Speech Signal Denoising using ICA-DWT and Different Shrinkage Algorithms

Bhaskar J. Saikia[*1], Ujwala Baruah[2], Y Surajkanta[3]
Dept. of Computer Science, National Institute of Technology, Silchar, Assam

***Abstract*- Speech signals are the least studied signal set in digital communication. The speech signals pass through air as their first channel of propagation and during this phase interact with noise. After applying the blind source separation method i.e. Independent Component Analysis, discrete wavelet transform is used to transform the signal into wavelets and apply various shrinkage methods to denoise the signal. The paper shows that out of various soft and hard thresholding methods, particle swarm optimization shows the best performance in low SNR conditions.**

Keywords- Speech Signals, ICA, DWT, Thresholding, SNR

I. INTRODUCTION

Speech signals are the acoustic signals [7] of range 20 Hz to 20 Kilo Hz and their presentation in digital domain brings them as a subject of digital communication. The study of speech signals is indispensible in its part in applications related mainly to identity authentication and animal's research. The noise in speech signal is generally classified in two sources: i.e. noise through a different source or due to the faulty nature of equipments. If the transmission channel or encoding/decoding algorithms are conceived, they would manipulate the nature of signals and hard to recognize. The most effective solution in this aspect lies in the structure of digital communication i.e. retransmission.

Noise is a complimentary signal for any source or system offered by channel. For example stochastic resonance occurred due to non-linear transmission of signal generates noise [1]. An important form of this phenomenon that is frequently considered involves a bistable dynamic system governed by a double-well potential, in which a periodic signal receives assistance from the noise to overcome the potential barrier and switch the system between its two stable states. In direct transmission mode noise is generally studied under two categories: periodic and aperiodic noise. Aperiodic noise accounts for all the general category of noises whose frequency and amplitude cannot be determined for any specific time. In digital signals these noises are termed as white, Gaussian, impulse etc. noises while in domain of speech signals babble noise and acoustic (environment) noise have considerable impact. Additive White Gaussian Noise (AWGN) is a noise channel. This channel effects on the transmitted signals when signals passes through the channel. This noise channel model is good for satellite and deep space communication but not in terrestrial communication because of multipath, terrain blocking and interference. AWGN is used to simulate background noise of channel. The mathematical expression as in received signal $r(t)$ is shown in Figure 1 which passed through the AWGN channel where, $s(t)$ is transmitted signal and $n(t)$ is background noise.

$$r(t) = s(t) + n(t) \tag{1}$$

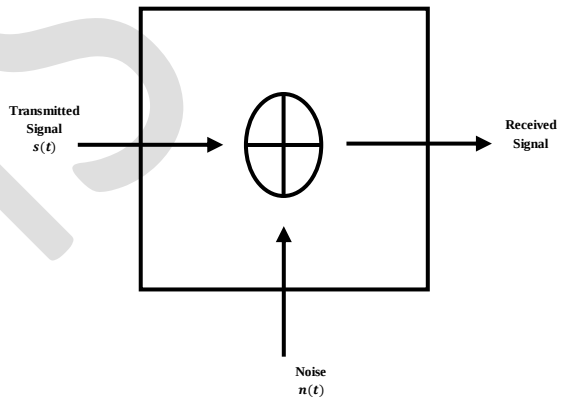

Figure 1: AWGN Channel

Due to wireless nature of transmission, the signals are interacted with channel noise and the destination terminal receives the mixture of both. If the signal is a speech, the audience in direct communication can filter the noise based on his experience of word recognition. But in machine language, suitable filter algorithms are exploited that differentiates the speech and noise based on their statistical properties. The filters for signal filtering are defined in broad categories of active and passive filters. Passive filters are systematic array of components that does not require power source for operation. The capacitors monitor the frequency and resistors define the amplitude or voltage of a passing signal. The efficiency of these filters are limited to the select only the specific band of a signal and cannot differentiate two signals from a same source (if they possess same frequencies). The active filters are rather a smart filter that classifies the nature of signal and can differentiate the required signal from a mixture. The design of active filters has grown to an automated approach powered with computational algorithms simulated on virtual environment

for testing. The algorithms are expected to minimize the effect of noise in a signal and most importantly retain the properties of original signal without considerable loss.

II. INDEPENDENT COMPONENT ANALYSIS

Independent Component Analysis (ICA) is the unsupervised computational and statistical method for discovering intrinsic hidden factors in the data. ICA exploits higher-order statistical dependencies among data and discovers a generative model for the observed multidimensional data. In the ICA model, observed data variables are assumed to be linear mixtures of some unknown independent sources (independent components).A mixing system is also assumed to be unknown. Independent components are assumed to be non-Gaussian and mutually statistically independent.ICA is generally the solution for feature extraction and patterns that represent image, media or time series.

The ICA model assumes that the observed sensory signals x_i are given as the pattern vectors $X = [x_1, x_2, \cdots, x_n]^T \in R^n$. The sample of observed patterns is given as a set of N pattern vectors $T = \{x_1, x_2, \cdots, x_n\}$ that can be represented as a $n \times N$ data set matrix $X = [x_1, x_2, \cdots, x_n] \in R^{n \times N}$ which contains patterns as its columns. The ICA model for the element xi is given as linear mixtures of m source independent variables s_j

$$x_i = \sum_{j=1}^{m} h_{ij} s_j, \quad i = 1, 2, \ldots, n \qquad (2)$$

Where, x_i is observed variable, s_j is the independent component (source signals) and h_{ij} are mixing coefficients. The independent source variables constitute the source vector (source pattern) vectors $s = [s_1, s_2, \cdots, s_n]^T \in R^m$. Hence, the ICA model can be presented in the matrix form

$$\mathrm{x} = Hs \qquad (3)$$

Where, $H \in R^{n \times m}$ is $n \times m$ unknown mixing matrix where row vector $h_i = [h_{i1}, h_{i2}, \cdots, h_{im}]$ represents mixing coefficients for observed signalx_i. Denoting by h_{ci} columns of matrix H we can write

$$\mathrm{x} = \sum_{i=1}^{m} h_{ci}, s_i \qquad (4)$$

The purpose of ICA is to estimate both the mixing matrix H and the sources (independent components) s using sets of observed vectors x. The ICA model for the set of N patterns x, represented as columns in matrix X, can be given as equation 3.

Where $S = [s_1, s_2, \cdots, s_n]$ is the m × N matrix which columns correspond to independent component vectors $s_i = [s_{i1}, s_{i2}, \cdots, s_{im}]^T$ discovered from the observation vector xi. Once the mixing matrix H has been estimated, we can compute its inverse$B = H^{-1}$, and then the independent component for the observation vector x can be computed by

$$s = B\mathrm{x} \qquad (5)$$

The extracted independent components s_i are as independent as possible, evaluated by an information-theoretic cost criterion such as minimum Kulback - Leibler divergence kurtosis, negentropy.

A. Pre-processing

Usually ICA is preceded by pre-processing, including centring and whitening.

B. Centring

Centring of x is the process of subtracting its mean vector $\mu = E\{x\}$ from x:

$$x = x - E\{x\} \qquad (6)$$

C. Whitening (sphering)

The second frequent pre-processing step in ICA is decorrelating (and possibly dimensionality reducing), called whitening. In whitening the sensor signal vector x is transformed using formula

$$y = W\mathrm{x}, \; so \; E\{yy^T\} = I_l, \qquad (7)$$

Where$y \in R^l$, is the $l - dimensional$ $(l \cdot n)$ whitened vector, and W is $l \times n$ whitening matrix. Whitening transforms the observed vector x in a linear manner to generate y (white) and its elements are uncorrelated with unity variance. Whitening allows also dimensionality reduction, by projecting of x onto first l eigenvectors of the covariance matrix of x.

Whitening is usually realized using the Eigen-value decomposition (EVD) of the covariance matrix $E\{yy^T\} \in R^{n \times N}$ of observed vector x

$$R_{\mathrm{xx}} = E\{\mathrm{x}\mathrm{x}^T\} = E_{\mathrm{x}} \Lambda_{\mathrm{x}}^{1/2} \Lambda_{\mathrm{x}}^{1/2} E_{\mathrm{x}}^T \qquad (8)$$

Here, $E\mathrm{x} \in R^{n \times n}$ is the orthogonal matrix of eigenvectors of $R_{\mathrm{xx}} = E\{\mathrm{x}\mathrm{x}^T\}$ and Λ is the diagonal matrix of its eigenvalues

$$\Lambda_{\mathrm{x}} = diag(\lambda_1, \lambda_2, .., \lambda_n) \qquad (9)$$

With positive eigenvalues$\lambda_1 \geq \lambda_2 \geq .. \geq \lambda_n \geq 0$. the whitening matrix can be computed as

$$W = \Lambda_{\mathrm{x}}^{-1/2} E_{\mathrm{x}}^T \qquad (10)$$

And consequently the whitening operation can be realized using formula

$$y = \Lambda_{\mathrm{x}}^{-1/2} E_{\mathrm{x}}^T \mathrm{x} = \mathrm{W}\mathrm{x} \qquad (11)$$

Recalling that,$x = Hs$, we can find from the above equation that

$$y = \Lambda_{\mathrm{x}}^{-1/2} E_{\mathrm{x}}^{T} \quad Hs = H_{\omega} s \tag{12}$$

We can see that whitening transforms the original mixing matrix H into a new one, H_{ω}

$$H_{\omega} = \Lambda_{\mathrm{x}}^{-1/2} E_{\mathrm{x}}^{T} \quad H \tag{13}$$

Whitening makes it possible to reduce the dimensionality of the whitened vector, by projecting observed vector into first $l\,(l \leq n)$ eigenvectors corresponding to first l eigenvalues $\lambda_1, \lambda_2, .., \lambda_l$, of the covariance matrix, E_{x}. Then, the resulting dimension of the matrix W is, $l \times n$ and the size of visible transformed vectors vector y from n to l is shrinking.

Whitening process' output vector can be taken as source to ICA algorithm. The whitened observation vector y is an input to unmixing (separation) operation

$$s = By \tag{14}$$

Where, B stands for original unmixing matrix. An reconstruction of the original observed vector x can be computed as,

$$\tilde{x} = Bs\ , \tag{15}$$

Where, $B = W_{\omega}^{-1}$

For the set of N patterns x forming as columns the matrix X We can provide the following ICA model

$$X = B\,S \tag{16}$$

Where, $S = [s_1, s_2, \cdots, s_n]$ is the $m \times N$ matrix which columns correspond to independent component vectors $s_i = [s_{i1}, s_{i2}, \cdots, s_{im}]^T$ discovered from the observation vector, x_i.Consequently we can find the set S of corresponding independent component vectors as

$$S = B^{-1}X. \tag{17}$$

Application of ICA as the blind source separation model was experimented to separate noise and signal based on their statistical properties of signals [2]. Multiple speech recognition through time-frequency masking can be applied to the ICA outputs to reduce remaining interferences [3] [4]. Same method along with drawbacks of Kurtosis algorithm is present in [5]. The literature on the ICA and wavelet methods exists mainly only in the environment of aperiodic noise. The ICA method only generates the independent components and thus requires an additional filter to denoise the signal. Speech enhancement with missing TF was based on ICA [8] as ICA can be data-driven, adaptive and linear representation in nature. The combination of ICA with wiener filter [9] was used to minimize mean square error, constrained ICA with Bessel Features [10] to extract subsets of desired independent source signals from a set of mixture of source signals. ICA with adaptive filters [11] was used for speech enhancement exploiting the properties of ICA and filtering the components before reconstruction. Liang Hong [12] tested Bayesian algorithm with ICA for maximum a posteriori (MAP) estimator and transformation of learned speech data via ICA. The limitations of linear filters are already mentioned and performance of adaptive filters such as wavelet filters depends on its application. According to Mihov, S. G. et al. [13] the soft thresholding techniques perform better than hard methods. Specific wavelet techniques like HAAR and Daubechies in an experimentation [14] showed mixed response for signals in range 8-20 dB. Adaptive wavelets [15], robust wavelet [16] denoising methods are some optimization techniques experimented in same reference. . In Bayesian denoising method [14] [15], optimal signal estimated for generation of minimum mean square estimators for specific features of speech signal. In addition with wavelet [16] assumes a minimal local regularity translates into constraints on the multifractal spectrum of the signal. The theoretical approach for this is based on Holder exponent [17] in a stochastic frame.

III. DISCRETE WAVELET TRANSFORM

The reconstruction of signal from ICA (in case of mixed signals) is paralyzed with some noise content present as the abrupt peaks in signals. The signals for further filtration are transformed into frequency domain for prediction of noisy peaks and disengaging them from signal. The limitations of Fourier transform and edge of DWT over it is a known fact. DWT decompose a signal into Detailed Coefficients (DC) and Approximate Coefficients (AC) of higher and lower frequencies respectively.

$$S_i = \frac{1}{\sqrt{M}} \sum_n s(t) \varphi_i\,(n) \tag{18}$$

$$S_j = \frac{1}{\sqrt{M}} \sum_n s(t) \psi\,(n) \tag{19}$$

Here, $i \geq j$, $s(t)$ is the output of ICA and $\varphi_i(n), s(t)\psi(n)$ are the functions of discrete variables $n = 0,1, \ldots M-1$. Equation 18 represents the approximate coefficients and equation 19 represents detailed coefficients.

IV. WAVELET THRESHOLDING

The AC of present wavelet is further scaled in 2nd level AC and DC (See figure 2). The removal of noise in DWT is performed by soft thresholding method considering its benefits over hard one [10]. In this method the denoising of a DC is given by [18]:

$$\widehat{w}_{j,k} = \begin{cases} sign(w_{j,k}) * (|w_{j,k}| - \lambda) & w_{j,k} \geq \lambda \\ 0 & w_{j,k} < \lambda \end{cases}$$

$$\lambda_j = \sigma\sqrt{2 \log N} \tag{20}$$

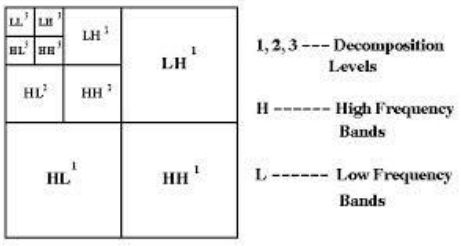

Figure 2: 3rd level of wavelet decomposition

The i^{th} level of DC is an orthonormal function against the noise function having standard deviation σ. N is the frequency values in DC component and this estimation. The models under inspection vary in their thresholding methods and rest of procedure is standard and default. After DWT calculation, the thresholding of wavelet components is performed and inverse wavelet transform is calculated at last.

The benefits of soft thresholding over hard thresholding are well known. They provide near minimax rate for a large range of Besov spaces [6]. The thresholding techniques in their mathematical form are described as follows:

A. VisuShrink

Originally developed by Donoho and Johnstone, the universal thresholding can be applied to a finite length L. VisuShrink thresholding computes the value through noise variance (σ) and number of wavelet coefficients (n). For the complete channel length (L), denoising of coefficients in index interval is given by $[2^L + 1, n]$. The threshold parameter (t) is given by $t = (2\log n)^{\frac{1}{2}}$

B. Sure Shrinkage

Stein's unbiased risk shrinkage or the sure shrinkage is employed to the complete vector of wavelet coefficients. In other words, the thresholding in sure shrinkage is performed on each level of scale. The thresholding computes the right threshold for diverse range of scales. The wavelets are arranged in ascending order with sum given by:

$$x_i = \sum_{r=1}^{n} a_i \quad (21)$$

The risk value is computed on all the coefficients of wavelet that are arranged in decreasing order. For each wavelet coefficient a vector element **c** is assigned to store the local variable values in MATLAB. The risk value for any local coefficient is given by

$$sure_i = \left(\frac{n_j - 2i}{n_j}\right) \quad (22)$$

The wavelet with minimum risk factor is best suited as the threshold value for a coefficient. For a complete range of coefficients, an absolute value is chosen as solution and rest all the threshold values are set to zero. This threshold is called as sure wavelet.

C. PSO based Shrinkage

The efficiency of thresholding depends on number of factors that are decided by the designer. The three main factors that are generally seen in most of researches are:

D. Wavelet choice for segmentation
E. Total number of decompositions
F. And the value of threshold at each individual level

PSO is a standard and widely used algorithm for optimization. The algorithm maintains several threshold values in its search space (length of signal). In every iteration step the candidate threshold is evaluated using optimization function: noise variance and number of coefficients. The candidate solution is matched against the minimum and maximum of objective function. At initial stage the solutions are chosen in random manner but from second stage, the solution is the enhanced filtration of first stage of solutions. The final stage of solution is either the end of iteration cycle or when the output is monotonic in nature.

D. Power Spectral Subtraction

To imply spectral subtraction, let us assume two basic things:

G. Speech magnitude spectrum and noise both are independent in nature against each other.
H. Also, their phase is independent with each other.

The power spectral subtraction for noisy signal corrupted by AWGN is

$$E\left\{\left|Y(e^{j\omega})\right|^2\right\} = E\left\{\left|X(e^{j\omega})\right|^2\right\} + E\left\{\left|N(e^{j\omega})\right|^2\right\} \text{ Or}$$

$$E\left\{\left|X(e^{j\omega})\right|^2\right\} = E\left\{\left|Y(e^{j\omega})\right|^2\right\} - E\left\{\left|N(e^{j\omega})\right|^2\right\} \quad (23)$$

Where,

$$E\{\cos(\Delta\theta)\} = 0$$

V. EXPERIMENTAL SETUP

The utterances were taken from a database composed of 30 speakers, 15 of them are male & other 15 are female speakers. From each speaker we have taken 3 to 4 seconds for training from 24 utterances per speaker with the introduction of session variability. Data recorded with using 16 bit resolution, 8 kHz, 44 kHz, 48 kHz sampling rate.

The training signal was corrupted by additive white Gaussian noise at SNR=5dB and preprocessed with the same speech enhancement method used to enhance the test signal. Simulations are programmed upon MATLAB version R2014a by PC with Intel Core 2 Duo CPU 2.0 GHz.

VI. RESULT

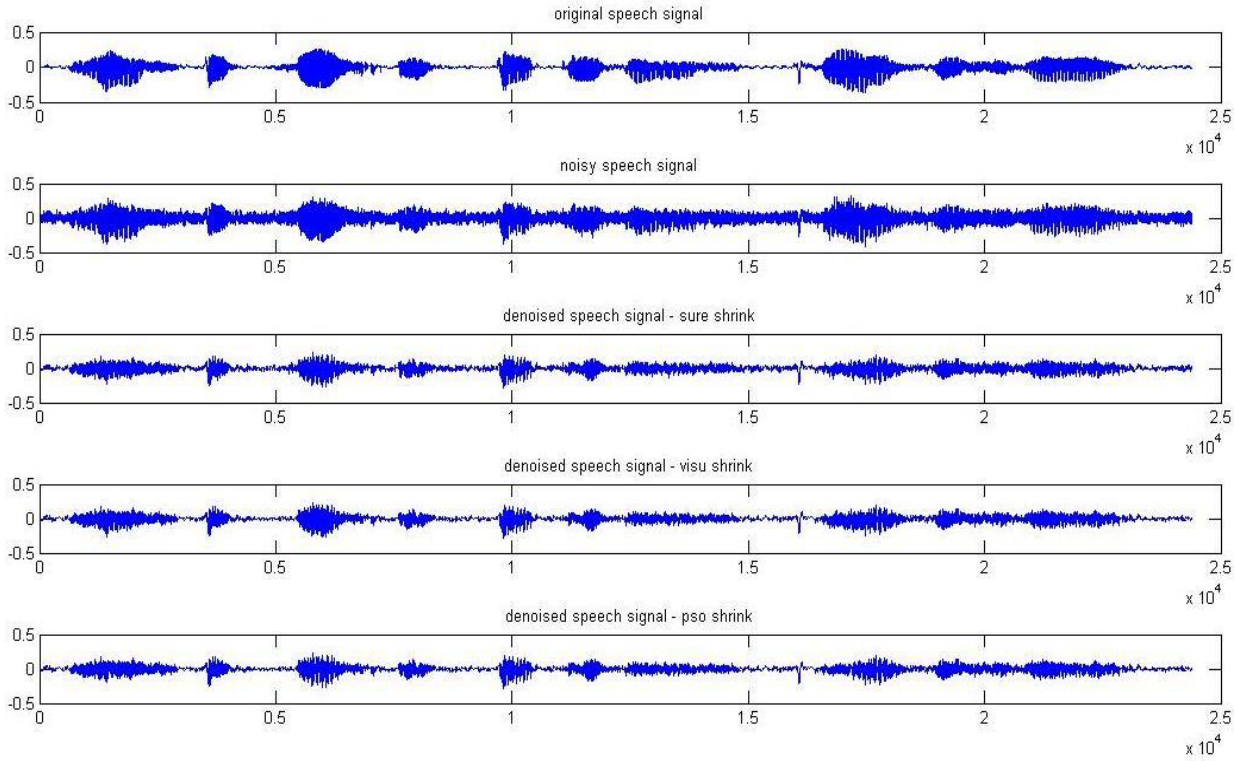

Figure 3: Speech Signals Output Through Different Techniques

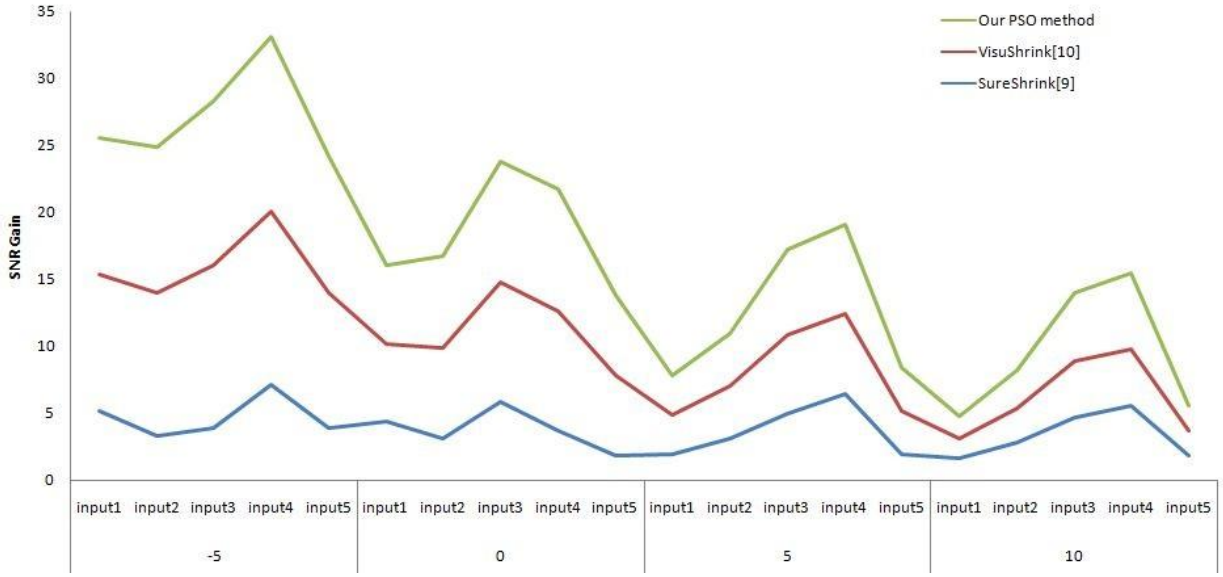

Figure 4: SNR Value Comparison Chart

Table 1: Comparative Analysis of Proposed Technique with Previous Methods

SNR(dB)	Input Speech	SureShrink[19]	VisuShrink[20]	Our PSO method
-5	input1	5.1726	10.1659	10.1669
	input2	3.2802	10.7768	10.7803
	input3	3.8645	12.2031	12.2099
	input4	7.1067	12.9873	12.9993
	input5	3.9241	10.1049	10.1051
0	input1	4.3706	5.8172	5.8391
	input2	3.1449	6.7682	6.7822
	input3	5.8764	8.9009	8.9960
	input4	3.6617	8.9447	9.1005
	input5	1.8163	6.0097	6.0126
5	input1	1.9936	2.8830	2.9692
	input2	3.1312	3.8624	3.9602
	input3	4.9761	5.9129	6.3147
	input4	6.3966	6.0001	6.6854
	input5	1.9579	3.1993	3.2059
10	input1	1.6625	1.4496	1.6637
	input2	2.7910	2.5613	2.8243
	input3	4.6545	4.2031	5.1484
	input4	5.5606	4.2640	5.5878
	input5	1.8360	1.8251	1.8724

A. Test Results for 0 dB (SNR)

Table 2: Output SNR for 0 dB Input SNR

SNRDB Gain VISU	SNRDB Gain SURE	SNRDB Gain PSS	SNRDB Gain PSO
14.3260	15.1237	9.0404	15.4273

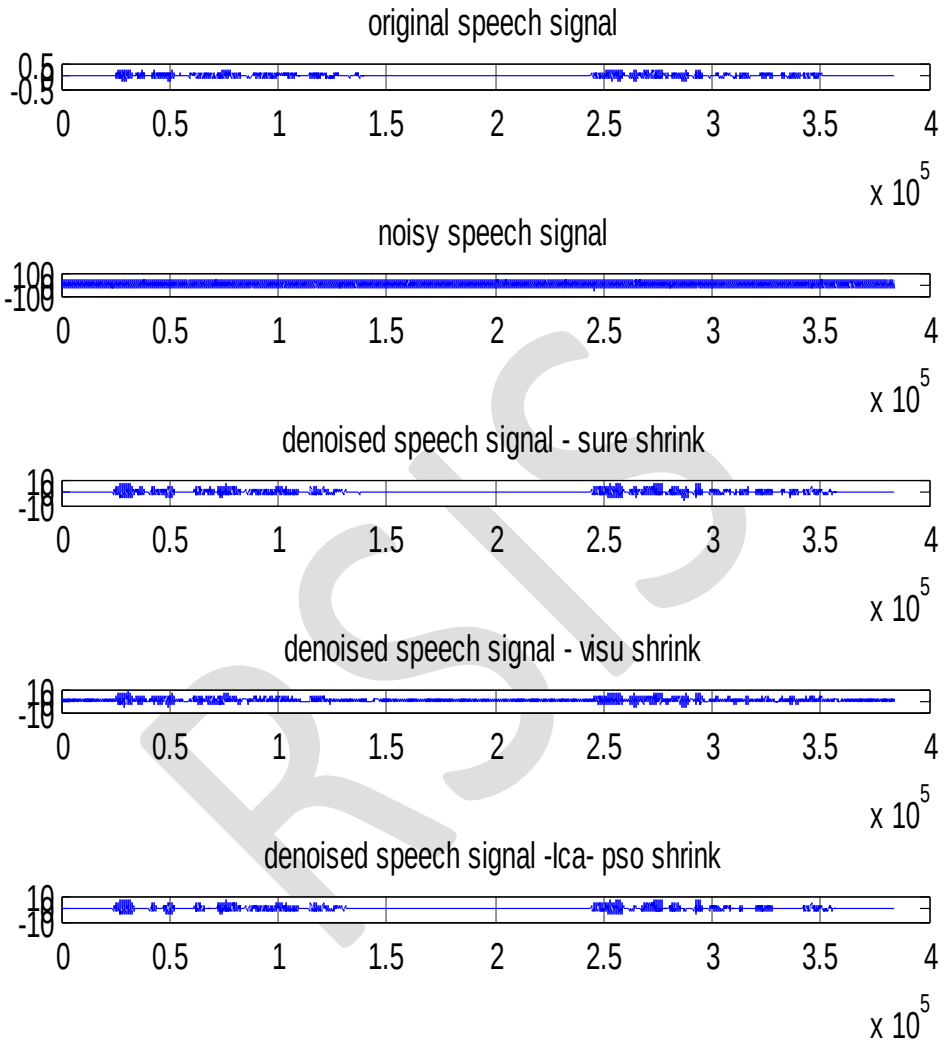

Figure 5: Output SNR for 0 dB

B. Test Results for 5 dB (SNR)

Table 3: Output SNR for 5 dB Input SNR

SNRDB Gain VISU	SNRDB Gain SURE	SNRDB Gain PSS	SNRDB Gain PSO
21.7690	22.5666	16.4832	22.8702

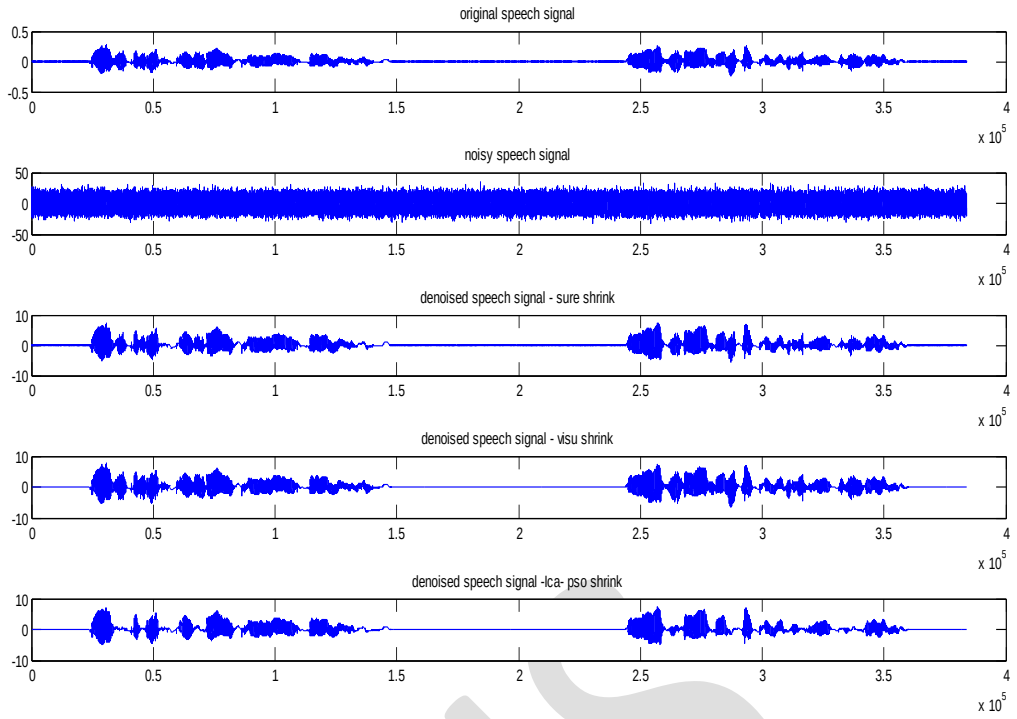

Figure 6: Output SNR for 5 dB

C. Test Results for 10 dB (SNR)

Table 4: Output SNR for 10 dB Input SNR

SNRDB Gain VISU	SNRDB Gain SURE	SNRDB Gain PSS	SNRDB Gain PSO
26.0079	26.7939	20.7108	27.0977

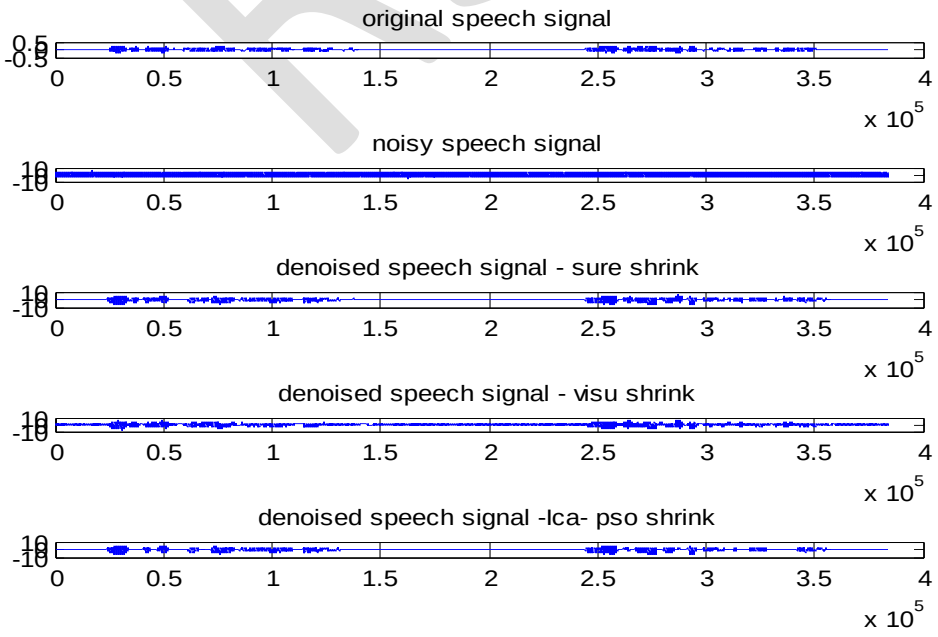

Figure 7: Output SNR for 10 dB

VII. EER CALCULATION

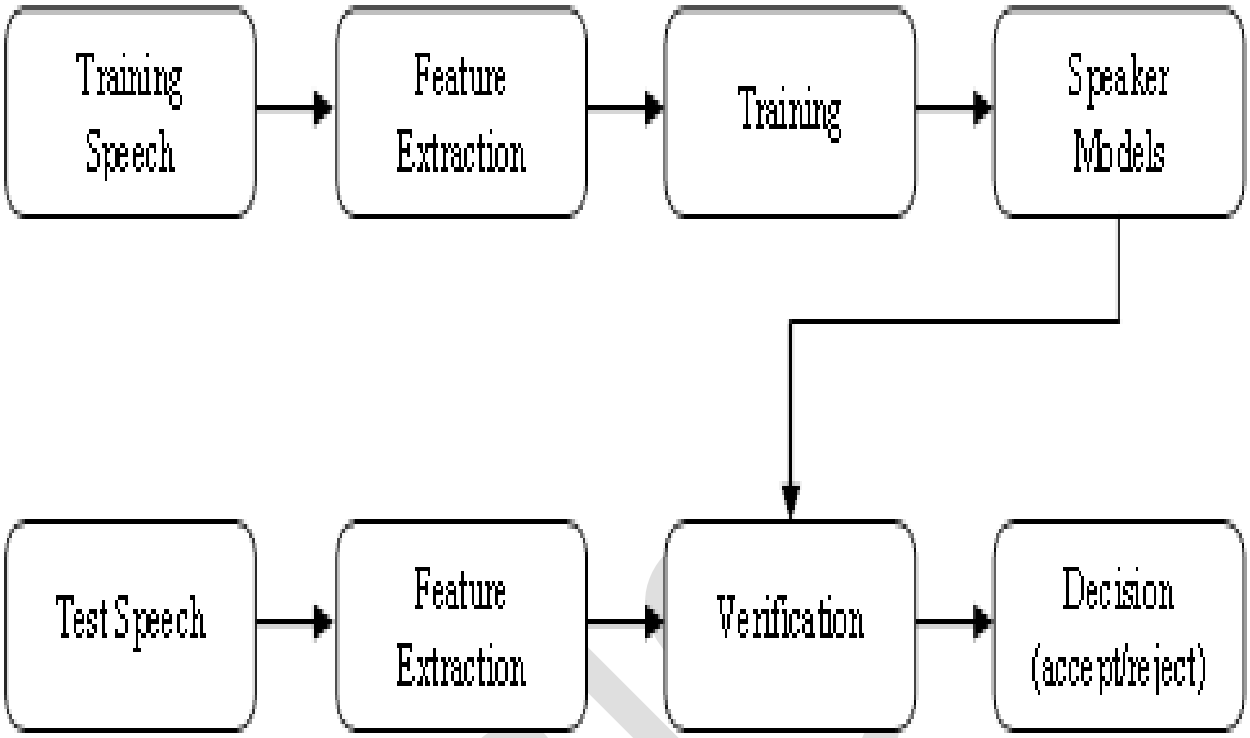

Figure 8: Basic Model for Speaker Verification System

Table 2: Equal Error Rate(EER) Calculation in a Speaker Verification System

	Genuine			**Imposter**			**Total Success Rate/Accuracy %**
	No. of observations	Accepted	Success rate %	No. of observations	Rejected	Success rate %	
Power spectral subtraction(PSS) [11]	2700	1940	71.85	2700	1950	72.22	72.04
SureShrink [9]	2700	2130	78.89	2700	2110	78.15	78.52
VisuShrink [10]	2700	2400	88.89	2700	2000	74.07	81.48
Proposed Method	2700	2150	79.63	2700	2400	88.89	84.26

VIII. CONCLUSION

The paper demonstrates the performance evaluation of various hard and soft thresholding methods. The first stage of filtering is processed through independent component analysis but the signals required a second stage of filtering for a better output. The speech signals when segmented into a granular level of frequency components, the shrinkage methods can be employed to discard the noisy peaks. In comparison of various methods the paper demonstrated that at higher SNR rate all methods are optimal in performance. But, with decrease in SNR ratio the performance of conventional soft thresholding methods does not reach the required mark. The output dB gain of PSO is highest among all the tested algorithms and thus authenticates the feasibility of proposed system. The equal error rate (EER) is calculated from an automated speaker verification system. Here we have considered a text dependent ASV system that employs 13 Mel frequency cepstrum co-efficients [MFCC] features and employs dynamic time warping [DTW] for time alignment [22]. The EER is measured over false acceptance and true rejections. Lower the EER better is the accuracy of the system. In future, the speech may be more corrupt and the high output SNR dB may be in demand. If high computational cost is acceptable, more proficient algorithms such as random forest may be applied.

REFERENCES

[1]. Vaudelle, F., Gazengel, J., Rivoire, G., Godivier, X., & Chapeau-Blondeau, F. (1998). Stochastic resonance and noise-enhanced transmission of spatial signals in optics: The case of scattering. *JOSA B*, *15*(11), 2674-2680.

[2]. Hongyan, L., & Guanglong, R. (2010, September). Blind separation of noisy mixed speech signals based Independent Component Analysis. In *2010 First International Conference on Pervasive Computing, Signal Processing and Applications* (pp. 978-0).

[3]. Hoffmann, E., Kolossa, D., & Orglmeister, R. (2011). Recognition of Multiple Speech Sources using ICA. In *Robust Speech Recognition of Uncertain or Missing Data* (pp. 319-344). Springer Berlin Heidelberg.

[4]. Vivek Anand. Dr. R. S. Anand.Dr. M. L. Dewal. Implementation of blind source separation of speech signals using independent component analysis. *International Journal of Computer Science and Information Technologies*, vol. 2, no. 5, pp. 2147-2151, 2011.

[5]. Karthikeyan, G., & Sahoo, R. (2007). Independent component analysis based blind signal separation for mixed speech signal *(Doctoral dissertation*). 10302066 , 2007.

[6]. Donoho, D. L. (1995). De-noising by soft-thresholding. *Information Theory, IEEE Transactions on*, *41*(3), 613-627.

[7]. Klasmeyer, G., & Sendlmeier, W. F. (2013). The classification of different phonation types in emotional and neutral speech. International Journal of Speech Language and the Law, 4(1), 104-124

[8]. Qiu Wan-hua, The decision-making in management and the application of entropy. Beijing: Mechanism Industry Press, 2002.

[9]. Liang Hong, Justinian Rosca, Radu Balan, Independent component analysis based single channel speech enhancement using wiener filter. Enhancement Using Wiener Filter, ISSPIT, 2003, https://www.cs.rochester.edu/u/rosca/preprints/2003/ISSPIT2003.pdf

[10]. Balakrishna lavu and Venkata Ashokkumar Potnuru, Speech Enhancement using Constrained-ICA with Bessel Features. Project Report, Blekinge Institute of Technology, January 2011, MEE10:114

[11]. Rutkowski, T., Cichocki, A., & Barros, A. K. Speech enhancement using adaptive filters and independent component analysis approach. In International Conference on Artificial Intelligence in Science and Technology, AISAT Vol. 2000, December 2000, pp. 191-196

[12]. Hong, L., Rosca, J., & Balan, R. Bayesian single channel speech enhancement exploiting sparseness in the ICA domain, na, 2004, pp: 1713-1716

[13]. Mihov, S. G., Ivanov, R. M., & Popov, A. N. Denoising Speech Signals by Wavelet Transform. Annual Journal Of Electronics, 2009, 712-715

[14]. Chavan, M. S., Chavan, M. M. N., & Gaikwad, M. S. Studies on implementation of wavelet for denoising speech signal. International Journal of Computer Applications, 3(2), 2010, 1-7

[15]. M. Tomic, Adaptive Wavelet Transforms with Application in Signal Denoising. https://www.fer.unizg.hr/_download/repository/Tomic,KDIclanak.pdf , pp 1-6

[16]. Sardy, S., Tseng, P., & Bruce, A. Robust wavelet denoising. Signal Processing, IEEE Transactions on, 49(6), 2009, 1146-1152

[17]. Attias, H., Deng, L., Acero, A., & Platt, J. C. A new method for speech denoising and robust speech recognition using probabilistic models for clean speech and for noise. In Interspeech, September 2001 (pp. 1903-1906)

[18]. C.A. Medina, a. Alcaim and J.A. Apolinario, Jnr. Wavelet denoising of speech using neural networks for threshold selection. ELECTRONICS LETTERS Vol. 39 No. 25, 11th December 2003

[19]. DONOHO, D.L., and JOHNSTONE, I.M.: 'Threshold selection for wavelet shrinkage of noisy data'. Proc. 16th Annual Conf. of the IEEE Engineering in Medicine and biology society, Maryland, USA, 1994, pp. 24a–25a

[20]. DONOHO, D.L., and JOHNSTONE, I.M.: 'Adapting to unknown smoothness via wavelet shrinkage', J. Am. Stat. Assoc., 1995, 90, (432),pp. 1200–1224.

[21]. BOLL, S.F.: 'Suppression of acoustic noise in speech using spectral subtraction', IEEE Trans. Acoustics Speech Signal Process., 1979, 27,pp. 113–120

[22]. REYNOLDS, D.A., QUARTIERI, T.F., and DUNN, R.B.: 'Speaker verification using adapted Gaussian mixture models', Digit. Signal Process., 2000, 10, pp. 19–41