Improving Speech Signal Intelligibility by Optimal Computation using Single-Channel Adaptive Filtering

Ohidujjaman
Lecturer
Department of Computer Science and Engineering
Dhaka International University
Dhaka, Bangladesh

Mahmudul Hasan
Assistant Professor
Department of Computer Science and Engineering
Comilla University
Comilla, Bangladesh

Mohammad Nurul Huda
Professor
Department of Computer Science and Engineering
United International University
Dhaka, Bangladesh

ABSTRACT
Numerous environmental sources of noise and distortion can degrade the quality of the speech signal in a communication system. This study explores the effects of these intrusive sounds on speech applications, introduces some techniques for reducing the influence of noise and enhances the acceptability and intelligibility of the speech signal. In our research, a noise reduction system incorporates a single microphone method in time domain to improve SNRs (signal to noise ratios) of noise contaminated speech.

Previously, noise reduction techniques estimate noise from the valley of the spectrum based on the harmonic properties of noisy speech, called minimum value sequences (MVS). Since the valleys of spectrum are inadequate to estimate noise reliably, we propose the estimated degree of noise (EDON) [1], [2] to adjust the amplitudes of the MVS. The salient features of the proposed method are a single-channel adaptive filter to reduce computational time and cost for optimal noise reduction, and to estimate noise continuously on a frame-by-frame basis without the aid of voice activity detector (VAD) [2]. For optimal noise reduction with a fewer number of iterations, an equation is derived from set values of SNRs and EDONs. To derive the proposed iteration number equation, we use the third degree parabola equation and least squares solution for the coefficients of EDON.

General Terms
Speech Signal Processing

Keywords
Adaptive filter; degree of noise; enhancement; iteration number; noise; speech signal; signal to noise ratio

1. INTRODUCTION
In most speech communication scenery, the presences of the background intrusions degrade the quality or intelligibility of speech signal. Noisy environment also reduces listeners' capability to understand what is said? ; In addition to interpersonal communication, speech can also be transmitted across telephone channels, loudspeakers or headphones. The quality of speech can also be influenced in data conversion (microphone), transmission (noisy data channels) or reproduction (louder speaker and headphones). The purpose of many enhancement methods is to reduce background noise, improve speech quality, or suppress channel or speaker interference. The goal of our thesis is to improve overall speech quality, hence to increase the intelligibility. We also concentrate on reducing back-ground noise to improve speech quality. The enhancement of superiority will focus on the speech that has additive white noise.

Enhancement algorithms can also be partitioned depending on whether a single-channel adaptive filter or dual-channel (or multi-channel) approach is used. For single-channel adaptive filter applications, only a single microphone is available. The applications in this situation are voice telephone or radio communications. In dual-channel algorithms, the acoustic sound waves arrive at each sensor at slightly different times (one is normally a delayed version of the other). In our study we deal with the first one and assume that: (1) Noise is white, (2) Noise distortion is additive, (3) Noise and speech signals are uncorrelated, and (4) Only one input channel is available (except for adaptive noise cancellation).

1.1 A Single Channel Noise Reduction Model
Most of the time, speech communication takes place in a noisy environment. In modern hands free speech communication environments often occurs the situation that the speech signal is superposed by background noise, Fig.1. This is particular the case if the speaker is not located as close as possible to the microphone. The speech signal intensity is decreased with growing distance to the microphone. It is even possible that background noise sources are captured at a higher level than the speech signal. The noise distorts the speech and words are hardly intelligible. One of the most important measures of speech intelligibility is the signal-to-noise ratio (SNR), which is the ratio between the power of the speech and the power of the noise expressed in decibels (dB). For instance, at classroom environment, the SNR ranges from +5dB to -7dB, and in cocktail parties the average SNR ranges from +1dB to -2dB [3], [7].

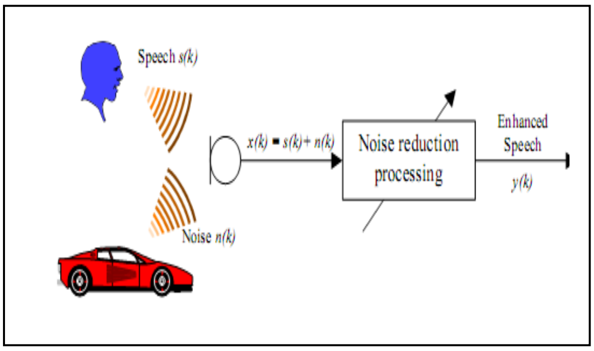


Fig 1: Schematic diagram of single-channel mixing of signals and the noise reduction system

1.2 Problems Statement

The most common problem in speech processing is the effect of interference noise in speech signals. Interference noise masks the speech signal and reduces its intelligibility. Intrusion sound comes from acoustical resources such as traffic, crowd, ventilation equipment, and usually echoes and reverberation. It can also be occurred electrically from distortion or tape hiss products and thermal noise.

In real world environment the presence of noise drastically degrades the performance of numerous speech applications that causes inaccurate information exchange.

Regardless of effort by the researchers, the task of noise reduction is incomplete, since few methods have determined simultaneously the following problems: (1) need to improve the SNR of noisy speech without distorted the speech content using single-channel method, (2) the need for second microphone to receive reference noise [8], (3) the need for detection of 'no speech region' in a noisy speech signal to estimate noise [9], (4) the failure to enhance unvoiced speech [3], (5) the inability to reduce dynamic (no stationary) interfering noise, and (6) need to diminish expensive computational time.

In this study, we present algorithms which attempt to resolve some of the above problems and diminish the effects of the remaining problems, and as well as find out the minimum iteration number to reduce maximum noise.

1.3 Objectives and Contributions

All speech and noise sources are acoustically combined (mixed to one single-channel) to simulate a real-noise environment. They reside in the same frequency band and may have similar correlation properties. Although multi-channel based noise reduction system provides a better performance than single-channel case **[6]**, but single-channel system is still an important issue when a single-microphone speech is only the available source. In our study we emphasis the following objectives:(1) study a single-channel noise reduction method of speech signal in a time domain, (2) observe the improvement of SNR of noise infected speech, (3) observe the speech quality and intelligibility, (4) study the effect of block length processing of adaptive filter and iteration number of filter update and, (5) study the proposed iteration number based on the estimated degree of noise, then investigate the iteration number in various frame.

We update the performance of combination the blind source separation (BSS) and weighted noise subtraction (WNS) method and propose single channel adaptive filter model. In the planned method, the estimated first noise is carefully weighted which effects low musical noise and high noise reduction score is obtained at the output. The filter coefficients are estimated by the least mean square (LMS) algorithm. The LMS algorithm is based on the steepest descent method and, is implemented in adaptive filtering. To acquire superior junction properties, we enlarge the data length from the preceding frames. A repetitive gradient-based hunt is used to acquire the optimal coefficients. At every time, an adaptation algorithm that fits the coefficients of the adaptive filter is calculated to minimize the residual. We notice that with increasing iteration, the computational time is increased. Again, for continuous speech sequences, excess numbers of iterations regenerate noise in the silent part and deteriorate the sound. Therefore, an effort is prepared to control the iteration, and we propose optimal iteration number for estimation process. We derive a function utilizing the Least-Squares (LS) method from the EDON **[1]** in each input SNR and the minimal numbers of iterations are required to obtain maximum noise reduction. Then, we propose a cascade processing of single channel adaptive filter. By adopting the proposed iteration number, the planned method turns into comparatively economical and the signal renewal trouble is reduced in the silent parts. In our study, we consider the signal and noise are uncorrelated. The main advantage of this approach is valid for any second- order signal processing, and hence, it is unnecessary to assume additional conditions on the signal such as state-space model, stationary. Moreover, using the second order statistics (decor relation), it is possible to separate dynamic sources **[12]**. Compare to other systems, the improved performance of the noise reduction system is achieved by the single-channel adaptive filtering. The block diagram of the proposed noise reduction system is shown in Fig. 2.

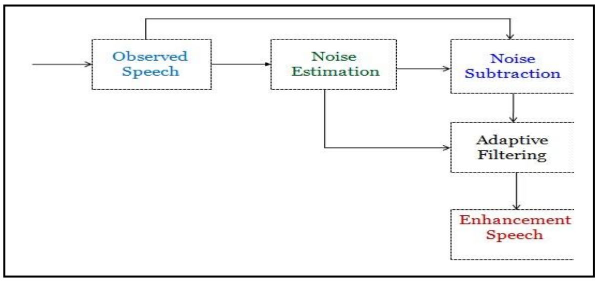


Fig 2: Block diagram of the proposed noise reduction system.

1.4 Applications of Proposed Method

Noise reduction techniques can be applied to reduce or remove the effect of unwanted signals to improve clarity of speech for transmission over the radio channel, or for operation of devices via a voice interface using a speech recognizer.

Hearing-impaired people usually have difficulty to understand speech because of background noise and reverberation. A major contribution of our proposed system is in ageing, whose ear is to become less sensitive to sound and less effective to detect limpidness speech. Therefore, a noise reduction method has to be carried out in the hearing aid. Most currently available hearing aids are of little use in assisting with improvement in this inadequate differentiation problem, and hence attempting to find successful speech processing schemes is an important area of research.

This research proposed an effective model to reduce maximum noise at fewer time and can be deployed in the following areas:(1) for telecommunication applications, (2) for better performance in automatic speech recognition(ASR) systems, (3) for mobile communication, (4) in multi-user mobile radio communications, (5) in the cockpits of the jet fighter aircraft and helicopters, (6) in cocktail party problem, (7) in the teleconferencing applications, (8) in car navigation systems, (9) in car interior communication systems **[3]** and,(10) in the auditorium.

2. EXISTING SCENARIOS

The subject of noise reduction for speech enhancement has a 40-years history in the field of telecommunications. Here, we discuss various types of noise reduction techniques as published in the literature. More attention is given to find out the disadvantages and drawbacks of traditional methods such as spectral subtraction (SS), blind source separation (BSS), adaptive noise cancellation (ANC), weighted noise subtraction (WNS)+BSS, and multi-channel based methods.

Spectral Subtraction (SS) method requires voice activity detector (VAD)[2], and also generates musical noise. Spectral subtraction works with the assumption that the noise is additive and stationary [5].

Blind Source Separation (BSS) has lack of a priori knowledge about the mixing system is compensated for by a statistically strong but physically plausible assumption of independence [9].

ANC needs simple hardware in applications, this would be difficult to implement in a real-time environment. In addition, long filter length can cause large mis -adjustment errors in the filter tap weights. Due to the feedback nature of an adaptive algorithm, the mis-adjustment errors can lead to echoes in the filtered speech. A desired signal corrupted by additive noise can often be recovered by an adaptive noise canceller using the LMS algorithm [4], ANC needs multi-channel adaptive filter.

Multi-channel based speech enhancement provides a better performance than single-channel case but have the disadvantage of requiring a second signal and lengthier computational times. To achieve an acceptable performance, it is needed a large number of microphones. Unfortunately, this is not practical in general in terms of spatial placement [6].

3. IMPROVED SINGLE CHANNEL ADAPTIVE FILTERING FOR SPEECH ENRICHMENT

The overall speech quality and intelligibility may be degraded, since speech sounds are masked by the noise and speech features. Today, a great deal of our personal communication is cellular phones and intercom devices. Noise corrupted speech does indeed force the user of such communication equipments to strain both hearing and voice. Altogether, acoustic noise dramatically decreases the performance of speech coding and speech recognition algorithms. It urges for effective speech enhancement methods (adaptive filter is required).The process is blind because we need no idea about either the signal or the noise. The principle of this investigation is to explore the utilization of the adaptive filter in order to improve speech enhancement performance and computational time. It is very important to investigate noise reduction problem using the adaptive filtering in the context of the BSS. Adaptive noise cancellation (ANC) technique deals with the enhancement of noise corrupted signals to enhance the performance of the BSS technique. We have proposed that the number of iterations to update filter coefficients is an important factor in such type of systems. In our research, we discuss the theory of adaptive filters to derive its implementation and propose number of iterations of the Least Mean Square (LMS) algorithm on basis of the estimated degree of noise (EDON). The main task of an adaptive filter is to update the filter coefficients/weights so that a residual signal is minimized. The finite impulse response (FIR) filter efficiently removes the environmental noise (broadband/narrowband) from the speech. However, the proposed method is single-channel adaptive filtering based on optimal iteration.

3.1 Adaptive Filter for Speech Enrichment

The main task of an adaptive filter is to update the filter coefficients/weights so that a residual signal is minimized **[4]**. If the signal is corrupted only by single or multiple fixed frequency components, the removing these unwanted elements of the speech signal would be a simple process and required a fixed iteration number of digital filters. However, in real environments this is rarely the case, because the variable frequency components are available. Fixed linear digital filters would be unable to recovery such problems and thus adaptive filtering processing technique is needed. The process is blind, since we need the one single-channel observed speech, where speech and noise are acoustically combined to make simulation of the real noisy environment.

The conventional adaptive filter consists of two inputs such as (1) estimated noise alone and (2) speech mixed with noise. A second channel is used to generate a speech-free measurement of the reference noise (estimated noise) **[1]** in the most of the previous studies. In this thesis, a similar method by single channel adaptive filtering approach based on the décor relation decisive factor is presented. The supreme potency of the projected method compared with other noise diminution technique is that no a priori information about either the signal or noise is required. Furthermore, a noise-dominant signal (estimated noise) is generated, which eradicates the necessity for the second channel.

The input signal x(n) is the sum of a desired signal d(n) and interfering noise v(n) as shown in fig.3.

Therefore desired speech signal:

$$d\,(n) = x\,(n) - v\,(n) \qquad (1)$$

=Noisy Speech –Noise

After side changing: $x(n) = d(n) + v(n)$ (2)

The error signal $e^t\,(n)$ is the difference between the desired d (n) and the estimated signal $d^t\,(n)$:

Error at t-th step: $e^t\,(n) = d\,(n) - d^t\,(n)$ (3)

Here, $d^t\,(n)$ = speech estimated at t-th step.

Estimated Speech $d^t\,(n)$: $d^t\,(n) = Wn^{\,t} * x\,(n)$ (4)

Moreover, the variable filter updates the filter coefficients Wn^t at every time instant: $Wn^{\,t} = Wn^{\,t} + \Delta Wn$ (5)

Where, ΔWn is the correction factor for the filter coefficients.

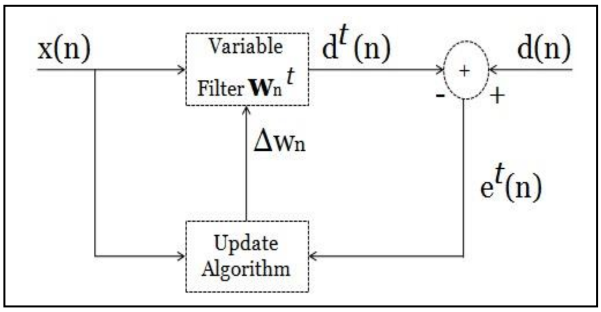
Therefore, $\Delta Wn = \eta * x(n) * \delta$ (6)

Hence,

$\Delta Wn = \eta * x(n)\,[d(n) - d^t(n)]$; $[\delta = d\,(n) - d^t\,(n)]$ (7)

=updating factor * input speech [desired speech - estimated speech]; [updating factor range $=0<\eta<1$].

The adaptive algorithm generates this correction factor based on the input x (n) and error $e^t\,(n)$ signals.



3.2 Derivation of Proposed Iteration Number for Adaptive Filter

At this section, we depict the projected iteration number assessment process for the least mean square (LMS) algorithm by the least-squares (LM) method [5]. The least mean square (LMS) algorithm is an iterative-based method. This procedure is cyclic through several iterations until the residual turns into sufficiently small in a statistical sense. Again, improper iteration increases output noise power that causes distorted speech. The major goal is to develop both the signal quality and computational time. We find the different minimum iterations for each input SNR. That is, a large number of iterations are required with a lower signal to noise ratio (SNR) than with a higher SNR. Using the proposed iteration numbers for both the lower and higher SNR is protracted .In the planned model, the iteration number is decided on the source of the estimated degree of noise of noisy speech. A function is equipped, by the least-squares (LM) method, from the estimated degree of noise (EDON) in each input SNR and the minimal iteration is required to obtain maximum noise reduction.

Figure 4 is the experimental result that shows the least-squares line indicating the relationship between proposed iteration number (I_{NP}) and estimated degree of noise [1],[2] (D_{mE}). In this experiment, we used five English vowel sounds articulated by 3 male and female speakers (training).

The minimum iteration number I_N required increases linearly with increasing degree of noise, when D_{mE} is between 0.267874 and 0.858 .I_N is practically constant when D_{mE} is less than 0.267874. In our research, we set as $I_N = 4$, when $D_{mE} > 0.267874$ at silent speech [1].

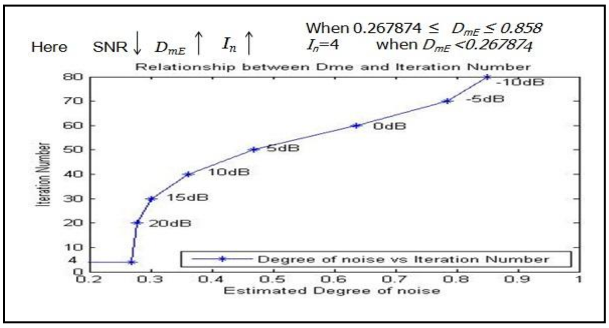


Fig 4: Relationship between I_N and D_{mE}.

According to the curve, suppose

$$Y=B_0+B_1X+B_2X^2+B_3X^3 \quad (8)$$

be the third degree parabola of best fit to set of n points (x_i,y_i); i=1,2,3……,n. Using the principle of least squares, we have to determine the constants $d=B_0$, $c=B_1$, b= B2, a=B3 so that the following table 1 is for iteration number(y) and values of x.

Table 1. Iteration number(y) and values of x

Y	X	X^2	X^3
20	0.277874076	0.0772	0.0215
30	0.301812176	0.0911	0.0275
40	0.360936494	0.1303	0.0470
50	0.469251476	0.2202	0.1033
60	0.636070106	0.4046	0.2573
70	0.7849689	0.6162	0.4837
80	0.850378035	0.7231	0.6149

Y=Iteration Number (I_n), X=Estimated Degree of Noise (D_{mE})

Taking coefficients of B_0, B_1, B_2 and B_3 from the equation $Y=B_0+B_1X+ B_2X^2+B_3X^3$; we get X Matrix:

$$X = \begin{pmatrix} 1 & 0.277874076 & 0.0772 & 0.0215 \\ 1 & 0.301812176 & 0.0911 & 0.0275 \\ 1 & 0.360936494 & 0.1303 & 0.0470 \\ 1 & 0.469251476 & 0.2202 & 0.1033 \\ 1 & 0.636070106 & 0.4046 & 0.2573 \\ 1 & 0.7849689 & 0.6162 & 0.4837 \\ 1 & 0.850378035 & 0.7231 & 0.6149 \end{pmatrix}$$

Taking values of Y from $Y=B_0+B_1X+ B_2X^2+B_3X^3$; we get Y Matrix:

$$Y = \begin{pmatrix} 20 \\ 30 \\ 40 \\ 50 \\ 60 \\ 70 \\ 80 \end{pmatrix}$$

We know that, the 'least squares solution' for input matrix X and output matrix Y is the following:

$$B= (X^T.X)^{-1}. X^TY \quad (9)$$

And

$$B = \begin{pmatrix} -122.25 \\ 834.1 \\ -1339.7 \\ 751.3 \end{pmatrix}$$

So the coefficients are a=751.3, b= -1339.7, c= 834.1, d= -122.25.Now we can adjust the third degree parabola equation to have the proposed iteration number equation:

$$I_{NP}=a^* D_{mE}^3-b^* D_{mE}^2+c^* D_{mE}-d \quad (10)$$

When $0.267874 \le D_{mE} \le 0.858$ [**2**]

I_n=4 when D_{mE} < 0.267874; [at the silent speech].

To analyze the iteration number, it requires a plot of the frame number against I_{NP} is shown in Fig. 5. It shows the various iteration numbers of the frame of speech signal.

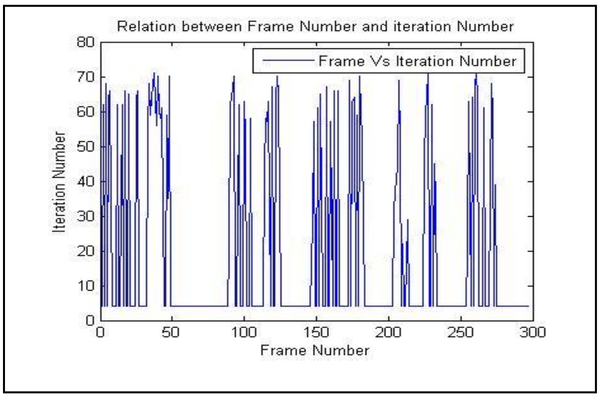


Fig 5: Relationship between frame number and iteration number

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

In our research, three types of experiments are conducted. The first is to evaluate the frame size, second one is to evaluate the windowing and DFT point and finally to evaluate the adaptive filter coefficient that is proposed method.

4.1 Setup for Proposed Method

This experiment was conducted by the propose technique, we consider the whole speech sequence to make the SNR decision. The projected algorithm is deployed using C programming language and the recital is examined using Matlab tools. The white noise sample is obtained by a normal random number generator (WGN [generate White Gaussian Noise] function from Matlab).The speech signal is contaminated by preservative noise at various signal to noise ratio(SNR) levels. The experiment uses English vowel sounds from female and male speakers. In this experiment, vowel sounds are collected from the TIMIT database [13]. Both the noise and speech are sampled at 10 kHz; earlier than they are digitally mixed to generate the noisy speech, is conducted to set the filter orders r_{12} and r_{21}. We set r_{12} = 0, then vary r_{21} from 0 to upwards. It shows that after r_{21} =100, the noise reduction is almost constant. In the experiment, white noise is degraded by the 0dB SNR. The orders of the filters are r_{12} = 0 and r_{21} = 100 and the step-size parameters are μ_1=1 and μ_2=0.007 used in the experiment. Noise reduction (NR) within the trial portion is estimated using the following equation (Equ.11).

$$NR(dB) = \frac{1}{K}\sum_{j=1}^{K} 10\log_{10} \frac{\sum_{n=k_j-l+1}^{k_j} (x(n)-s(n))^2}{\sum_{n=k_j-l+1}^{k_j} (\hat{s}_1(n)-s(n))^2} \qquad (11)$$

$k_1 = l$ and $k_{j+1} = k_j + l'$

Where s(n) , x(n) and $\hat{s}_1(n)$ are the clean speech ,the noisy speech and the enriched speech in a discrete time n, respectively. The constants *K, l, and l′* are the total number of segments, the frame length, and frame shift, respectively.

4.2 Parameter Setting for the Function to Estimate INP

The table 2 shows the average noise reduction of English vowel sounds in various iteration numbers. In the experiment, we use five English vowel sounds from 3 male and 3 female speakers. The table also shows for each input SNRs, the average estimated degree of noise. Using data tabulated in the Table 2, we notice the average D_{mE} in each input SNR for different vowel sounds and decided the minimum iteration number I_{Nmin} required for maximum noise reduction as demonstrated in Table 3.

Table 2. Noise reduction of vowel sounds in various iteration

SNR	Av.D_{mE}	I_N=5	I_N=10	I_N=20	I_N=30	I_N=40	I_N=50	I_N=60	I_N=70	I_N=80
-10	0.85	15.25	16.51	17.85	18.68	19.22	19.65	19.85	19.81	**19.91**
-5	0.78	12.76	13.84	15.13	16.03	16.33	16.65	16.79	**16.87**	16.78
0	0.64	10.96	11.98	12.93	13.57	13.93	14.19	**14.29**	14.22	14.25
5	0.47	8.96	9.77	10.55	10.93	11.16	**11.29**	11.26	11.36	11.30
10	0.36	6.06	6.67	7.28	7.47	**7.47**	7.42	7.53	7.54	7.50
15	0.30	3.53	4.01	4.29	**4.47**	4.44	4.46	4.58	4.56	4.51
20	0.28	1.89	2.15	**2.39**	2.37	2.46	2.49	2.47	2.18	2.57

Sounds from 3 male and female speakers are averaged

Table 3. Minimum iteration number required for various noise levels

SNR	D_{mE}	I_{Nmin}
-10	0.8503378	19.91054
-5	0.784969	16.86575
0	0.63607	14.28648
5	0.469251	11.28853
10	0.360936	7.46857
15	0.301812	4.47165
20	0.277874	2.39259

4.3 Noise Reduction Results

Table 4 represents the average noise reduction results of sentences. Ten English sentences are degraded by synthesized white noise at various noise levels. The single channel adaptive filter is carried out using the observed signal y_1 (n) =x (n) and estimated second noise y_2 (n).Using the proposed method, a better noise reduction result is obtained specially in the low SNR parts, but in the high SNR parts, the result is unsatisfactory.

Table 4. Average noise reduction results of English sentences in db

Noise Reduction Results in db		
SNR	**White Noise**	**D_{mE}**
-10	15.9646	0.8503378
-5	14.6251	0.784969
0	13.3777	0.63607
5	11.3203	0.469251
10	8.5344	0.360936
15	5.7086	0.301812
20	2.1717	0.277874

4.4 Signal to Noise Ratio Results

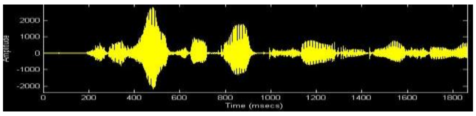
We have carried out experiments with synthesized vowel sounds and synthesized noise. Table 5 shows the average SNR improvement of the real vowel sounds degraded by the synthesized white noise at various noise levels. The ordinary SS, BSS+WNS and proposed model provides average eliminated noise 6.5dB, 8.7dB and 10.8dB at 0dB SNR . Table 5 shows that the performance decreases with increasing SNRin (input SNR).

Table 5. Average result of ordinary SS, ordinary WNS+BSS, and Proposed Method

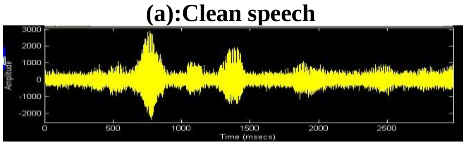
VOWEL	0dB			5dB			10dB			15dB		
	SS	**WNS + BSS**	**Proposed method**	**SS**	**WNS + BSS**	**Proposed method**	**SS**	**WNS + BSS**	**Proposed method**	**SS**	**WNS + BSS**	**Proposed method**
/a/	5.9	7.1	8.0468	5.3	5.9	6.2887	4.0	4.7	5.5989	2.4	3.2	4.4396
/e/	6.9	9.0	10.7282	6.4	7.8	8.2831	5.3	5.9	6.1772	3.8	4.2	3.9466
/i/	6.8	9.8	10.6965	6.7	8.7	8.4470	5.8	7.2	7.2903	4.6	5.5	3.5948
/o/	6.2	7.8	9.1427	5.3	6.4	4.6310	3.9	4.7	3.5244	2.3	2.8	2.3421
/u/	6.9	9.6	15.1757	6.8	9.0	11.6381	6.2	7.4	10.2106	5.3	5.9	7.8137
Avg.	6.5	8.7	**10.75798**	6.1	7.6	**7.85758**	5.0	5.98	**6.5602**	3.7	4.3	**4.42736**

Results of white noise reduction are averaged

Using a single-channel adaptive filter method, it is very complicated to progress mutually the speech eminence and intelligibility at the same time. That is why, the more the noise is eliminated, the more the speech drops intelligibility. We endeavor to a better manage the tradeoff between speech distortion and noise reduction. The clean speech and noisy speech sequences are as shown in figure 6 respectively. In this figure, the speech sentence “She had your dark suit in greasy wash water all year” from TIMIT corpus **[13]** is degraded by white noise at 5dB SNR. Then we scrutinize the effect of the number of filter update iterations on the lengthy speech sequence as shown the results in fig. 7.

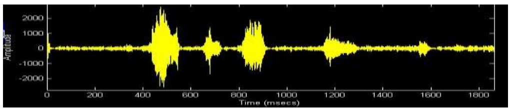


(a):Clean speech

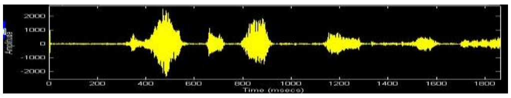


(b): Nosy speech

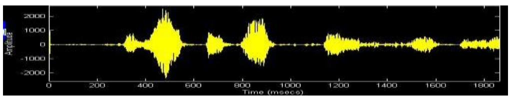
Fig 6: Speech waveforms; (a) Clean speech and (b) noisy speech at 5dB SNR.



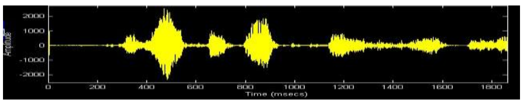
(a):Iteration=2



(b):Iteration=10



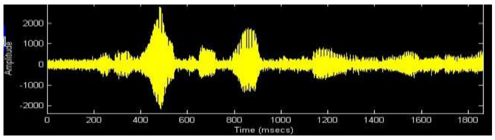
(c):Iteration=30



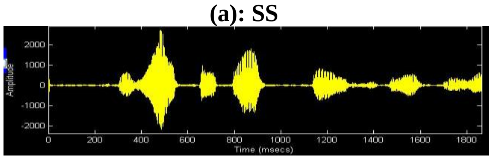
(d):Iteration=50

Fig 7: Enhanced speech waveforms for different iteration numbers in the single channel adaptive filter method, (From top) I_N =2, 10, 30, and 50.

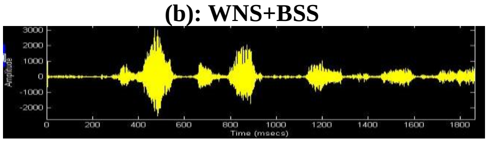
Finally, the last noise-reduction speech is obtained by proposed (I_N=I_{NP}) iteration number as shown in Fig. 8(c).The fig.8 (a, b) also shows results of SS and the WNS+BSS.



(a): SS



(b): WNS+BSS

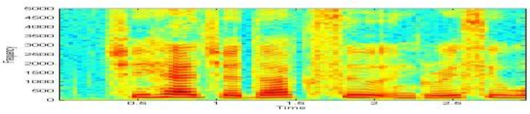


(c): A single channel adaptive filter

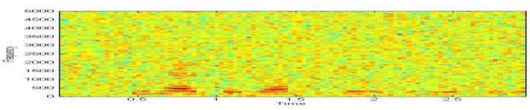
Fig 8: Enhanced speech wave forms for different methods. (a) The SS, (b) the WNS+BSS, and (c) Proposed (when I_N = I_{NP}) methods.

We examine the noise diminution performance in the voiced and the silent part independently. It is examined for each input signal to noise ratio (SNR), noise diminution in the voiced part is less than that in the silent part. The analysis results are as shown in figure 9. The speech sample is same as in the case of figure 6 and 7 are used.

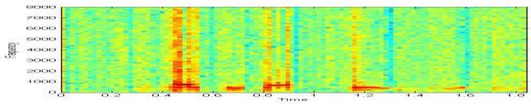
We assess the speech spectrograms and confirm the quality of the enriched speech obtained by the proposed model. The spectrograms of (a) Clean speech (same speech as in fig. 6), (b) Noisy speech (white noise at 5 dB SNR), and (c) to (f) the outcomes with different iteration number as shown in fig. 9. Figure 10 also shows speech spectrograms of SS, WNS+BSS and proposed (I_N = I_{NP}) methods.



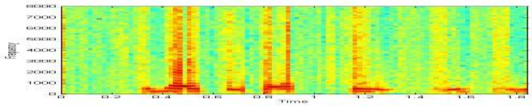
(a):Clean speech spectrogram



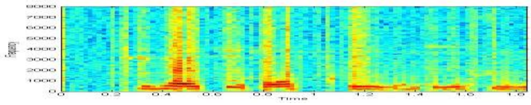
(b): Noisy speech spectrogram



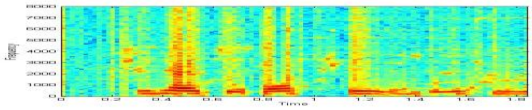
(c): Iteration=2



(d): Iteration=10

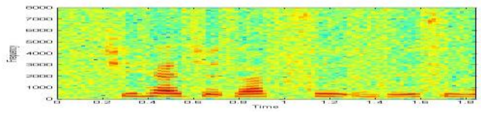


(e): Iteration=30

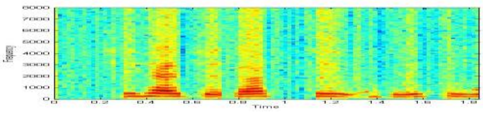


(f): Iteration=50

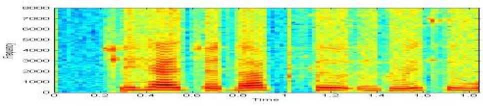
Fig 9: Spectrogram of various Iteration number in 5dB speech signal



(a): SS



(b): WNS+BSS



(c): Proposed Method

Fig 10: Spectrogram of SS, WNS+BSS and Proposed method.

5. CONCLUSION

In this study, the proposed method does not use voice activity detector (VAD). It introduced the DON, which was a good tool for various adjustment settings for the experiment. The proposed fixed iteration based single-channel adaptive filtering technique overcomes the requirement for second channel and lengthier computational times. Noise reduction score is greater for wide band noises; also in high and mid frequency levels. The proposed iteration numbers are produced in good noise reduction performance in both speech and silent parts using less computational time.

- Higher noisy speech requires higher number of iterations in filter.
- Less noisy speech requires fewer numbers of iterations.
- An equation, iteration number=f (degree of noise) is derived.
- Our proposed method reduces more noise compared to others.

Future works are stated as below**:**

- Noise reduction by means of multi-channel adaptive filtering in concern reducing cost and time.
- Study to reduce the range of Estimated Degree of Noise.
- Color noise reduction from speech signal in real environment except white noise.
- Noise reduction by considering large amount of user created speech databases.
- Study the minimum iteration number for each data value.

6. REFERENCES

[1] Hamid, M.E., Hasan, M., Inamurrahman, M., and Ghribi,W., "An Improved Sad Algorithm of Noisy Speech Using Estimated Degree of Noise",canadian journal on signal processing issn: 1923-1709, july 15, 2011, canada.

[2] Hasan, M., Hamid, M.E.," A parametric formulation to Detect Speech Activity of noisy speech using EDON"; iccit – 2010, iccit – 2010, paper id-155, page- 252 - 255; 23-25 december, 2010, dhaka, bangladesh.

[3] Huang, Y., and Benesty, J., Audio Signal Processing for Next-Generation Multimedia Communication Systems, Kluwer Academic Publishers, 2004.

[4] Lim, J. S., and Oppenheim, A. V., "All-pole Modeling of Degraded Speech," IEEE Trans. Acoust., Speech Signal Process., vol 26, pp 197- 210, June 1978.

[5] Jutten, C., and Hérault, J., "Independent Component Analysis versus Principal Component Analysis", Signal Processing IV, pp. 643-646, 1988.

[6] Thi, H. L. N., and Jutten, C., "Blind Source Separation for Convolutive Mixtures", Signal Processing, vol.45, pp.209-229, 1995.

[7] Hamid, M. E., Ogawa, K., and Fukabayashi, T., "Improved Single Channel Noise Reduction Method of Speech by Blind Source Separation", Acoust. Sci. & tech., Japan.

[8] Hamid, M. E., and Fukabayashi, T., "A Two-Stage Method for Single-Channel Speech Enhancement", IEICE Trans. Fundamentals, vol. E89-A, no.4, pp. 1058-1068, April 2006.

[9] Hamid, M. E., and Fukabayashi, T., "Noise Reduction of Speech Signal using Noise Subtraction and Blind Source Separation", Annual Report of Graduate School of Electronic Science and Technology, Shizuoka University, vol. 27, pp. 31-37, 2006.

[10] Hamid, M. E., Ogawa, K., and Fukabayashi, T., "Performance of Two-Stage Noise Reduction Method by Iteration Number of Separation Filter Update in Processing Block", Proc. Technical Report of IEICE, pp.19-24, Wakayama, January 2006.

[11] Hamid, M. E., Ogawa, K., and Fukabayashi, T., "Improvement of Nose Reduction of Speech using Blind Signal Separation",Proc. ASJ spring meeting, pp. 639-640, Tokyo, March 2006.

[12] Hamid, M. E., Ogawa, K., and Fukabayashi, T., "Wide Band Noise Reduction of Speech using Noise Subtraction and Blind Source Separation", Proc. WESPAC IX, Seoul, June 2006.

[13] TIMIT Acoustic-Phonetic Continuous Speech Corpus site https://catalog.ldc.upenn.edu/LDC93S1.