SPEECH SIGNAL ENHANCEMENT USING FIREFLY OPTIMIZATION ALGORITHM

G. Manmadha Rao and K N P V R Dinesh Gupta
Department of Electronics and Communication Engineering,
GMR Institute of Technology, Rajam, A.P, INDIA

ABSTRACT
The speech signal enhancement is essential to obtain clean speech signal from noisy signal. For multimodal optimization, the natural-inspired algorithms such as Firefly Algorithm (FA) are better. The proposed algorithm contains preprocessing module, optimization module and spectral filtering module. Here, Loizou’s and Aurora databases are considered for signals. In this paper the Perceptional Evolution of Speech Quality (PESQ) and Signal-to-Noise ratio (SNR) of the enhanced signal are calculated to evaluate the performance of Firefly Algorithm.

Key words: Multi model optimization, natural inspired algorithms, SNR, PSO, Firefly Algorithm, and Perceptional Evolution Speech Quality.

http://www.iaeme.com/IJMET/issues.asp?JType=IJMET&VType=8&IType=10

1. INTRODUCTION
Speech is most importantly used for human communication. The main objective of speech signal enhancement is to ameliorate the quality of speech when is disgrace by the noises. Speech enhancement [1] focused on improvement 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 [8] is used in mobile communications, Speech to text translating systems, less quality recordings, speech recognition systems, and to improve the performance of listening. It is a simple problem of signal processing. Speech enhancement is basing on background noise and environmental state. If the background noise present in the signal it is very difficult to hear. Generally a signal-to-noise ratio of about 0-10dB higher than normal hearing listener is required to obtain the same level of understanding the speech signals. Therefore, multi microphone and signals noise trimming strategies have been developed for advanced listing systems. The enhancement of original speech signal [3] in the presence of stationary noise using an array of microphones has been examined for several years. Algorithms for speech signal enhancement in multiple applications like hand free devices, mobile phones, etc are mostly used for speech enhancement for suppress of background noise. In the presence of room vibrations the speech distortion cannot be reduced.

http://www.iaeme.com/IJMET/index.asp 120 editor@iaeme.com
In speech signal enhancement the types of distortions can be divided into two types. Those are 1) The speech signal affected by itself due to the distortion and 2) due to the background noise the distortion can be effected. By these two distortions, listeners are getting to be effected the most by speech distortion when making judgment of overall speech quality. The most commonly distortion in speech is caused by additive noise, and it is not depends on clean speech. The Speech Enhancement algorithms [7] are mainly classified as, 1) Hidden Markov Model (HMM) and 2) transformation of signals, that is MMSE [15]-[18] estimation, spectral subtraction [7]-[14] and subspace based methods etc.

So many different noise reduction methods proposed previously. Existing approaches contains advanced methods such as kalman filtering [6], spectral subtraction [13], and Ephraim mullah filtering techniques. For the coefficient thresholding approach wavelet based techniques are used for speech signal enhancement. The alternative of traditional optimization techniques are firefly optimization algorithm [11] and particle swarm optimization (PSO) techniques.

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

The speech signal enhancement signal is mostly required as the signals are degraded when passing through the medium and interferes. In this paper optimal mask generation and hybridization of the spectral filtering [6] is carried out with the aid of Minimum Mean Square Error (MMSE) [15] firefly [4] and PSO [5]. The proposed technique contains three modules those are pre-processing module, mask generation module, and spectral filtering module.

2.1. Pre-Processing Module

In the preprocessing module, the input signals is prepossessed first, here Hamming windowing technique is used and followed by the FFT [5]. Initially the input signal is spitted as overlapping frames, and each frame contains the duration of 0.025ms. The block diagram of preprocessing module is as shown in Fig1.

\[
hm(k) = x - y \cos(2\pi k/k-1)
\]
Where, \( a=0.54, \ b=0.46, \ K \) is the width of the samples in the identical windowing function and \( M \) is integer for \( 0 < m < M-1 \). After the windowing technique followed by the Fast Fourier transforms (FFT) is to obtain time domain to frequency domain signal. Let the input filtering signals in the \( i_{th} \) frame be represented as \( w^{(i)}_0, w^{(i)}_1, \ldots, w^{(i)}_{M-1} \) and Fourier transform is showed in below equation 2.

\[
w^{(i)}_k = \sum_{n=0}^{M-1} w^{(i)}_n e^{-i2\pi nm/M}, \text{where } k = 0, M - 1
\]  

(2)

Here the at the start power spectrum is represented by \( \Lambda \) and is given by \( \Lambda = \sum_{k=0}^{M-1} w^{(i)}_j \) is taken mean of the transformed sequence. Then, the noise power spectrum is represented by \( \tau \) and is obtained \( \sum_{j=0}^{M-1} w^{(i)}_j \). The process is continued for all frames \( F_i \), where \( 1 \leq i \leq n \).

### 2.2. Optimization Module

In the optimization module the resource noise speech [17] is divided into noise frames or speech signal frames. Here, Particle Swarm Optimization (PSO) is considered.

For the population based stochastic search algorithm we considered Swarm based Optimization [2] algorithm and it is best for search space algorithm. It provides results to the complicated nonlinear optimization troubles. The main advantage of PSO is very cheaper and simpler compare to other optimization algorithms [12]. In Particle swarm optimization each population is called a swarm and each member of the population is called particle.

PSO algorithm steps:

- Initially it generates a random population. In this case the initial population consists of value interval \([0, 1]\).
- For each and every particle we measure the position and velocity.
- After the measurement of position and velocity of particles we identify the best position and best velocity. This process is repeating for all iterations.
- Upgrade the current velocity, and it is add it to the swarm particle and get the modern particle.

\[
V_{t+1} = v_t + 1/2\alpha v_{t} + 1/6\alpha(1-\alpha)v_{t-2} + 1/24\alpha(1-\alpha)(-\alpha)v_{t-3} + \psi_2(\rho_b - \rho) + \psi_2(\omega_b - \rho)
\]  

(3)

\[
\rho_{t+1} = \rho_t + v_{t+1}
\]  

(4)

After the entire particle updated, assess using fitness function. If the fitness function is contented, the process stop otherwise the entire process is go over again from step3.

The fitness [1] in this paper depends on three terms. For measuring the fitness in this case, the values are changed to zero or one. It can be denoted by \( z \), if \( z > 0.5 \) it is changed to 1, otherwise 0. The initial noise power spectrum is represented by \( \Lambda \) and noise spectrum variance is represented by spectrum distance can be calculated using equation 5.

\[
SD^{(t)} = 20 \log_{10} \Lambda^t - log \left\| W^{(i)}_j \right\|, \text{where } 0 < i \leq n \text{ and } 0 < j < M - 1
\]  

(5)

The fitness terms are

- Fitness1 = mean spectral distance between signal frames and \( \Lambda \).
- Fitness2 = corr(all frames) / [corr(noise frames)+corr(signal frames)]
- Fitness3 = [no. of noisy frames + no. of signaling frames] / no. of noisy frames
- Fitness = Fitness1 * Fitness2 * Fitness3
2.3. Firefly Algorithm

Dr. Xin She Yang was introduced Firefly Algorithm [4] in the year of 2007 at Cambridge University. The Firefly Algorithm was based on the flashing behavior of the fireflies. The swarm based algorithms such as PSO and Artificial Bee Colony Optimization [13] are very similar to firefly algorithm. This algorithm is much simpler in both implementation and concept wise.

In this paper, the firefly algorithm follows three unique rules. Those are

- All the fireflies must be same sex so the all fireflies will not dazzle to other fireflies regardless of their sex.
- The attractiveness is proportional to their brightness, in any two fireflies in the population, the brighter firefly will attract to the lesser brighter firefly.
- The shining of the Firefly is represented by the landscape of the objective function.

The flow of Firefly optimization is as shown the Fig 2.

![Firefly Algorithm](image)

**Figure 2 Firefly Algorithm**

Firefly algorithm pseudo code can be prepared by depending on the three unique rules.

**Pseudo code for FA:**

- Let \( f(x) \) be the Objective function, here \( x=(x_1,\ldots,x_d) \)
- All the fireflies initial population is created;
- light intensity \( I \) value is calculated;
- Light absorption coefficient \( \gamma \) is measured;
- While \( t<\text{Max Generation} \)
- For \( i = 1 \) to \( n \) (all \( n \) fireflies);
For j=1 to n (all n fireflies)
   If (Ij > Ii), move firefly i towards j;
   end if
   to calculate new solutions and note down the new light intensity values;
   End for j;
   End for i;
   Rank all the fireflies and find the best one;
   End while;
   Post process results and visualization;
   End procedure;

2.4. Spectral filtering (SF) Module
The spectral filtering [16] module contains the MMSE [18] technique. In this each of the signal frames is multiplied with gain factor \( w^j = \{w_0^{(i)}, w_1^{(i)}, \ldots, w_{M-1}^{(i)}\} \) to enhance the speech signal. The algorithm is changed with employment of firefly and Particle Swarm Optimization [2] for division of the input noise speech signal into respective different frames having the time duration 0.025ms as discussed earlier instead of the normal way in MMSE [18] for determine spectral distance and getting a threshold.

\[
\Lambda^{(i)} = 9 * \Lambda^{(i)} + \| W_j^{(i)} \| / 10 \tag{6}
\]

\[
\Gamma^{(i)} = 9 * \Gamma^{(i)} + \| W_j^{(i)} \|^2 / 10 \tag{7}
\]

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

\[
G = \left( \frac{c \sqrt{B}}{\gamma_{new}} \right) e^{-B/2} \ast (1 + B) \ast \text{Bessel}(0,B/2) \ast \text{Bessel}(1,B/2) \tag{8}
\]

\[
Xm_j^{(i)} = w_j^{(i)} \ast G \tag{9}
\]

3. RESULTS AND DISCUSSION
In this section the quality performance of the signal in different noisy environments is evaluated. The simulated plots of signal in different noisy environment with the help of Firefly Optimization and Particle Swarm Optimization are observed.

3.1. Experimental set up and Database Information
In this we considered the signal and noises of Loizou’s database [19] for experimentation. The database was introduced to ease assessment of speech improvement techniques. The noise signal can be taken taken from the AURORA database and comparing train noise, babble, car, exhibition hall, restaurant, and street noises.

3.2. Evaluation Metrics
Evaluation contains PESQ [1] and SNR. PESQ is a testing technique for automatic measurement of the speech quality. The PESQ is comes under a group of standards for objective voice signal quality testing. PESQ can be applied to provide end to end quality test measurement for a system, or characterized single system component. The Perceptual Evaluation Speech Quality score is calculated as a linear combination of the average
disturbance value \( (D_{avg}) \) and the average asymmetrical disturbance value \( (A_{avg}) \) is given in equation 9.

\[
PESQ = b_0 + b_1 D_{avg} + b_2 A_{avg} \text{ where } b_0 = 4.50, b_1 = -0.01, b_2 = -0.0309 \tag{9}
\]

Signal to Noise Ratio (SNR) compares the level of desired signal and level of background noise in desired signal [11]. The signal to noise ratio is defined as the ratio of signal power to the noise power.

### 3.3. Simulation Results

In this section, original signals from Loizus database and noises from AURORA database are considered and evaluated using PSO and Firefly Algorithms. The respective simulated results of original signal and different noises like car noise, exhibition noise, restaurant noise, babble noise, street noise and train noise are illustrated in Fig 3 to Fig 16. By observing the simulation results, the proposed Firefly Algorithm is found to be superior as it gives better speech enhancement [8]-[9] and noise suppression.
Figure 7 Simulation plots for restaurant noise using PSO

Figure 8 Simulation plots for airport noise using PSO

Figure 9 Simulation plots for exhibition noise using PSO

Figure 10 Simulation plots for babble noise using Firefly Algorithm

Figure 11 Simulation plots for car noise using Firefly Algorithm

Figure 12 Simulation plots for restaurant noise using Firefly Algorithm
3.4. Detailed Analysis
In this section, PESQ and SNR performance measures are calculated. This analysis is carried out for a noise level of 0dB and the noises considered are Babble noise, Train noise, Car noise, Street noise, Restaurant noise, Exhibition noise and Airport noise.

Table 1 comparison results for PSO & Firefly Algorithm

<table>
<thead>
<tr>
<th>Noises</th>
<th>PSO SNR</th>
<th>PSO PESQ</th>
<th>Firefly SNR</th>
<th>Firefly PESQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babble</td>
<td>33.0446</td>
<td>1.9447</td>
<td>35.8635</td>
<td>2.2938</td>
</tr>
<tr>
<td>Train</td>
<td>33.1637</td>
<td>1.8584</td>
<td>35.9079</td>
<td>2.6522</td>
</tr>
<tr>
<td>Car</td>
<td>33.1536</td>
<td>1.9093</td>
<td>35.1627</td>
<td>2.6275</td>
</tr>
<tr>
<td>Restaurant</td>
<td>32.9157</td>
<td>1.8994</td>
<td>35.2689</td>
<td>2.3845</td>
</tr>
<tr>
<td>Exhibition</td>
<td>32.9834</td>
<td>1.8919</td>
<td>35.7853</td>
<td>2.4112</td>
</tr>
<tr>
<td>Airport</td>
<td>33.0267</td>
<td>1.8837</td>
<td>36.30616</td>
<td>2.1202</td>
</tr>
<tr>
<td>Street</td>
<td>32.8604</td>
<td>1.9061</td>
<td>35.8294</td>
<td>2.5216</td>
</tr>
</tbody>
</table>
The Experimental results/ performance measures, SNR and PESQ for signal with Babble noise using PSO are 33.046 and 1.9447 respectively. Whereas with Firefly Algorithm, the performance measures are SNR, 35.8635 and PESQ, 2.2938, which are found to be better in all the noisy environments compared to PSO.

4. CONCLUSIONS
In this paper, hybridization of spectral filtering and optimization algorithm is carried out for effective speech enhancement. The signal with different noises is processed using PSO & Firefly Algorithms. The Performance measures Perceptual Evaluation of Speech Quality (PESQ) and Signal to Noise Ratio (SNR) are calculated and Firefly Optimization Algorithm found to be superior as it gives better results than Particle Swarm Optimization (PSO) in all noisy environments.

REFERENCES

http://www.iaeme.com/IJMET/index.asp editor@iaeme.com
Speech Signal Enhancement Using Firefly Optimization Algorithm


