An Improved Technique for Speech Signal De-noising Based on Wavelet Threshold and Invasive Weed Optimization Algorithm

Ali Shaban **Haider. J. Abd**
University of Babylon, College of Engineering, Department of electrical engineering
eng_ali47@yahoo.com haiderlaser@yahoo.com
Noora Salim
Alqasim Green University, College of water resources engineering
noora.salim27@gmail.com

Abstract

Speech signals play a significant role in the area of digital signal processing. When these signals pass through air as a channel of propagation, it interacts with noise. Therefore, it needs removing noise from corrupted signal without altering it. De-noising is a compromise between the removal of the largest possible amount of noise and the preservation of signal integrity. To improve the performance of the speech which displays high power fluctuations, a new speech de-noising method based on Invasive Weed Optimization (IWO) is proposed. In addition, a theoretical model is modified to estimate the value of threshold without any priority of knowledge. This is done by implementing the IWO algorithm for kurtosis measuring of the residual noise signal to find an optimum threshold value at which the kurtosis function is maximum. It has been observed that the proposed method appeared better performance than other methods at the same condition. Moreover, the results show that the proposed IWO algorithm offered a better mean square error(MSE) than Particle Swarm Optimization Algorithm (PSO) for both one and multilevel decomposition. For instance, IWO brought an improvement in MSE in the range of 0.01 compared with PSO for multilevel decomposition.

Keywords: Signal de-noising, Discrete wavelet transform, Invasive Weed Optimization, Kurtosis.

الخلاصة:

إشارات الكلام تلعب دورا هاما في مجال معالجة الإشارات الرقمية. عندما تمر هذه الإشارات عن طريق الهواء كقناة للانتشار، فانها تتفاعل مع الضوضاء. ولذلك، فإنه يحتاج إزالة الضجيج من الإشارة التالفه بدون تاثر تلك الاشاره. طريقة تقليل الضوضاء تمثل حلا وسطا بين إزالة أكبر قدر ممكن من الضجيج والحفاظ على سلامة الإشارة. لتحسين أداءالكلام الذي يعرض تقلبات عالية الطاقة، طريقة جديدة لتقليل الضوضاء تقوم على أساس تحسين الأعشاب الغازيه (IWO) كانت قد اقترحت. بالإضافة إلى ذلك، تم تعديل نموذج نظري لتقدير قيمة العتبة دون أي معرفه مسبقه. ويتم ذلك من خلال تنفيذ خوارزمية IWO لقياس التفرطح للإشارة الضوضاء المتبقية لإيجاد قيمة العتبة المثلى االتي تكون عندها قيمة دالة التفلطح اعظم مايمكن. وقد لوحظ أن الطريقة المقترحة أبدت أداء أفضل من الطرق الأخرى تحت نفس الظروف. علاوة على ذلك، فقد بينت النتائج أن خوارزمية IWO المقترحة قدمت أفضل متوسط مربع الخطأ (MSE) من خوارزمية تجمع الجسيمات لمستوى واحد ولمتعدد المستويات. على سبيل المثال، خوارزمية IWO جلبت تحسنا في MSE في حدود 0.01 مقارنة مع خوارزمية تجمع الجسيمات ولمتعددة المستويات.

الكلمات ألمفتاحيه: تقليل ضوضاء الاشارة ، اشارات الكلام ، خوارزمية تحسين الاعشاب الغازيه ، متوسط مربع الخطأ، العتبه العكسيه

1. Introduction

Speech enhancement methods can be used to improve the quality of the speech processing equipment such as mobile telephony, digital hearing aids and human-machine and make them more efficient under noisy environment (Ganesh and Dinesh, 2011). Speech signals are the acoustic signals and have the corrupted noise. The noise is generally classified into two sources i.e. noise through a channel or due to the wrong nature of devices. Additive White Gaussian Noise (AWGN) is a noise channel. It pollutes the transmitted signals when signals passes through it (Bhaskar *et.al.*, 2015).To reduce the defect of this noise, different methods have been reviewed to enhance the speech signal and reduce corrupted noise (Sumithra *et.al.*, 2009; Mallat *et.al.*, 1992; Shuqi *et.al.*, 2009; Zhiyong *et.al.*, 2010; Arvind *et.al.*).Fourier domain was for long time the optimal manner to suppress the noise (Sumithra *et.al.*, 2009).

Recently, methods that based on the wavelet transformation have become growing in popularity (SU Li *et.al.*, 2009; Gao *et.al.*, 2011; KS Thyagarajan, 2008).Wavelets offer effective tool for non-linear filtering of signals affected by noise. The main aim of the wavelet transforms is to separate the high and low frequency components in the signal made it works as a superior technology in the area of signal de-noising.

In the speech signal, the noises are frequently localized at high frequency components. Therefore, the needing for wavelet transform becomes very useful to decompose the signal into its different frequency components and then extract the noise by thresholding it.The thresholding criteria in the de-noising process is still the main challenging and should be selected carefully (Gao and Yan, 2011; KS Thyagarajan, 2008). If the value of the threshold was high, this will lead to destroy the signal data, while the low value of the threshold will keep the noisy data. The threshold selection has been derived in a Bayesian method using generalized Gaussian distribution (GGD) as a probabilistic model of the signal wavelet coefficients (Pankaj and Swati, 2011; Grace *et.al.*, 2000). The mean square error (MSE) is utilized as a fitness functions for the optimization algorithms (Gopinath *et.al.*, 2014;Xing and SiweiLyu, 2014).

In this work, IWO algorithm is proposed as an optimized solution to estimate the optimal threshold. It depends on criteria for fitness function that count on the kurtosis measuring for the estimated residual noise signal. In addition, an inverse threshold function was used to evaluate this residual noise from the detail coefficients of the DWT of the noisy signal. The algorithm supposes that there is a single value for the threshold called optimum threshold that maximizes kurtosis value of the residual noise which is then discovered by IWO algorithm. The proposed algorithm was applied on the speech signal in the presence of DWT de-noising method and compared with PSO optimized algorithm. The results prove that IWO algorithm revealed higher speech signal with lower MSE than PSO algorithm under similar system.

2. Speech signal separation using Discrete Wavelet Transform

The process of discrete wavelet transform is separation or decomposition process where the levels approximate sequence is decomposed into the next one of levels approximate sequences and detail sequences. This analytical method is applicable to the rich low frequency part of signals, such as images, voice and so on. A noisy signal with gaussian noise is formulated as:

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

Where, $nSig$: refers to the observed noisy signal, Sig: indicates the unknown original signal and $Noise$: is the gaussian noise with zero mean and finite variance σ^2.The target is to extract the noise, to obtain an evaluated $\{ eSig \}$ of the original $\{ Sig\}$ with lower value of mean square error (MSE):

$$MSE = \frac{1}{N}\sum_i^N (Sig_i - eSig_i)^2 \tag{2}$$

Where N represents the length of signal (should be integer power of **2**) (Grace *et.al.*, 2000).

Different de-noising methods are suggested to resolve this problem, but the most efficient one was by a method of wavelet transform. Wavelet transform is a famous instrument for signal analysis. In this method, the signal decomposed to more than one segment which belongs to various frequency components as in Fig. (1). To do that the signal should compare with a group of wavelet basis functions and then search their similarities in frequency contents (Gao and Yan, 2011).

Let W indicates to the orthogonal Discrete Wavelet Transform (DWT) matrix, the wavelet coefficients is:

$W \times nSig = W \times Sig + W \times Noise$

$$W_{nSig} = W_{Sig} + W_{Noise} \qquad (3)$$

Where, W_{Noise} is the noise since the transformation is orthogonal (Grace *et.al.*, 2000).

The wavelet de-noising method begin with trimming each coefficient of the detail subbands (cD) with a definite threshold to get threshold version of detail sub bands (Z). Then (Z) reconstructed with the approximation coefficients (cA) to produce the estimated signal where:

$$eSig = W^{-1} \times [cA_J\,, Z] \qquad (4)$$

Where, W^{-1}: is the Inverse Discrete Wavelet Transform (IDWT) operator (KS Thyagarajan, 2006).

There are two major threshold functions that always used.

A) Soft-threshold function (also named the shrinkage function), which is defined as:

$$Z = \psi(cD, T) = sign(cD).\times\, max\{(|cD| - T), 0\} \qquad (5)$$

It takes the argument and shrinks it toward zero by the threshold *T*.

B) *Hard-threshold function* (popular alternative) , which can be defined as:

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

Where $L(\,)$ is a logic function (0 or 1), this function takes 1 if its value larger than threshold *T* otherwise, it takes zero (Grace *et.al.*, 2000).

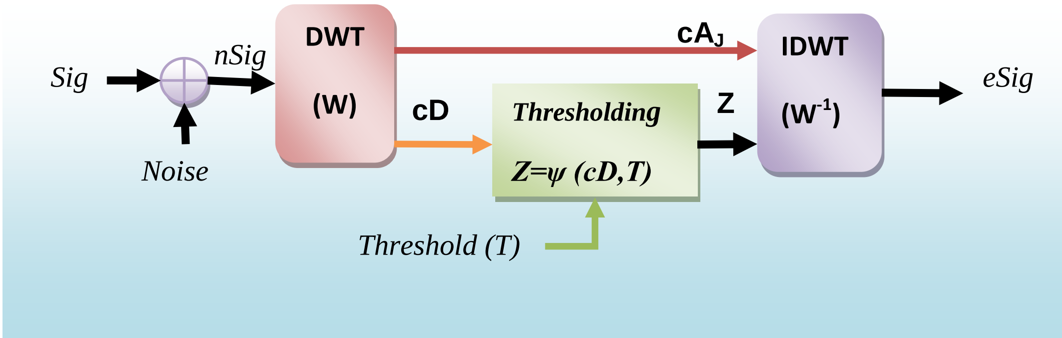


Figure (1): Conventional wavelet de-noising method.

3. De-noising threshold selection methods

The available wavelet de-noising methods are differing from each other by the way of selecting the threshold. A few threshold selection methods that have been investigated. These methods can be described as seen in sections (3.1, 3.2 and 3.3)

3.1. Visu-Shrink threshold: It is evaluated by expression depicted in equation (7)

$$T_{Visu} = \sigma_N \sqrt{2log(L)} \qquad (7)$$

Where σ_N represents a noise variance and L is a length of signal.

This method have used to minimize the maximum error over all possible L-sample signals

(Grace *et.al.*, 2000).

3.2. Sure-Shrink threshold: T_{Sure} is evaluated by the equation (8)

$$T_{Sure} = min\{t_J\,, \sigma\sqrt{2\log(L)}\} \qquad (8)$$

Where t_J represents the threshold value at J_{th} decomposition level in wavelet domain (Mantosh and Hari, 2013).

3.3. Bayes Shrink threshold: It is one of the significant methods. The Bayes method has a better mean square error (MSE) than the Sure-Shrink (S. Grace *et al.*, 2000). The Bayes Shrink threshold is given as mention in equation (9)

$$T_{Bayes} = \frac{\sigma^2}{\sqrt{\sigma_{Sig}^2}} \tag{9}$$

Where σ^2 *is* the noise variance and σ_{Sig}^2 is the variance of original signal.

4-Proposed scheme for speech signal de-noising

Fig (2) describes the proposed signal de-noising technique. Previously, methods of signal de-noising assume that there is some known knowledge for both signal and noise distributions with given parameters to estimate the value of threshold. Practically, only the noisy signal that observed is determined. Therefore, the wavelet de-noising method is developed without depending on the priority of knowledge as in Fig.(2). The developed model uses the Kurtosis statistic of the residual noise signal to obtain the optimum value for threshold at which the Kurtosis function becomes maximum, and then uses the IWO algorithm to reach the optimal threshold after specific iterations.

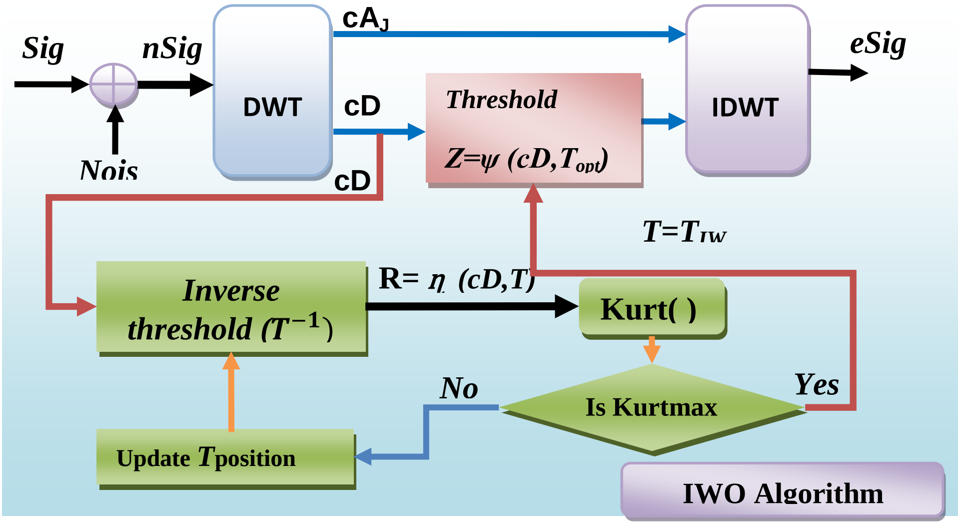


Figure (2): The proposed de-noising signal model

The normalized ***Kurtosis function*** for a random variable x is defined as in eq.(10) (Gopinath *et.al.*, 2014) :

$$kurt(x) = \frac{E((x-m_x)^4)}{\left(E((x-m_x)^2)\right)^2} - 3 \tag{10}$$

Where:*E(x)* is the predicted value of *x*.

The implemented method starts with applying DWT to noisy signal (*nSig*) to decompose it into approximation and detail coefficients. Then anew function is modified to extract this noise from the detail coefficients. Furthermore, this function is candidate as ***inverse threshold*** function and works to shrink the input by T if its absolute value smaller than 2T, otherwise, set into T.

$$R = \eta(cD, T) = sign(cD).\times \min\{|cD| - T\ , T\} \tag{11}$$

5- IWO Algorithm

Invasive Weed Optimization algorithm is a term can be collected from: the colony, seeds and invasive weeds in nature. It is based on weed biology and ecology. It was observed that by adjusting the properties of the invasive weeds, leads to a robust optimization algorithm (Andrze, 2002). To characterize the algorithm process, a new terms used to describe this algorithm should be introduced. Each individual containing a set of optimization variable is named a seed. Each seed develop to a flowering plant within colony. The term of a plant is single individual after estimating its fitness.

The IWO algorithm steps can be explained in Fig.(3). Producing of seeds in IWO algorithm are being disseminated in the search space by using normally distributed random numbers with mean equal to the location of the producing plants and varying standard deviations. The standard deviation (SD) can be expressed by equation (12).

$$\boldsymbol{\sigma_{iter}} = \frac{(\boldsymbol{iter_{max}} - \boldsymbol{iter})^{\boldsymbol{n}}}{(\boldsymbol{iter_{max}})^{\boldsymbol{n}}}\left(\boldsymbol{\sigma_{initial}} - \boldsymbol{\sigma_{final}}\right) + \boldsymbol{\sigma_{final}} \tag{12}$$

where $\sigma_{iter\ max}$ is the maximum number of iterations, $\sigma_{initial}$ and σ_{final} are defined initial and final standard deviations, respectively and n is the nonlinear modulation index.

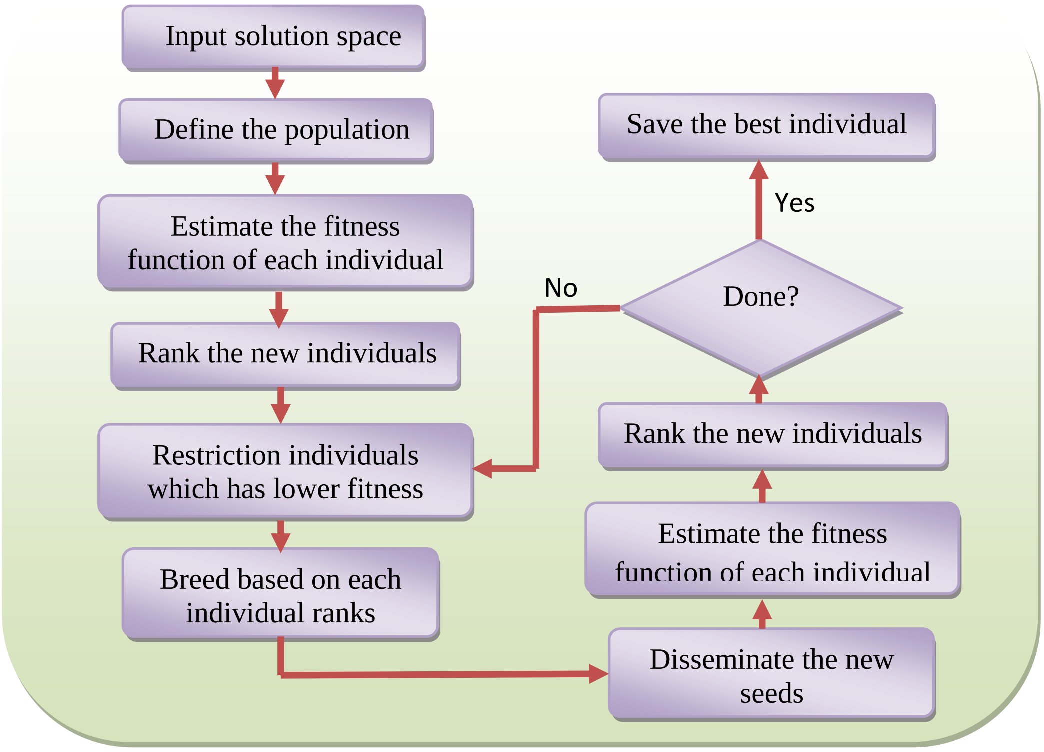


Figure (3): Flowchart describes IWO algorithm

6-Results and discussion

To investigate the performance of the proposed de-noising method, MATLAB2014A program have been used to implement the system shown in Fig. (2). The findings can be divided into four parts as follows stated through sections (6.1 to 6.4)

6.1 Analysis of original speech signal

In the proposed algorithm speech signal is used to test the proposed model, this signal has N=80000 symbol length with different frequency range as shown in Fig.(4). The tested signal is affected by Gaussian noise with SNR= 10,15, 20, and 25 to get a noisy signal $nSig$ from each one.

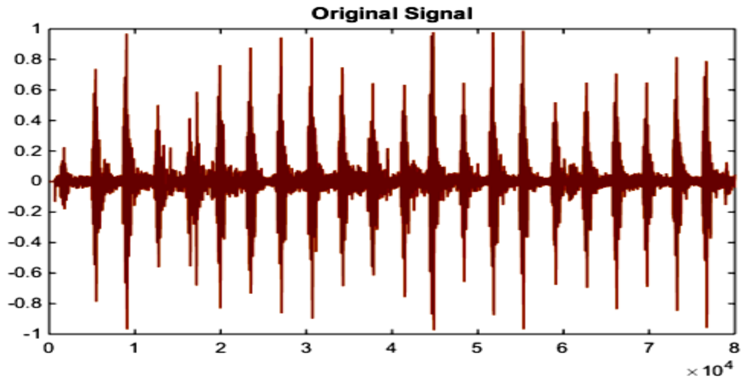


Figure (4): Speech signals tested in simulation.

6. 2 Kurtosis statistics behavior

In this work, a one level Haar DWT was used to decompose the noisy signal into approximation and detail coefficients with 1600 samples. The kurtosis function calculated for detail coefficients after thresholding process by the inverse soft threshold function Eq.(10). It can observed from Fig.(5) for a tested speech signal with different *SNR* that there is a single value for threshold called optimum threshold (T_{opt}) at which the kurtosis measuring function of residual noise (***R***) be maximum.

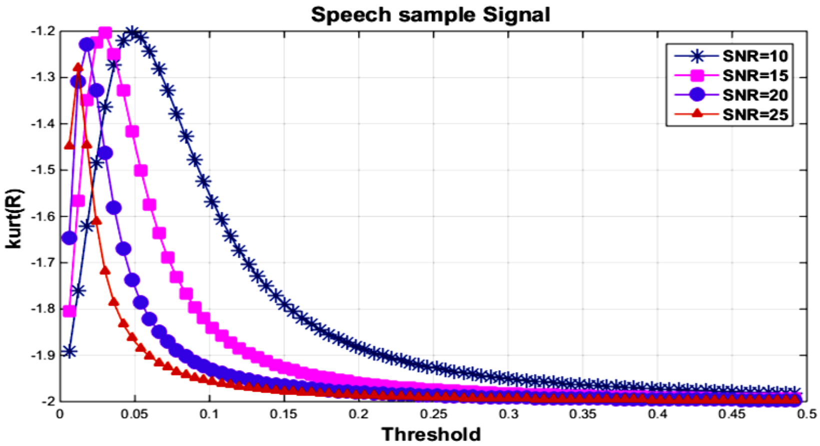


Figure (5): Kurtosis measuring of residual noise at different SNR levels for speech signal

To validate the proposed method, the optimum threshold value that obtained from kurtosis measuring is compared with the available thresholds such as Bayes Shrink and Visue-Shrink Eq.7 to Eq.11.The results shown in Table (1) stated that there is an optimal threshold obtained for each SNR .

Table (1) Comparison among different optimum thresholds methods

SNR dB	10 dB	15 dB	20 dB	25 dB
T_{opt}	0.048	0.03	0.018	0.012
T_{Bayes}	0.049046	0.027581	0.01551	0.0087217
Tvisue	0.036146	0.020326	0.01143	0.0064278

6.3. Optimum threshold value using IWO algorithm

6.3.1 IWO algorithm for one level DWT

A one level decomposition of a speech signal with SNR=10dB has been assumed as an example case. Here the signal separated into approximation and detail levels with 40000 samples as shown in Fig.(6). After applying the proposed algorithm, the optimum threshold value was $T_{IWO} = 0.048$ with maximum kurtosis= -1.2 for the residual noise. The convergence behavior of IWO algorithm in Fig.(7).

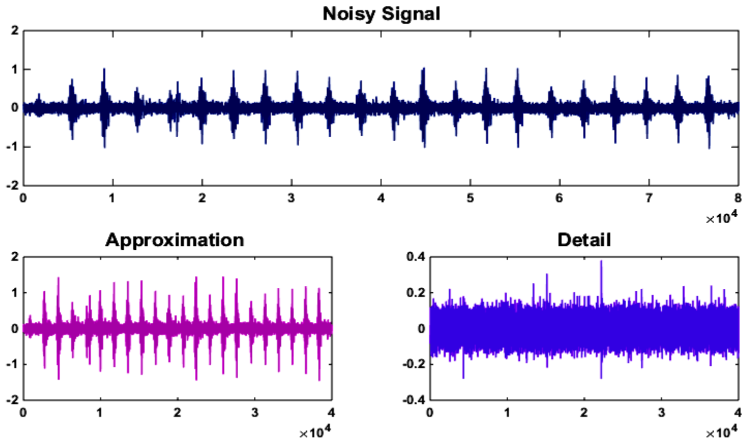


Figure (6): One level DWT decomposition at SNR=10dB

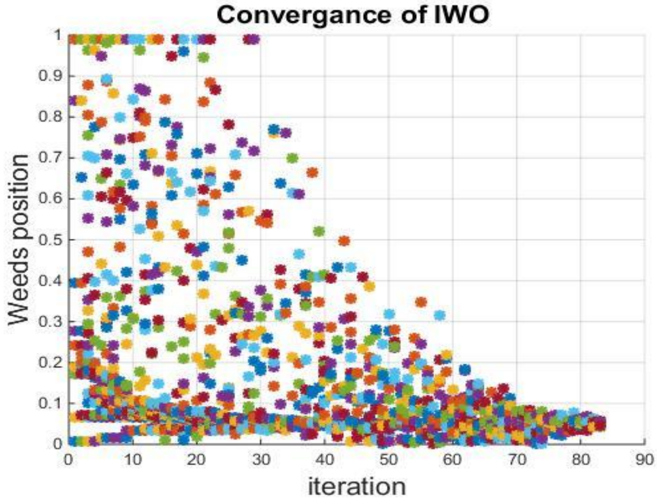


Figure (7): Convergence behavior of IWO of one level DWT decomposition at SNR=10dB.

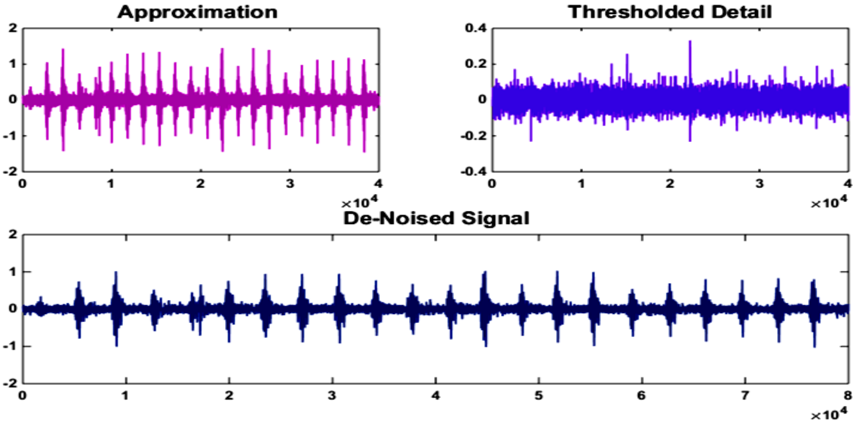
The de-noising process using IWO is illustrated in Fig.(6). It's obvious that the de-nosing performance was done using one level DWT. The same procedure used for this signal at four different *SNR* level and results recorded in Table (2).To confirm the performance of IWO algorithm, the proposed signal de-noising method was examined with recent algorithm like PSO and compare it with IWO(Dinesh K. Gupta *et al*., 2015). The comparison was done in term of both threshold value and MSE at the same system parameter. The findings prove that IWO algorithm offered a lower MSE than PSO for all SNR. For example, at10 dB SNR, the MSEs of IWO and PSO were 0.037 and 0.047 respectively as shown in Table (2)

Table (2) Comparison between IWO and PSO algorithms for different SNR

SNR dB	IWO			PSO		
	T_{IWO}	MSE	iter	T_{PSO}	MSE	iter
10 dB	0.048425	0.037959	80	0.037082	0.047403	22
15 dB	0.028418	0.021845	60	0.0027879	0.023343	23
20 dB	0.016868	0.012716	50	0.0016608	0.015001	24
25 dB	0.010706	0.0076943	40	0.0047545	0.0083785	28

6.3.2 IWO algorithm for multilevel DWT

For more verification, the IWO algorithm was applied with five levels decomposition. The detail coefficients of a five sub bands have been chosen as in Fig.(9) for a speech signal with SNR=10 dB .



Figure(8):De-noised signal using proposed algorithm in case of one level DWT at *SNR*=10dB.

After applying the suggested algorithm, the maximum kurtosis values, threshold and number of iteration for each detail subbands are:

Detail coefficient: cD1 : T1 = 0.19244 | num iter1 = 45 | Kurtmax1 = -1.1973
Detail coefficient: cD2 : T2 = 0.10297 | num iter2 = 45 | Kurtmax2 = -1.1994
Detail coefficient: cD3 : T3 = 0.060679 | num iter3 = 43 | Kurtmax3 = -1.2572
Detail coefficient: cD4 : T4 = 0.037165 | num iter4 = 40 | Kurtmax4 = -1.3074
Detail coefficient: cD5 : T5 = 0.022496 | num iter5 = 42 | Kurtmax5 = -1.3979

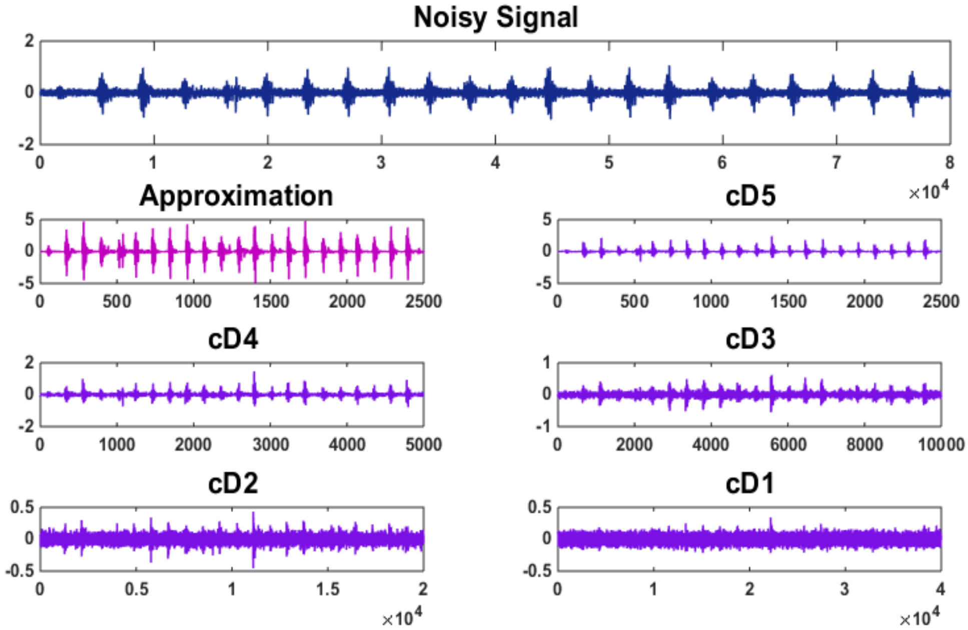
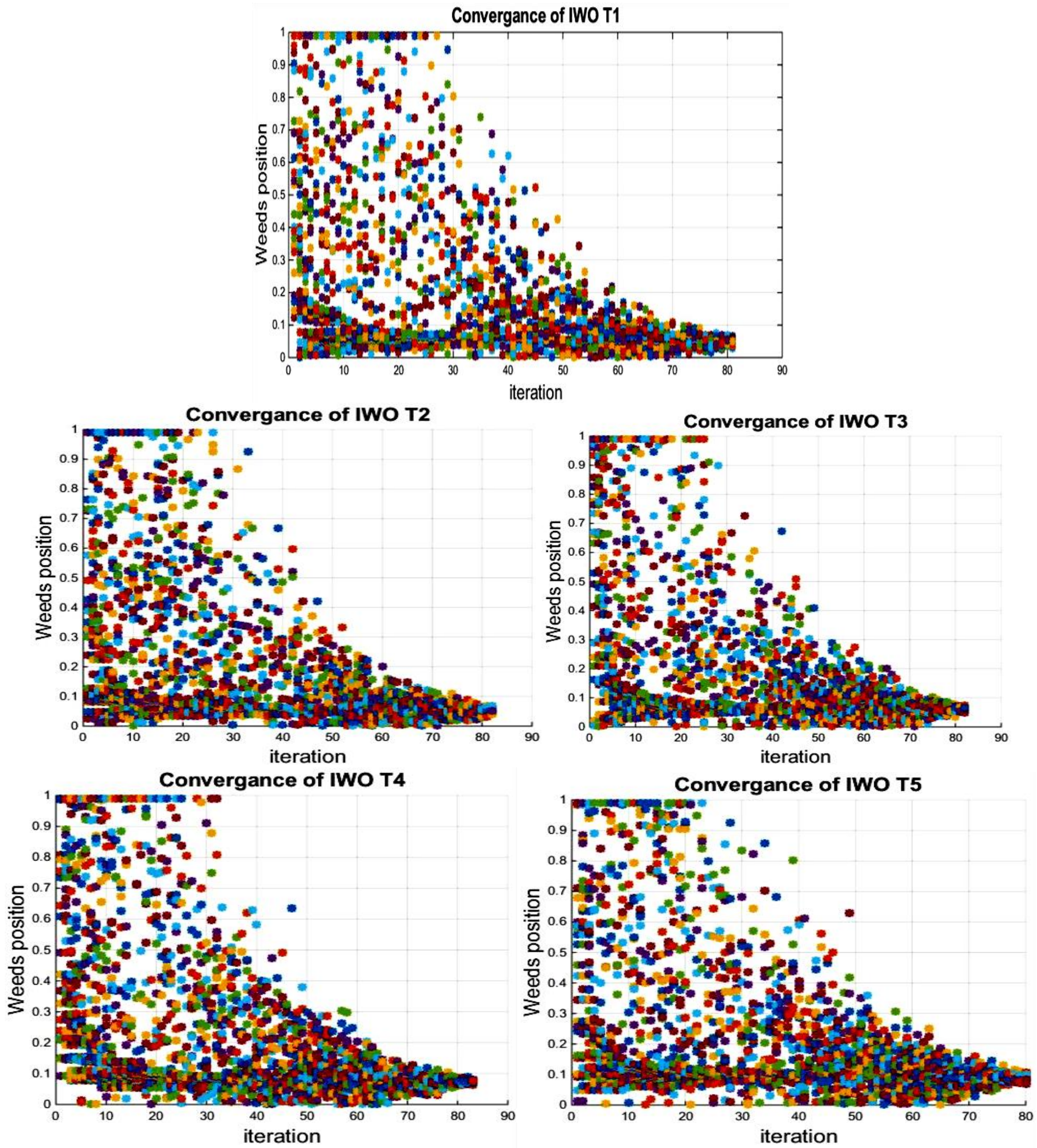


Figure: (9) Five level decomposition for speech signal with SNR=10.

The IWO algorithm behavior and de-noising process for multi level decomposition are illustrated in Figs (10) and (11) respectively. **Table (3)** shows the comparison between IWO and PSO of multi level decomposition for various SNR. From table (3), it's clearly that IWO algorithm revealed a best MSE for all decomposition levels than PSO algorithm.



Figure(10): Convergence behavior of IWO in case of five level DWT with noisy sine wave at *SNR*=10dB.

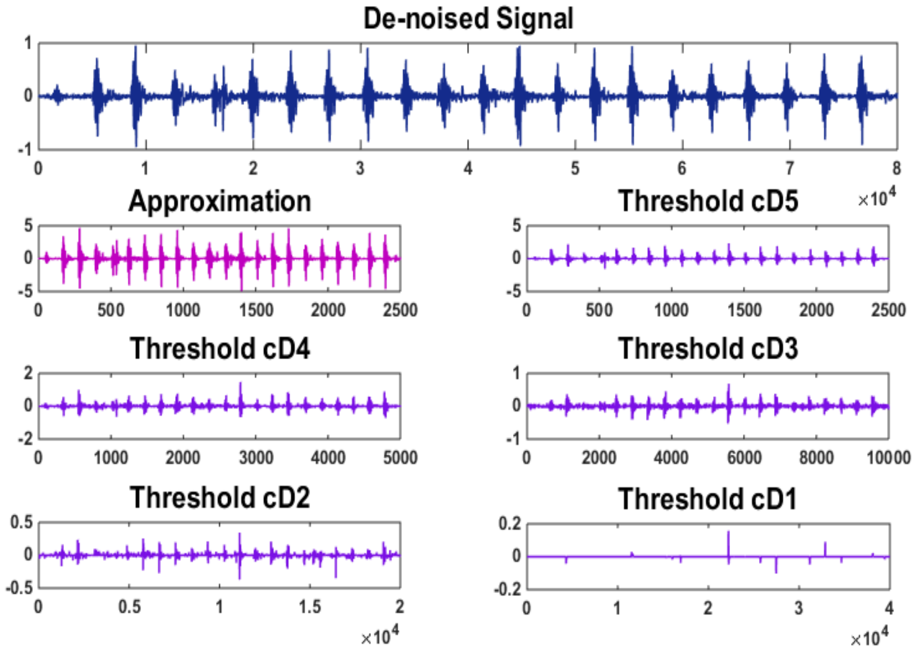


Figure (11): De-noised signal and threshold detail coefficients using proposed algorithm in case of five level DWT at *SNR*=10dB.

Table (3): MSE comparison between IWO and PSO algorithms in term of MSE in case of one and five levels DWT

SNR	IWO		PSO	
	MSE_{L5}	MSE_{L1}	MSE_{L5}	MSE_{L1}
10	0.024585	0.037959	0.025322	0.047403
15	0.017992	0.021845	0.018747	0.023343
20	0.013497	0.012716	0.014845	0.015001
25	0.007596	0.0076943	0.0097479	0.0083785

7. Conclusion

To enhance the speech signal, a new speech de-noising method based on Invasive Weed Optimization (IWO) has been proposed and implemented for both one and multilevel decomposition. As well a theoretical model is modified to evaluate the value of de-noising threshold by estimate kurtosis function of the residual noise signal. It has been noticed that the suggested method introduced better performance than other methods at the same environment. Furthermore, the results confirm that IWO algorithm released a better MSE than the PSO algorithm. For example, for multilevel decomposition, IWO algorithm offered an enhancement in MSE around 0.01 compared with PSO algorithm. Finally, IWO algorithm released optimum threshold which revealed a significant improvement in speech signal.

References

Andrzej CICHOCKI Shun-ichi AMARI, 2002, " Adaptive Blind Signal and Image Processing Learning Algorithms and Applications", John Wiley &Sons LTD

Arvind R. Yadav1*, Anand2 R. S., Dewal3 M. L., and Sangeeta Gupta4, Performance analysis of discrete wavelet transform based first-order statistical texture features for hardwood species classification,

Bhaskar J. Saikia, Ujwala Baruah, and Y Surajkanta, 2015, Speech Signal Denoising using ICA-DWT and Different Shrinkage Algorithms, IJRSI , Volume II, Issue IV, April

Dinesh K. Gupta, Satish Pawar, and Yogendra Kumar Jain, 2015, "Wavelet based Multilevel Sub-Band Adaptive Thresholding for Image De-noising using Modified PSO Algorithm", IARJSET, Vol.2, pp 24-30.

Ganesh R. Naik and Dinesh K. Kumar, 2011, Improving the quality of the audio sources using Gaussianity reduction technique, International Journal Of Electronics28 Jun

Gao, R.X. and Yan, R., 2011, "Wavelets Theory and Application for Manufacturing", Springer, New York Dordrecht Heidelberg London .

Gopinath V., Dr.NishatKanvel ,and Dr. T.K.Shanthi, 2014, Novel Image Compression Scheme with Wavelet Packets Optimization Using Artificial Bee Colony Algorithm"IJETT – Volume 4 Issue 8- August.

Grace Chang S., Bin Yu and Martin Vetterli, 2000, Adaptive Wavelet Thresholding for Image Denoising and Compression"IEEE TransactiononImage Processing, Vol. 9, No. 9, SEP. 2000, pp. 1532-1546

KS Thyagarajan, 2008,"Digital Image Processing with Application to Digital Cinema" Focal Press, Taylor& Francis 2008.

Mallat S. and Hwang W. L., 1992, Singularity detection and processing with wavelets," IEEE Trans.on Infomation Theory, vol. IT-38, pp. 617-643,

Mantosh Biswas and Hari Om, 2013, "An Image Denoising Threshold Estimation Method", Advances in Computer Science and its Applications (ACSA) Vol. 2, pp. 377-381.

Pankaj Hedaoo and Swati S Godbole, 2011, Wavelet Thresholding Approach for Image De-noising ,International Journal of Network Security & Its Applications (IJNSA), Vol.3, pp.16-21,

Shaya Karimkashiand Ahmed A. Kishk, 2010, Invasive Weed Optimization and its Features in Electromagnetics, IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, VOL. 58, NO. 4, APRIL 2010

Shuqi Wang and Yin Shi, 2009, An Improved speech denoising Algorithm Based on Adaptive Least Mean Square, International Conference on Industrial and Information Systems

SU Li, and ZHOU Xue-mei, 2009, Application of improved wavelet thresholding method for de-noising gyro signal, Journal of Chinese Inertial Technology, 17(2), 231-235.

Sumithra M G; A. and Thanushkodi K; B. 2009, Performance Evaluation of Different Thresholding Methods in Time Adaptive Wavelet Based Speech Enhancement, IACSIT International Journal of Engineering and Technology Vol.1,No.5,December, ISSN: 1793-8236

Xing Zhang, and SiweiLyu, 2014, Using Projection Kurtosis Concentration of Natural Images for Blind Noise Covariance Matrix Estimation, IEEE, Computer Vision Foundation.

Zhiyong He, Xuhong Guo, and Maoqing Zhang, 2010, Detection and Removal of Impulsive Colored Noise for Speech Enhancement, Proceedings of the IEEE International Conference on Information and Automation June 20 - 23, Harbin, China.