International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 08, August 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

Speech Signal Enhancement for Babble Noise Using Optimization Algorithm

K N P V R Dinesh Gupta[1], Dr. G. Manmadha Rao[2]

[1]*ECE Department, M.TECH Scholar, DECS, GMRIT, (India)*

[2]*ECE Departments, Professor, GMRIT, (India)*

ABSTRACT

The speech signal enhancement is for obtain clean speech signal from noisy signal environment. For multimodal optimization we better to use natural-inspired algorithms. In this paper we use Firefly algorithmWe compare the firefly algorithm with particle swarm optimization technique (PSO). The proposed algorithm contains three important module techniques. Those are preprocessing module, optimization module and spectral filtering module. The signals are taken from Loizou's database for evaluating proposed technique. In this paper we calculate the perceptional evolution of speech quality (PESQ) and signal to noise (SNR) of the optimized signalsignal. The results of firefly algorithm and PSO are to be compare then we observe that the proposed technique is better than the existing technique.

Key words - speech signal enhancement, Particle Swarm Optimization(PSO), Firefly Algorithm, Perceptional Evolution speech quality(PESQ), Signal to Noise Ratio(SNR).

I. INTRODUCTION

Speech is mostly importantly used for human communication. The goal of speech signal enhancement is to improve the quality of speech is degraded by the noises. Speech enhancement [1] aims to improve the performance of speech communication systems from the noise speech. Mostly speech signal enhancement applications in the areas of speech recognition and speaker identification systems. Speech signal enhancement applied in mobile radio communications, speech to text converting systems, low quality recordings, speech recognition systems, and to improve the performance of hearing. It is a classical problem of signal processing. Speech enhancement is depends on background noise and environmental conditions. If the background noise present in the signal it is very difficult to hearing. Generally we require a signal to noise ratio of about 5-10dB higher than normal hearing listener to achieve the same level understanding the speech signals. Therefore, multi microphones and signals noise reduction strategies have been developed for modern hearing systems. The enhancement of desired speech signal in the presence of stationary noise [11] using an array of microphones has been examined for many years. Algorithms for speech signal enhancement used for different applications like mobile phones, hand free devices etc. mostly used systems for SE to achieve a suppression of disturbing background noise [11]. But do not reduce speech distortion due to room vibrations.

International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 08, August 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

In speech signal enhancement the types of distortions can be divided into two types. Those are 1) the distortions that affect the speech signal itself and 2) the distortions that affect the background noise. By these two distortions, listeners seem to be influenced the most by the speech distortion when making judgment of overall speech quality. The most commonly distortion in speech can be caused by additive noise, which is independent of clean speech. The SE algorithms [6] can be divided into two classes those are, 1) the class based on hidden markov model(HMM) and 2) the classes based on transformation of signals, such as MMSE [12] estimation, spectral subtraction and subspace method.

II.PROPOSED HYBRIDIZATION OF SPECTRAL FILTERING WITH OPTIMAL BINARY MASK TO SPEECH SIGNAL ENHANCEMENT

The speech signal enhancement signal is much required to the degradation of the signals passing through the channel and interferers. In this paper hybridization of the spectral filtering and optimal mask generation is carried out with the aid of minimum mean square error (MMSE) [12] and firefly or PSO. The proposed technique contains three modules those are preprocessing module, optimization module, and optimization module.

A. PREPROCESSING MODULE

In this module, the input signals is prepossessed first we using windowing technique and followed by the Fourier transformation [5]. Initially the signal is split the input signal as overlapping frames, and each frame contains the duration of 0.025ms.

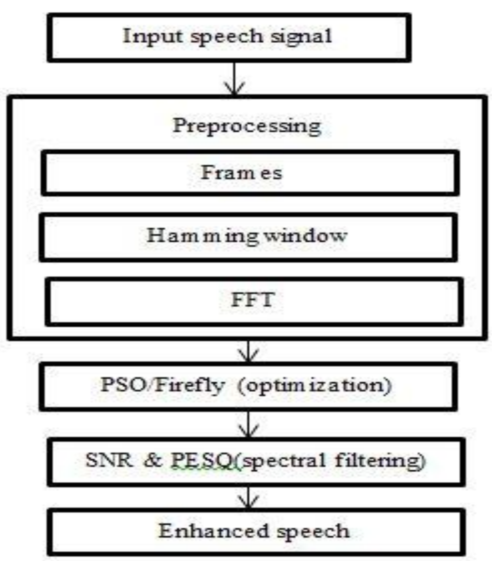

Fig1: Block diagram of speech signal enhancement

The block diagram of preprocessing module is shown in below fig1.

The input speech signal is denoted by S by having a total duration of T ms and the frames be represented by Fi, where $1 \le i \le T/0.025$ each having 0.025 ms. It can be represented by S = {F1 F2……… Fη}, when η=T/0.025 the

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

frames are windowed by using the hamming window technique. By sing the hamming window technique we can eliminate the unwanted signal components and to get sharper peaks. By using the hamming window technique we minimize the maximum side lobes. And the function is defined by:

$$hm(k) = a - b\cos(2\pi k/k\text{-}1) \ldots\ldots(1)$$

Where, a=0.54, b=0.46, K is the width of the samples in the symmetrical window is function and M is integer for 0 < m < M-1. After the windowing technique followed by the Fast Fourier transforms (FFT) this can be used for converting time domain into frequency domain signal. Let the input windowed signals in the i_{th} frame be represented as $w_0^{(i)}, w_1^{(i)}, \ldots \ldots, w_{M-1}^{(i)}$ and Fourier transform is given by:

$$w_k^{(i)} = \sum_{k=0}^{M-1} w_n^{(i)} e^{-i2\pi k\frac{m}{M}}, where k = 0, M-1 ..(2)$$

Here the initial power spectrum is denoted by Λ and is given by $\Lambda = \sum_{k=0}^{M-1} w_j^{(i)}$ is taken mean of the transformed sequence. Then, the noise spectrum is denoted by τ and is given by $= \sum_{j=0}^{M-1} w_j^{(i)}$. The process is carried out for all frames Fi, where $1 \leq i \leq \eta$.

B. Optimization Module

In this module the input noisy speech is divided into noise frames or speech frames using windowing and FFT techniques. This step is performed instead of the normal way of thresholding in MMSE resulting in a hybridized algorithm.

The PSO [2] algorithm is a population based stochastic search algorithm. It provides solutions to the complex nonlinear optimization problems. The main advantage of PSO is it is very cheaper while comparing to other optimization algorithms. In PSO each member of the population is called particle, and each population is called a swarm.

PSO Algorithm Steps:

1. Initially it generates a random population. In this case the initial population consists of value interval [0, 1].
2. Compute the position and velocity of each and every particle.
3. Compute the best velocity for each particle and the best velocity for all particles in the iterations.
4. Update the new velocity, add it to the swarm particle and get the new particle.

$$V_{t+1}=v_t+1/2\alpha v_{t-1}+1/6\alpha(1-\alpha)v_{t-2}+1/24\alpha(1-\alpha)(-\alpha)v_{t-3}+\Psi_1(\rho_b-\rho)+\Psi_2(\omega_b-\rho) \ldots..(3)$$

$$\rho_{t+1} = \rho_t + v_{t+1} \ldots\ldots\ldots\ldots.. (4)$$

5. After updating all the particles, evaluate using fitness function is satisfied, the process ends otherwise the whole process is repeated from step3.

The fitness [1] in this paper depends on three terms. For calculating the fitness in this case, the values are converted to zero or one. It can be represented by z, if z > 0.5 it is converted to 1, otherwise 0. The initial noise power spectrum is denoted by Λ and noise spectrum variance is denoted by is calculated. And spectrum distance is given by:

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

$$SD^{(t)} = 20\log_{10}\Lambda^{t} - log\left\|W_j^{(i)}\right\|, where\ 0 < i \le {}^{n}and\ 0 < j < M - 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots(5)$$

The fitness terms are

Fitness1 = mean spectral distance between signal frames and Λ.

Fitness2 = corr(all frames) / [corr(noise only frames)+corr(signal frames)]

Fitness3 = [no. of noise frames + no. of signal frames] / no. of noise frames

Fitness = Fitness1 * Fitness2 * Fitness3

Firefly Algorithm:

The firefly algorithm [3] was proposed by Dr. Xin She Yang in the year of 2007 at Cambridge University. It was inspired by the flashing behavior of fireflies. The firefly algorithm was very similar to the swarm based algorithms such as Particle Swarm Optimization and Artificial Bee Colony Optimization. This algorithm provided to be much simpler both implementation and concept wise.

The firefly algorithm in this paper can followed three idealized rules. Those are given below

1. All the fireflies are unisex so that one firefly will be attracted to other fireflies regardless of their sex
2. .Attractiveness is proportional to their brightness, if any two fireflies, the less bright one will move towards the brighter one.
3. The landscape of the objective function determines the brightness of the firefly.

The flowchart for Firefly algorithm is shown below Fig2.

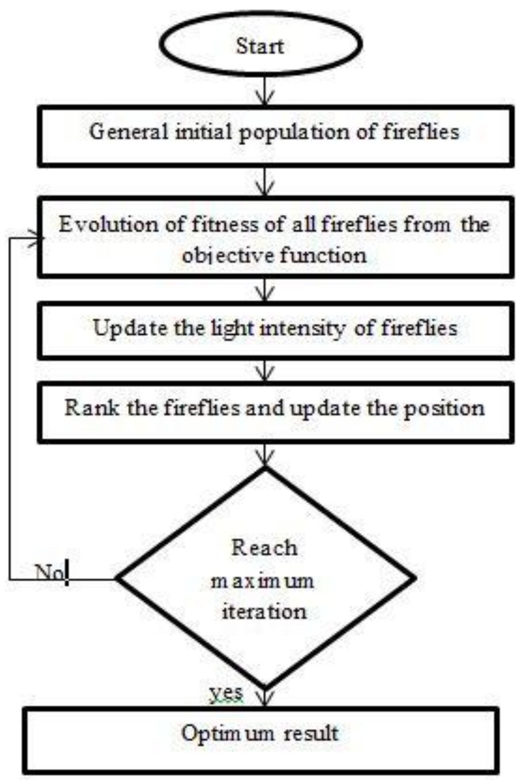

Fig2: Flowchart for Firefly Algorithm

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

Based on these three rules the pseudo code of the Firefly algorithm can be prepared.

Steps for Firefly Algorithm:

1. Define objective function
2. Initialize the population of all fireflies
3. Define light absorption coefficient
4. Compare the brightness of nearest fireflies
5. The less brighter one firefly moves towards brighter firefly
6. If there is any nearest neighbor firefly then it moves randomly
7. Determine the new solution and find light intensity
8. Rank the fireflies according to light intensity and find current best
9. Repeat the steps 4,5,6,7,8.
10. Maximum iterations stop the process.

C. Spectral Filtering Module

The spectral filtering module contains with the minimum mean square error (MMSE) [12] technique in this each of the signal frames is multiplied with gain factor so as to enhance the speech signal. ($w^i = \{w_0^{(i)}, w_1^{(i)}, \dots \dots, w_{M-1}^{(i)}\}$)The algorithm is modified with employment of firefly and PSO for classifying the input noisy speech signal into respective frames as discussed earlier instead of the normal way in MMSE of finding spectral distance and setting a threshold to classify. The block diagram of hybridized algorithm is shown in below figure 3.

$$\Lambda^{(i)} = 9 * \Lambda^{(i)} + \| W_j^{(i)} \| / 10 \ \dots\dots\dots\dots\dots\dots(6)$$

$$\Gamma^{(i)} = 9 * \Gamma^{(i)} + \| W_j^{(i)2} \| /10 \dots\dots\dots\dots\dots\dots\dots(7)$$

The gain factor is found out with the help of apriori SNR and apostiriori SNR.

$$G = \{(c * \sqrt{B})/\gamma_{new}\} * e^{-B/2} *(1 + B) * \text{Bessel}(0,B/2) * \text{Bessel}(1,B/2)\dots\dots\dots\dots\dots(8)$$

$$Xm_j^{(i)} = w_j^{(i)} \times G \dots\dots\dots\dots\dots\dots\dots(9)$$

III. RESULTS AND DISCUSSION

In this section comparative analysis is also made by comparing other existing techniques.

A. Evaluation Metrics

Evaluation metrics consisting perceptual evaluation of speech quantity (PESQ) [1] and signal to noise ratio (SNR). PESQ is a test methodology for automated assessment of the speech quality. It comes under a family of standards for objective voice quality testing. PESQ can be applied to provide end to end quality assessment for a network, or

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

characterized individual network component. The PESQ score is calculated as a linear combination of the average disturbance value and the average asymmetrical disturbance value is given by the equation below:

$$PESQ = b_0 + b_1 D_{avg} + b_2 A_{avg} \text{ where } b_0 = 4.50,\ b_1 = -01,\ b_2 = -0.0309 \quad \ldots\ldots\ldots\ldots\ldots(9)$$

Signal to noise ratio compares the level of desired signal to the level of background noise [11]. It is defined as the ratio of signal power to the noise power. And it can be expressed in decibels.

B. Simulation Plots

In this section the simulation plots obtained for the proposed technique is given. The figures include original signals, noise signals and enhanced signals estimates from various techniques. The figures given the simulation results for the different types of noise such ass car noise, exhibition noise, restaurant noise, babble noise, street noise and train noise. The plots are taken having the noise level at 0dB. From the simulation results, we can observe that our proposed technique that is firefly algorithm has better noise suppression and speech enhancement.

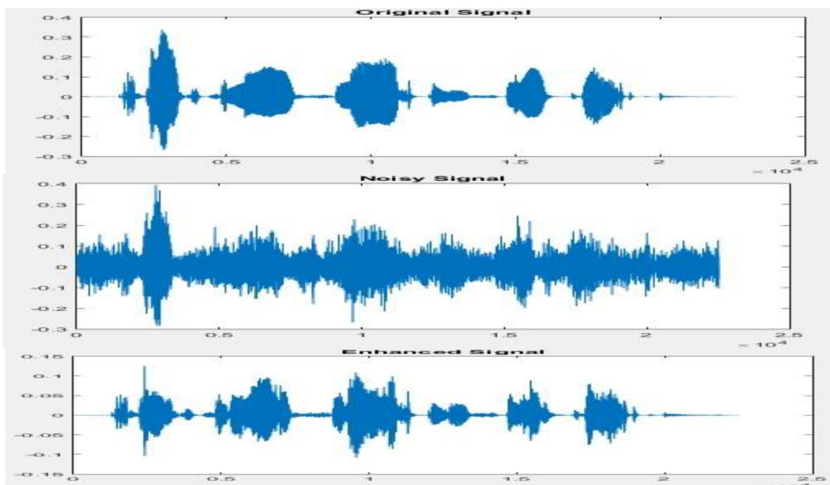

Fig3: Simulation plots for babble noise using PSO

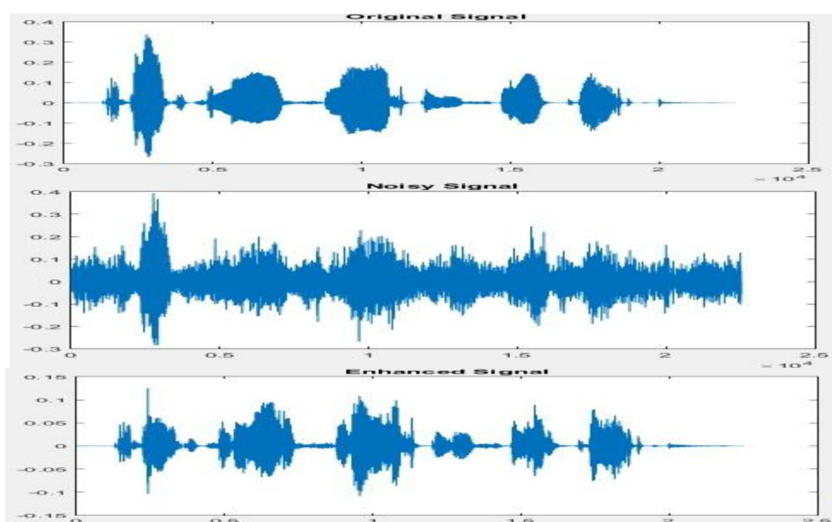

Fig4: Simulation plots for babble noise using Firefly Algorithm

D. Detailed Analysis

Detailed analysis is carried out using the calculation of PESQ and SNR. This analysis is carried out for a noise level at 0dB. The noise is the babble noise.

	PESQ	SNR
PSO	1.849	32.9673
Firefly	2.014	36.1421

The experimental results for PSO for babble noise the SNR is 32.9673 and PESQ is 1.849. The experimental results for Firefly algorithm for babble noise the SNR is 36.1421 and PESQ is 2.014.

IV. CONCLUSION

In this paper hybridization of spectral filtering and optimization algorithm is carried out containing MMSE and firefly for effective speech enhancement. The proposed technique consists of three modules those are preprocessing module, optimization module, and spectral filtering module. lozus's database and auarora database are used for evaluating the proposed technique using standard evaluation consists of PESQ and SNR. Hence we can observe that proposed technique that is firefly optimization algorithm better evaluation that he existing technique that is PSO.

V. REFERENCES

[1]. R. senthamizhSelvi, G.R. Suresh, "Hybridization of spectral filtering with particle swarm optimization for speech signal enhancement", International Journal of Speech Technology, springer science and business media New York, Vol 19, Issue 1, pp 19–31, Mar 2016.

[2] K. Prajna · G. S. B. Rao, K. V. V. S. Reddy, R. Uma Maheswari, "A new dual channeapproach to speech enhancement based on Accelerated Particle Swarm Optimization (APSO)", International Journal of Speech Technology, Vol 17, Issue 4, pp 31-351, Dec 2014.

[3] AdilHashmi, NishantGoel, ShrutiGoel, Divya Gupta, Firefly Algorithm for Unconstained Optimization, IOSR Journal of Computer Engineering (IOSR-JCE), Vol 11, Issue 1, pp 75-78, June 2013

[4] LalehBadriAsl and VahidMajid" Speech Enhancement Using Particle Swarm Optimization Technique", International Conference On Measuring Technology And Automation 2010

[5] Ghasemi, J., &Mollaei, M. R. K. A new approach for speech enhancement based on eigen value spectral subtraction. Signal Processing, 3(4), pp 34–41, 2009

[6] Ephraim, Y., & Van Trees, H. L. . A signal subspace approach for speech enhancement. IEEE Transaction Speech and Audio Processing, 3(4), pp 251–266, 1995

[7] Choi J.H., & Chang J.H, using acoustic environment classification for statistical model based speech enhancement. Speech communication, vol 54, pp 477-490, 2012

[8] X. S. Yang "Firefly algorithm for multimodal optimization, stochastic Algorithms", Foundations and Applications (SAGA 2009, Lecture notes in computer science, vol 5792), pp-169-178, 2009

[9] E. Bonabeau, M. Dorigo, G. Theraulaz, Swarm Intelligence: From Natural to Artificial Systems (Santa Fe Institute Studies in the Sciences of Complexity, NY: Oxford University Press, 1999).

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

[10] Pei-Wei TSai, Jeng-Shyang Pan, Bin-Yih Liao, Shu-Chuan Chu, Enhanced Artificial Bee Colony Optimization, International Journal of Innovative Computing, Information and Control, Volume 5, Number 12, December 2009.

[11] BOLL S F. Suppression of acoustic noise in speech using spectral subtraction[J]. IEEE Trans. Acoustics, Speech, Signal Processing, 1979,27(2):113-120.

[12] Ephraim Y, Malah D. Speech enhancement using a minimum meansquare error short time spectral amplitude estimator. IEEE Transactions on Acoustics, Speech, Signal Processing, 1984, 32(6): 1109-1121.