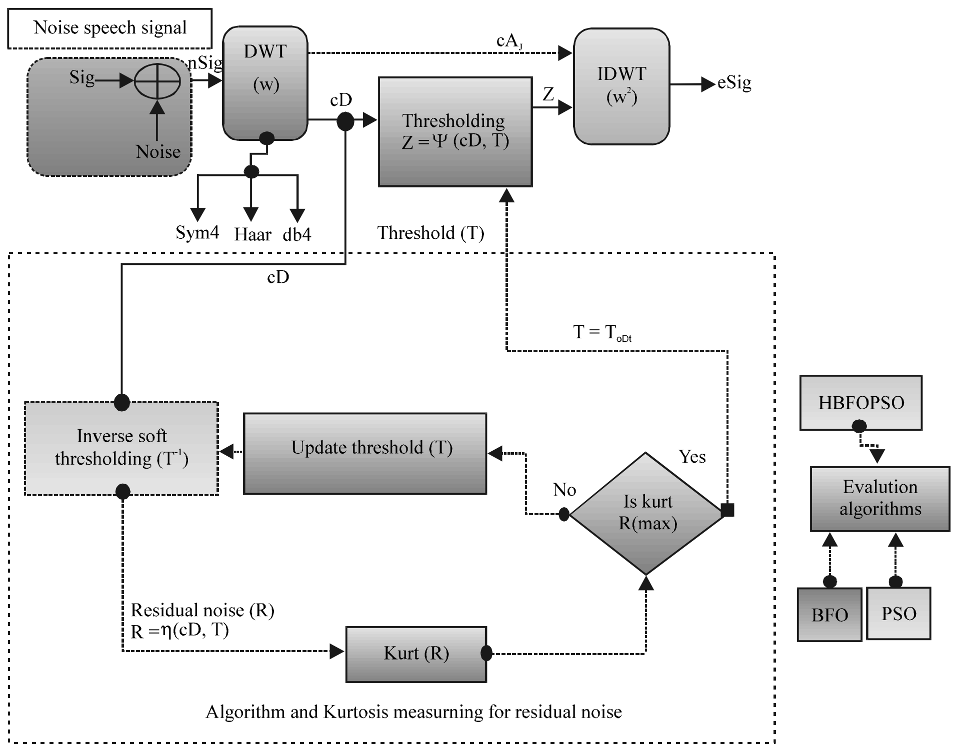

Fig. 2: The proposed denoising model

where, x is a random variable, E (.) denotes the expectation operation. Here, zero-mean $E(x^2) = 0$ is assumed. The executed method begins with applying DWT on the noisy speech signal to extract the wavelet coefficients which are the Approximation coefficients (cA) and the Details coefficients (cD).Then applying soft thresholding to the noisy coefficients in order to get the noiseless coefficients. Finally, apply the inverse threshold (R) to decrease the input coefficient by threshold T if its value <2T otherwise, set its value into T.

$$R = \eta(cD.T) = Sign(|cD|).\min(|Cd|-T.T) \quad (8)$$

Optimization algorithms

Particle Swarm Optimization (PSO): PSO is an optimization algorithm presented by Kennedy and Eberhart (1995), inspired by the birds natural foraging behavior to find an optimal solution. In the PSO, the particles are distributed randomly in the searching area with Velocity (V_i) was explained by Eq. 9 and varying their positions with Eq. 10. These particles follow the best one which is nearer to the aim and has better fitness value (Bhandari *et al.*, 2016).

$$v_i[t+1] = wv_i[t]+c_1r_1(x_{i,best}x:[t]+c_2r_2(x_{g.best}[t]-x_i[t] \quad (9)$$

$$x_i[t+1] = x_i[t]+v_i[t+1] \quad (10)$$

where, x_i is a particle position at each iteration, position vector is $x_i[t]$ is a fitness function, c_1, c_2: learning coefficients, r_1, r_2: random values, w is weight inertia, v_i is the ith particles velocity. ($pbest^t_i$) personal best is the position vector for the best fitness and ($gbest^t_i$) global best is the global best answer of all particles in the population.

Bacterial Foraging Optimization (BFO): BFO is an optimization technique introduced by Prof. K.M. Passino in 2002 depends on social behavior of *E. Coli* bacteria presented in human intestine. It follows several stages as a chemotaxis, reproduction, swarming and dispersal and elimination events (Sharma and Gary, 2012).

In chemotaxis steps, *E. coli* bacteria movement are performed by two various ways, first one is a swimming which indicates to the movement in same direction, the second one is the tumbling means the movement in a random direction. The bacterium movement after each step is achieved by Eq. 11 (Xing and Gao, 2014).

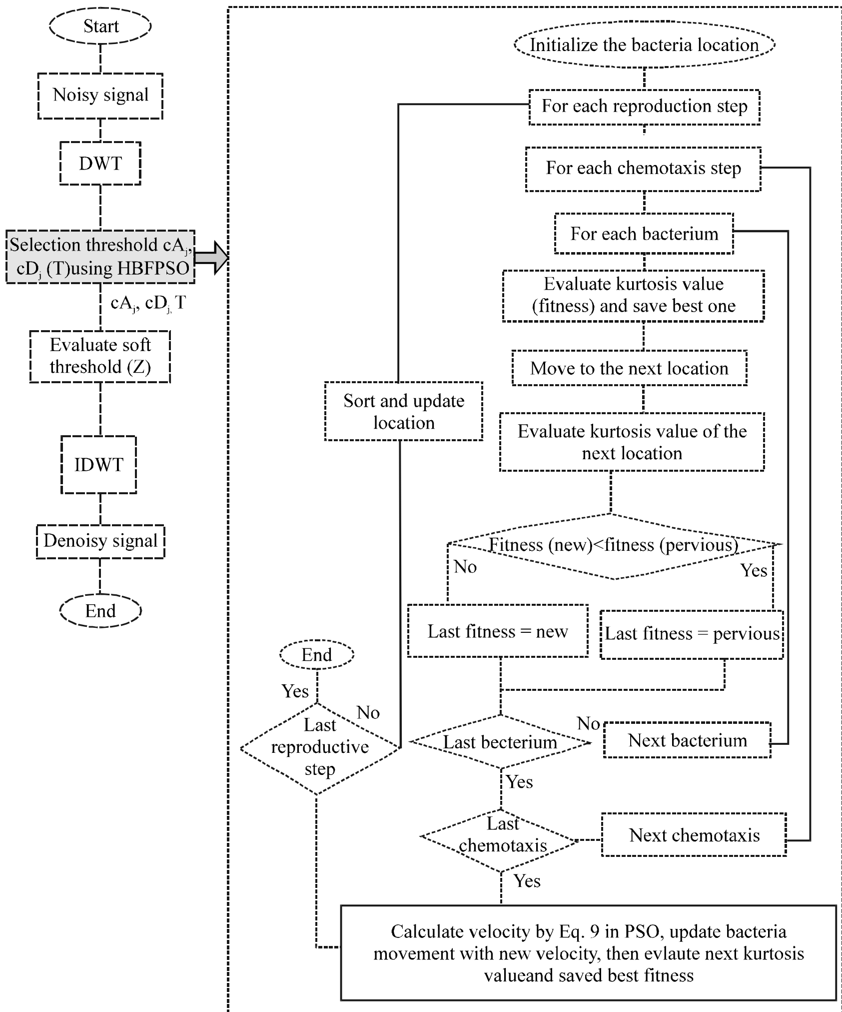

Fig. 3: Flowchart of denoising method based on HBFPSO

$$P^i(j+1,k,l) = P^i(j,k,l)+C(i)\frac{\text{Delta}(i)}{\sqrt{\text{Delta}(i)^T\,\text{Delta}(i)}} \quad (11)$$

where, $P^i(j, k, l)$ means theith bacterium location at (jth) chemotactic steps, (kth) reproductive steps and (lth) elimination steps and C(i) is step size. Through swarming step, bacteria move out their places in a ring of cells.

Through reproduction steps, the bacteria of worst health die and another bacteria divided into two and placed in same position, therefore, bacteria in population remain constant. Then sort bacteria as a higher cost means lower health showed in Eq. 12.

$$J^i_{\text{health}} = \sum_{j=1}^{Nc+1} J(i,j,k,l) \quad (12)$$

where, i = 1, ..., S, S: is a bacteria number in population, Nc: is a chemotaxis steps number and J () is a fitness function. Dispersion process is obtained after a specific reproduction steps.

Hybrid optimization algorithm (BFPSO): Hybrid BFPSO algorithm is presented to raise the convergence speed and minimize the local. In this hybrid combination, PSO achieves a global search and creates a near optimal answer very quickly and then followed by a local search by BFO which presents the optimal solution of high accuracy (Cheng and Lien, 2012).

PSO has an essential drawback of existence trapped in optimum local, although, it has high speed whereas BFO has a disadvantage which is a very low convergence speed, however, it has not trapped in local optimal. The main implementing steps of HBFPSO can be summarized as follows:

Algorithm 1; HFDSO algrithm:

- Step 1) initialize bacterium location
- Step 2) For each one of four operations like chemotaxis, swarming, reproduction and elimination and dispersal steps estimate kurtosis values and saved best one.
- step 3) Estimate kurtosis value for next location
- Step 4) Choose the best one between present and previous value
- Step 5) Calculate velocity by Eq. 9 in PSO
- Step 6) Update bacteria movement with new velocity, then evaluates new kurtosis value and saved the best fitness

The optimized threshold value is obtained after specific steps are shown in the flowchart that stated in Fig. 3.

RESULTS AND DISCUSSION

To investigate the proposed denoising method performance system is implemented by MATLAB 2016 as shown in Fig. 3.

Original speech signal analysis: Speech signal utilizes to check the suggested denoising technique, this speech signal has L = 435925 symbol length with various frequency ranges as seen in Fig. 4. The tested speech signal is corrupted by AWGN noise with several SNR = 5, 10 and 15.

Kurtosis function of residual noise: In this study a one level with db4 wavelet filter type has been employed for decomposition each noisy signal to different approximates coefficients and details, each of them with (217966 samples). Kurtosis criteria evaluated for detail part after thresholding them by inverse threshold as in Eq. 7. It is noticed from Fig. 5 of tested signal with various SNR levels that there is optimal threshold (T_{opt}) obtained when kurtosis value is maximum.

Threshold estimation using HBFPSO algorithm

HBFPSO algorithm for one level: A one decomposition level of noisy speech signal at SNR = 10 dB are considered as an instance case. The signal decomposed into approximations and detail coefficients each of which with 217966 samples as shown in Fig. 6. After applying the proposed algorithm, the optimal threshold value is THBFPSO = 1.6235 with m aximum kurtosis value is a -1.2005 of residual noise signal. The convergence behavior of HBFPSO algorithm is shown in Fig. 7.

The denoising process utilizing HBFPSO is explained in Fig. 6 and denoising results are recorded in Table 1. To confirm the HBFPSO algorithm performance, the proposed signal denoising way are tested with single algorithm as BFO and PSO and compares its results with HBFPSO. The comparison based on output SNR, MSE and iterations number at same condition. The results verify that HBFPSO algorithm presented a lesser (MSE) than BFO and PSO for all noise levels. The denoising results verify that HBFPSO algorithm offered a lower MSE than BFO and PSO for all SNR used. HBFPSO appeared good performance compared to others. For instant at 10 dB SNR, MSE of HBFPSO was 0.019107 while for PSO and BFO the MSE were 0.021389, 0.019457, respectively as appeared in Table 1.

Finally T_{HBFPSO} value used to threshold the detail coefficients and the resultant is reconstructed with approximation coefficients using IDWT to obtain de-noised signal as shown in Fig. 8.

HBFPSO algorithm of Multi DWT level: For more certification, the HBFPSO algorithm is operated on three decomposition levels. Detail coefficients (cD) of a three levels are seen in Fig. 9 with SNR = 10 dB.

After utilizing the proposed algorithm, the obtained values of threshold, number of iteration and highest kurtosis for each are:

- Detail coefficient (cD1):T1 = 0.75878 | niter = 10 | Kurtmax1 = -1.697
- Detail coefficient (cD2):T2 = 0.25579 | niter =10 | Kurtmax2 = -1.6935
- Detail coefficient: cD3:T3 = 0.26218 | niter = 10 | Kurtmax3 = -1.2615

The convergence behaviorand denoising process of HBFPSO algorithm for three decomposition levels are showed in Fig. 10 and 11. Table 2 illustrates the HBFPSO, PSO and BFO comparison of three decomposition levels for different SNR levels. From Table 2, it's obviously that HBFPSO hybrid algorithm exposed a least MSE in all cases than BFO and PSO algorithms.

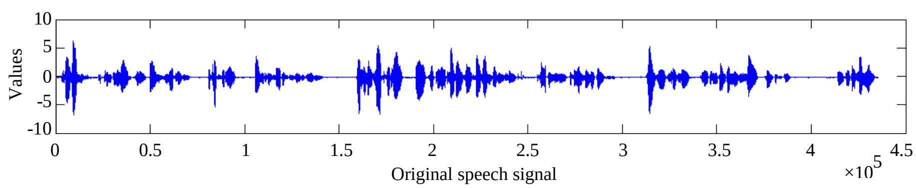

Fig. 4: Original speech signal

Table 1: Comparison between (HBFPSO, PSO and BFO) results for various SNR

	HBFPSO		BFO		PSO	
Wavelet types/SNR1	MSE	SNR2	MSE	SNR2	MSE	SNR2
Haar						
5	0.06109	7.7458	0.061539	7.714	0.062032	7.6794
10	0.021191	12.344	0.021407	12.3	0.02195	12.1911
15	0.007089	16.7466	0.007925	16.0399	0.008448	16.3377
Sym4						
5	0.0583	7.9488	0.058589	7.9274	0.067543	7. 3097
10	0.019011	12.8155	0.019055	12.8054	0.021472	12.2868
15	0.006701	17.3437	0.006718	17.3329	0.006934	17.1951
dB4						
5	0.058176	7.9581	0.06206	7.6774	0.067149	7.3351
10	0.019107	12.7935	0.019457	12.7148	0.021389	12.3036
15	0.006666	17.3663	0.006674	17.3616	0.006917	17.2058

Table 2: Comparison between (HBFPSO, PSO and BFO) algorithms for different SNR of multi-level

	HBFPSO		BFO		PSO	
Wavelet types/SNR1	MSE	SNR2	MSE	SNR2	MSE	SNR2
Haar						
5	0.030927	10.7022	0.03107	10.6821	0.031595	10.6093
10	0.014472	14.0004	0.015983	13.569	0.017368	10.3236
15	0.005140	17.0682	0.007699	16.6295	0.011466	15.0115
Sym4						
5	0.023442	11.9056	0.058589	7.9274	0.025994	11.4568
10	0.010065	15.57720	0.019055	12.8054	0.016103	13.5363
15	0.0045685	19.0078	0.006718	17.3329	0.005786	17.9812
dB4						
5	0.022999	11.9508	0.02436	11.7388	0.027879	11.1527
10	0.010593	15.3553	0.014457	14.0048	0.018694	12.8886
15	0.004540	19.0344	0.005459	18.2343	0.006125	17.7339

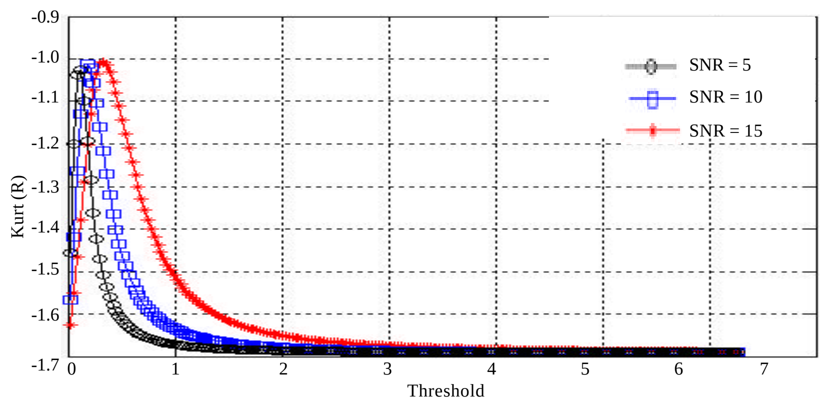

Fig. 5: Kurtosis measurning of residual noise at different SNR levels for speech signals

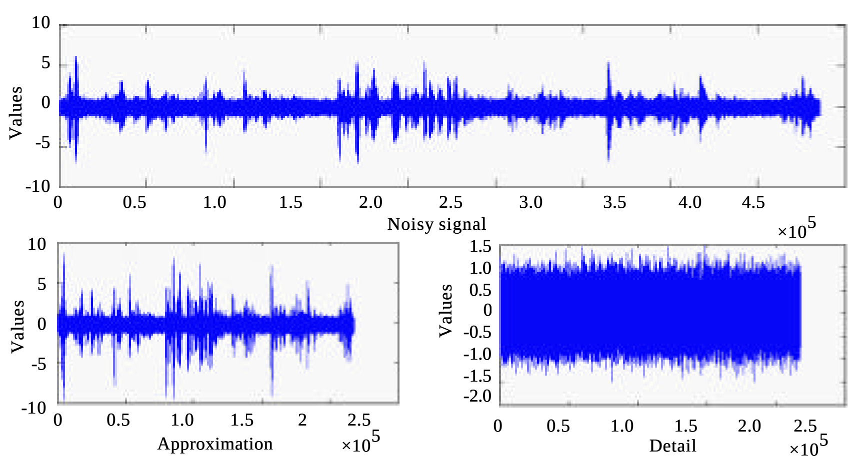

Fig. 6: One decomposition level at SNR = 5 dB

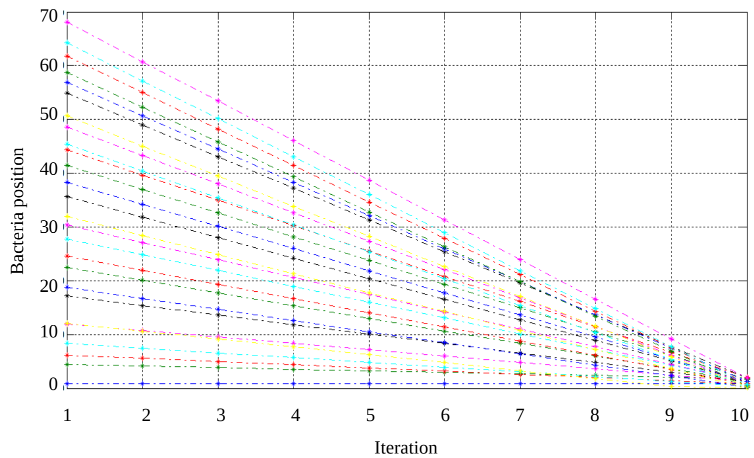

Fig. 7: HBFPSO convergence behavior for one decomposition level (SNR = 5 dB)

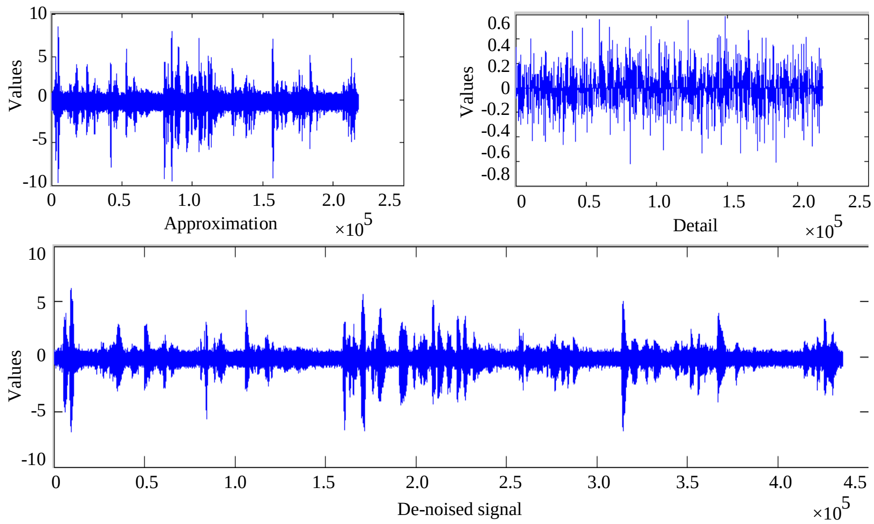

Fig. 8: Denoised signal of one decomposition level at SNR = 5 dB

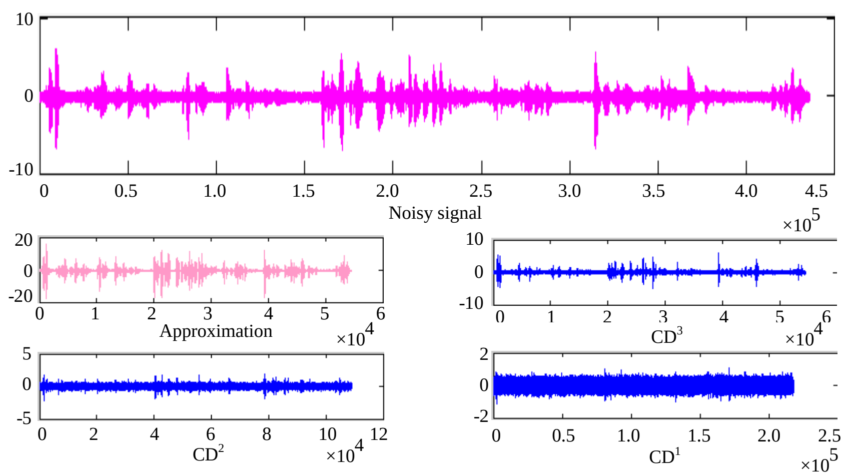

Fig. 9: Three decomposition level with SNR = 5

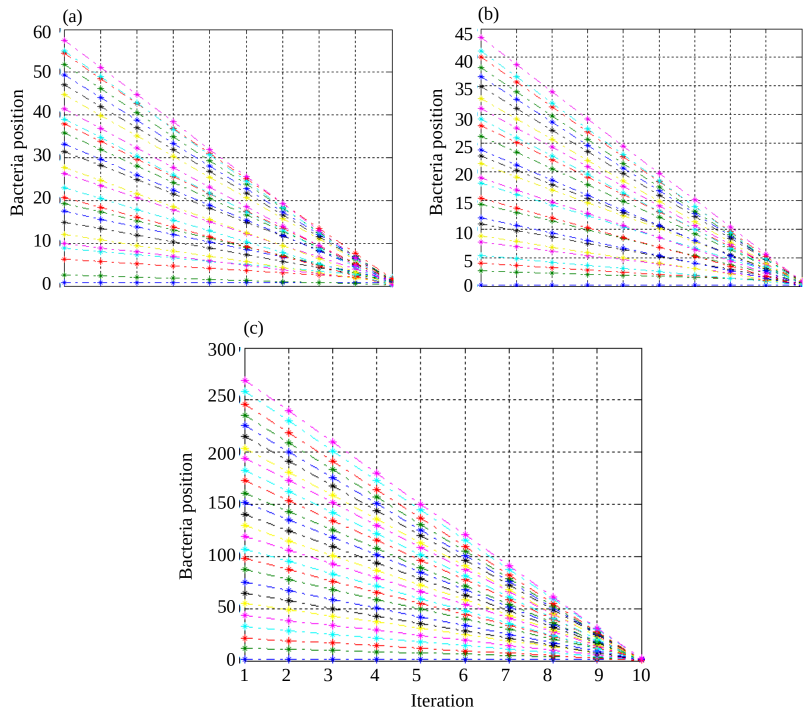

Fig. 10: HBFPSO Convergence behavior at three level; a) Convergance of HBFOPSO of T1; b) Convergance of HBFOPSO of T2 and c) Convergance of HBFOPSO of T3 (SNR = 10 dB)

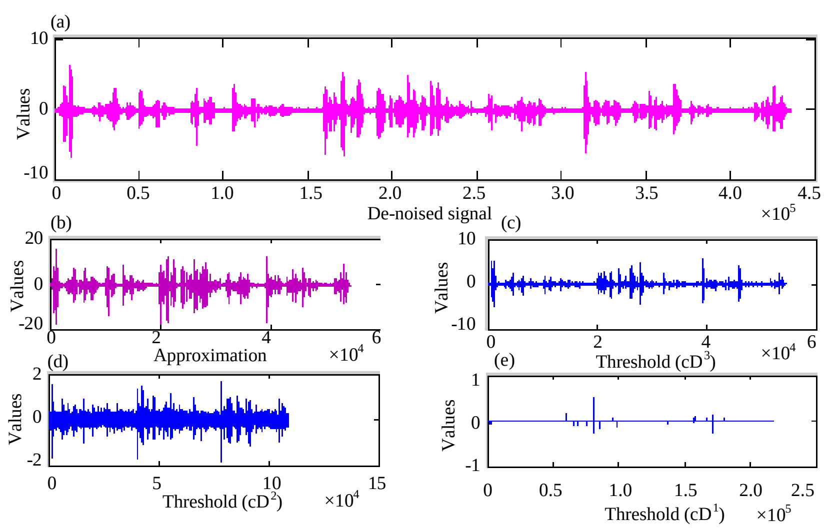

Fig. 11: Denoised signal, approximation and threshold detail coefficients of three level at SNR = 10 dB

CONCLUSION

A new denoising technique based on a hybrid optimization algorithm (HBFPSO) is suggested and executed at one and three decomposition levels. In addition, theoretical models are modified to estimate the denoising threshold by estimate the maximum kurtosis value of Residual (R) noise signal. The numerical results confirmed the effectiveness of this algorithm. Furthermore, the HBFPSO algorithm achieved a least MSE than the BFO and PSO algorithms for both one and three decomposition level. For instance, in three decomposition levels, HBFPSO algorithm presented an improvement in MSE around 0.004 compared with single algorithms that used at the same condition.

RECOMMENDATIONS

Future research on this approach will include modified hybrid thresholding methods with other noise types like babble, realistic and F16 noise also using multi-wavelet transform instead of wavelet transform.

REFERENCES

Arora, P. and M. Bansal, 2012. Comparative analysis of advanced thresholding methods for speech-signal denoising. Intl. J. Comput. Appl., 59: 28-32.

Arora, P., K. Singh, B. Singh and S.M. Sahib, 2017. Denoising of speech signals using wavelets. Intl. J. Res. Appl. Sci. Eng. Technol., 5: 58-63.

Bhandari, A.K., A. Kumar, G.K. Singh and V. Soni, 2016. Performance study of evolutionary algorithm for different wavelet filters for satellite image denoising using sub-band adaptive threshold. J. Exp. Theor. Artif. Intell., 28: 71-95.

Cengiz, Y. and Y.D.U. Arioz, 2016. An application for speech denoising using discrete wavelet transform. Proceedings of the 2016 20th International Conference on Biomedical Engineering Meeting (BIYOMUT), November 3-5, 2016, IEEE, Izmir, Turkey, ISBN:978-1-5090-5830-3, pp: 1-4.

Chang, S.G., B. Yu and M. Vetterli, 2000. Adaptive wavelet thresholding for image denoising and compression. IEEE Trans. Image Process, 9: 1532-1546.

Chen, X., S. Li and W. Wang, 2015. New de-noising method for speech signal based on wavelet entropy and adaptive threshold. J. Inf. Comput. Sci., 12: 1257-1265.

Cheng, M.Y. and L.C. Lien, 2012. Hybrid artificial intelligence-based PBA for benchmark functions and facility layout design optimization. J. Comput. Civil Eng., 26: 612-624.

Gao, R.X. and R. Yan, 2010. Wavelets: Theory and Applications for Manufacturing. Springer, Berlin, Germany, ISBN:978-1-4419-1545-0, Pages: 221.

Graf, S., T. Herbig, M. Buck and G. Schmidt, 2016. Kurtosis-controlled babble noise suppression. Proceedings of the 2016 International Symposium on Speech Communication 12. ITG, October 5-7, 2016, VDE, Paderborn, Germany, ISBN:978-3-8007-4275-2, pp: 1-5.

Kumari, V.S.R. and D.K. Devarakonda, 2013. A wavelet based denoising of speech signal. Intl. J. Eng. Trends Technol., 5: 107-115.

Liu, H., R. Zhang, Y. Zhou, X. Jing and T.K. Truong, 2017. Speech denoising using transform domains in the presence of impulsive and Gaussian noises. IEEE. Access, 5: 21193-21203.

Munegowda, B.K., 2016. Performance and comparative analysis of wavelet transform in denoising audio signal from various realistic noise. Ph.D Thesis, Edinburgh Napier University, Edinburgh, Scotland, UK.

Oktar, M.A., 2017. Offline and real time noise reduction in speech signals using the discrete wavelet packet decomposition. Ph.D Thesis, University of the West of England, Bristol, England.

Saikia, B.J., U. Baruah and Y. Surajkanta, 2015. Speech signal denoising using ICA-DWT and different shrinkage algorithms. Intl. J. Res. Sci. Innovation, 2: 66-75.

Santosh, D.H.H., V.S. Pendyala, V.L. Kumar and N.S. Rao, 2012. Statistical approach for noise removal in speech signals using LMS, NLMS, Block LMS and RLS Adaptive filters. Intl. J. Model. Optim., 2: 351-355.

Shaban, A., H.J. Abd and N. Salim, 2017. An improved technique for speech signal De-noising based on wavelet threshold and invasive weed optimization algorithm. J. Univ. Babylon, 25: 1413-1423.

Sharma, V., S.S. and T. Garg, 2012. A review of bacterial foraging optimization and its applications. Intl. J. Comput. Appl., 1: 9-12.

Sumithra, A. and B. Thanushkodi, 2009. Performance evaluation of different thresholding methods in time adaptive wavelet based speech enhancement. IACSIT. Intl. J. Eng. Technol., 1: 42-51.

Thyagarajan, K.S., 2005. Digital Image Processing with Application to Digital Cinema. Focal Press, Waltham, Massachusetts, USA., ISBN-13:978-0-240-80729-4, Pages: 371.

Wang, S. and Y. Shi, 2009. An improved speech denoising algorithm based on adaptive least mean square. Proceedings of the 2009 International Conference on Industrial and Information Systems, IIS'09, April 24-25, 2009, IEEE, Haikou, China, ISBN:978-0-7695-3618-7, pp: 293-296.

Wu, Z., 2014. Signal De-noising method based on particle swarm algorithm and Wavelet transform. Tehnicki Vjesnik, 21: 1001-1008.

Xing, B. and W.J. Gao, 2014. Innovative Computational Intelligence: A Rough Guide to 134 Clever Algorithms. Springer, Gewerbestrasse, Switzerland, ISBN: 978-3-319-03403-4, Pages: 260.