Indian Journal of Science and Technology, Vol 9(39), DOI: 10.17485/ijst/2016/v9i39/100594, October 2016
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Speech Enhancement using Kalman Filter with Preprocessed Digital Expander in Noisy Environment

G. Manmadha Rao* **and U. Santosh Kumar**

Department of ECE, GMR Institute of Technology, Rajam - 532127, Andhra Pradesh, India;
manmadharao.g@gmrit.org, uskscientist@gmail.com

Abstract

Objective: The primary objective of the Speech Enhancement algorithms is to enrich the superiority of speech. The superiority of speech is articulated in two factors, clarity and other is intelligibility. **Methods/Statistical Analysis:** The method to improve the quality of speech in this paper is proposed based on computationally efficient AR modeled Kalman Filter with digital compressor/expander. This approach is based on reconstruction of noisy speech signal using digital expander and further enhancement with Auto Regressive modeled Kalman filter. Findings: The results of proposed method in terms of SNR and intelligibility are found to be better compared to earlier methods like spectral subtraction; wiener filter and Kalman filter methods. Application/Improvement: This study suggests that improvement in speech signal recovery in noisy environment helping researchers for developing efficient devices in the field of Speech recognition systems, Speech based authentication systems, audio processing devices and so on.

Keywords: Digital Compressor/Expander, Digital Filters, Intelligibility, SNR, Spectral Subtraction, Speech Enhancement, Wiener Filter

1. Introduction

Speech is the primary mode of Communication among all human being. It is very efficient and effective way of communication. The Speech processing is widely used in many applications like mobile phones, VOIP, Teleconferencing systems, voice enabled security devices, household appliances, speech recognition systems, hearing aids, biomedical signal processing, ATM machines and computers[1]. Various types of additive noise in real-world environments often corrupt speech. Unfortunately characteristics of this additive noise is difficult to estimate due to it has different characteristics in different environments. So speech enhancement is very much required. There are many speech enhancement techniques in stationary and non- stationary noisy environments like spectral subtraction, wiener filtering, and subspace based methods etc. Spectral subtraction is fundamental method; recovery of the signal is done by estimating the magnitude spectrum or power spectrum of a signal detected in additive noise and later subtracting the average noise spectrum from the noisy signal spectrum[2]. Spectral subtraction method is the basic method where the noisy suppression can be done by estimating the spectral characteristics of the noise and separating from the noisy speech spectral characteristics. Major drawback is musical noise and also noise during the silence period[3].

In Wiener filter based speech enhancement method using system transfer function characteristics original speech signal is recovered by minimizing Mean Square Error (MSE)[4]. With this method speech quality is improved, but the musical noise still has an influence on the quality of speech[5]. As said by Pallival and Basu, the speech signal parameters are estimated before it gets disturbed by white noise[6]. Based on the speech model parameters and noise variance estimation a new time adaptive algorithm is used by[7]. Gannot proposed the Espectation-Maximisation algorithm which repeatedly calculates the Spectral parameters of speech and noise

*Author for correspondence

parameter[8]. There are lot of changes made to the basic above stated algorithms by many authors but those doesn't meet the expectations. Those algorithms are good at SNR but giving low intelligibility in speech, and those which are good at intelligibility are giving low SNR.

In this paper to overcome above stated problem a new Iterative adaptive Kalman filter based method with preprocessing of digital audio effecting technique called digital expander is proposed. In this method speech signal from a series of noisy speech signals and noise signal is considered as modeled as the AR process[9]. This time-varying Auto Regressive (AR) speech model coefficients are estimated based on Linear Prediction Coefficient estimation (LPC). In addition to coefficients estimation this paper solved problem of de-noising the colored noise. We made an assumption that the noise is also an autoregressive process[10]. So we estimated its AR coefficients and variances by LPC in the same way.

2. Mathematical Depiction

One of the most efficient speech enhancement techniques is Kalman filtering method, in which speech signal is generally presented as AR model and visualized in the state-space domain. Kalman filter techniques are analyzed in two steps. The first step is to estimate the noise, signal model parameters and driving variances and then the parameters are updated. This method is introduced because of low SNR and speech intelligibility degradation in Kalman filter based speech enhancement.

2.1 Iterative Adaptive Kalman Filter

The speech signal $x(n)$ and the additive noise or disturbance $d(n)$ are articulated in terms of p th order Autoregressive (AR) model shown below

$$x(n) = \sum_{i=1}^{p} a_i x(n-i) + u(i) \tag{1}$$

$$d(n) = \sum_{j=1}^{p} b_j x(n-j) + w(i) \tag{2}$$

And Noisy speech from above Equation (1), (2) can be expressed as

$$y(n) = x(n) + d(n) \tag{3}$$

Where $x(n)$ is the speech signal, $d(n)$ is the additive noise and $y(n)$ is the noisy speech.

a_j and b_j are AR model parameters.

State-space form of above AR model is

$$x(n+1) = A(n)x(n) + (u(n), 0, \dots\dots, 0)^T \tag{4}$$

$$A(n) = \begin{bmatrix} a_1(n) & \cdots & a_p(n) & 0 & & \cdots & 0 & 0 \\ 1 & \cdots & 0 & 0 & & \cdots & 0 & 0 \\ \vdots & \ddots & & & & & & \vdots \\ 0 & 1 & & & 0 & \cdots & 0 & 0 \\ 0 & & 0 & 1 & & \cdots & 0 & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & \cdots & 0 & 0 & & \cdots & 1 & 0 \end{bmatrix} \tag{5}$$

$$d(n+1) = B(n)d(n) + (w(n), 0, \dots\dots, 0)^T \tag{6}$$

$$B(n) = \begin{bmatrix} b_1(n) & \cdots & b_{q-1}(n) & b_q(n) \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix} \tag{7}$$

From the above discussion the augmented state vector and driving noise vectors $X(n)$, $W(n)$ respectively.

$$X(n) = \begin{pmatrix} s(n) \\ v(n) \end{pmatrix}, \qquad W(n) = \begin{pmatrix} u(n) \\ w(n) \end{pmatrix} \tag{8}$$

From Equation (3) and (4)

$$X(n+1) = (F(n) * X(n)) + (G * W(n)) \tag{9}$$

$$y(n) = C^T * X(n)$$

Where $F(n) = \begin{pmatrix} A(n) & 0 \\ 0 & B(n) \end{pmatrix}$ $G = \begin{pmatrix} e_s & 0 \\ 0 & e_v \end{pmatrix}$ $C = \begin{pmatrix} e_s \\ e_v \end{pmatrix}$

$e_s = (1, 0, \dots 0 \dots\dots 0 \dots 0,)^T$ with d+1 dimension and e_v with q dimension. Now we can optimally suppresses the disturbing noise by calculating Kalman filter basic parameters such as variance and gain[8]. The process of Iterative Kalman filter[8] is in two steps 1. Estimation: State vector propagation, parameter covariance matrix propagation and 2. Updation: Compute Kalman gain, state vector update, parameter covariance matrix update.

2.2 Digital Expander

Digital Expander is mainly used to control the Dynamic range. With logarithmic function, dynamic range can be calculated by taking ratio between maximum to minimum amplitude of the signal.

Mathematically Dynamical range can be obtained by using the below equation

$$op(n) = g(n) * ip(n-D) \tag{10}$$

where $op(n)$ is output signal, $g(n)$ is gain factor and $ip(n-D)$ is delayed input signal

After applying logarithm to the Equation (10)

$$op_{dB}(n) = g_{dB}(n) + ip_{dB}(n-D) \tag{11}$$

Gain factor is affected by level measurement, Static curve, and Attack and release time. Level measurement can be measured by calculating the absolute value and peak values of input signal. By comparing the absolute and peak values output of the level measurement can be calculated using Equation (12) and (13).

If absolute value is greater than peak value of input then

$$op_{peak}(n) = (1 - Attack\ Time)ip_{peak}(n-1) + Attack\ Time(ip(n)) \quad (12)$$

$other\ wise$

$$op_{peak}(n) = (1 - Release\ Time) * ip_{peak}(n-1) \quad (13)$$

Static curve is obtained by using change in input level to the change in weighting level. This is done with the help of compression factor.

$$compression\ factor = \frac{\Delta ip}{\Delta op} \quad (14)$$

It can be written in terms of slope and compression factor as

$$slope\ of\ the\ static\ curve = 1 - \frac{1}{compression\ factor} \quad (15)$$

Finally the attack and release times can be generated using gain smoothing factor g(n). The block diagram of compressor and expander show in Figure 2. Based on comparison between Input signal with threshold, smoothen gain factor is calculated. The basic structure is similar to the limiter.

3. Implementation and Evaluation of Proposed Method

Figure 1 shows the flow of proposed method based on it Mat lab code is developed, where Clean Speech and different noise level speeches are considered from the Noizeus Database and random values are added to the clean speech in order to generate Noisy Speech. A Digital Expander is used to the Noisy Speech at expand Factor of 0.5 as shown in Figure 2 and then applied to Kalman filter for better Speech Enhancement. Here different real time noisy signals taken from Noizeus database. Here Digital Expanded Noisy speech signal and noise signals are represented by an AR model of order P = 20. For every frame of 25 ms duration, these AR coefficients are updated and analyzed using the linear prediction analysis method (LPC).

The additive measurement noise is expected to be stationary during the each small frame. LPC coefficient estimation order was taken as 13 for noisy speech and

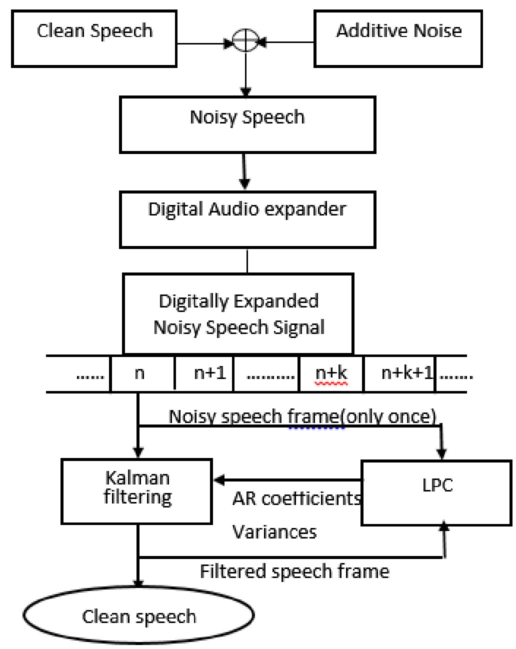

Figure 1. Block diagram for iterative kalman filter with digital expander.

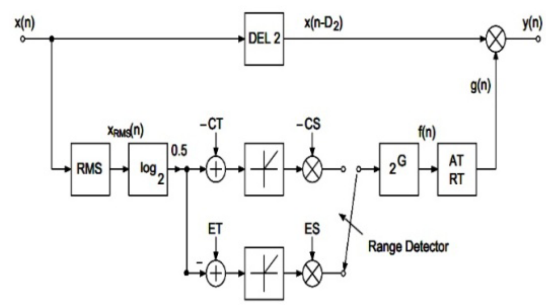

Figure 2. Block diagram of digital expander.

7 iterations for Kalman filter to give high SNR with low intelligibility. To overcome this drawback digital expander is applied to the noisy speech signal with expand factor of 0.5. It is found to be good in SNR and intelligibility.

4. Results

In this paper, we have observed and tabulated the results of Spectral Subtraction, Wiener Filter Kalman filter methods and compared with Iterative Kalman filter with digital expander method. Here different real time noise signals (From NOIZEUS database) of 0 dB, 5 dB, 10 dB and 15 dB are considered for performance analysis, with Hanning window in above three stated algorithms. Compared to all these methods, proposed algorithm is exhibiting the best results in terms of SNR and intelligibility. The corresponding wave forms are shown below. Simulation results shows that the proposed digital expander with kalman filter based technique is effective for speech enhancement compare to other filter based techniques. Iterative Kalman filter, proposed method results and other basic methods output wave forms are illustrated in Figure 3–11. Figure 11 shows output waveforms of Kalman and proposed method, Table 1 represents corresponding SNR

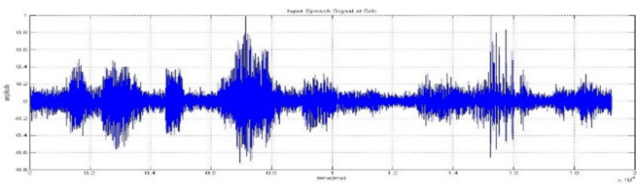

Figure 3. Input signal_5db.

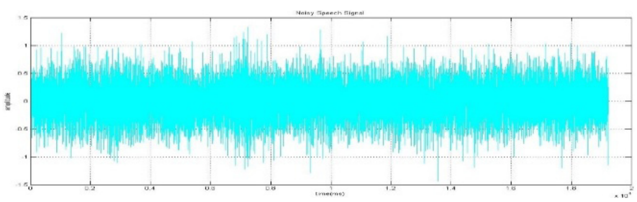

Figure 4. Noisy speech signal.

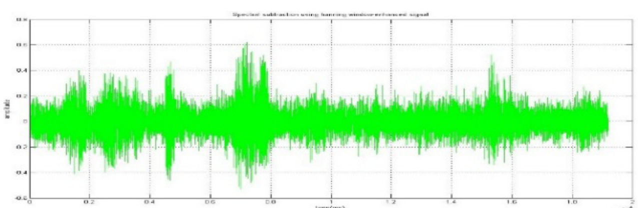

Figure 5. Enhanced signal (at 5dB) using spectral subtraction.

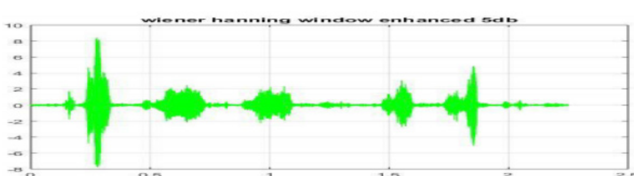

Figure 6. Enhanced signal using wiener filtering (at5dB).

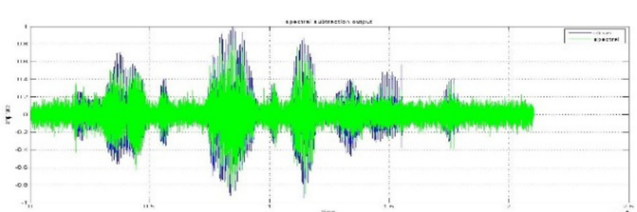

Figure 7. Clean speech vs spectral recovered speech.

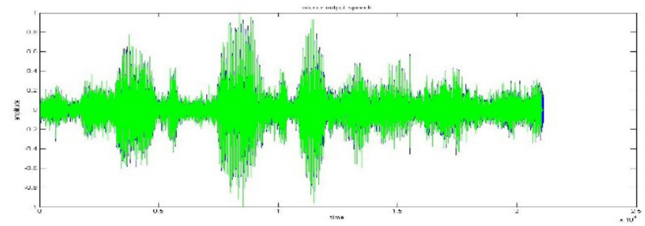

Figure 8. Clean speech vs wiener recovered speech

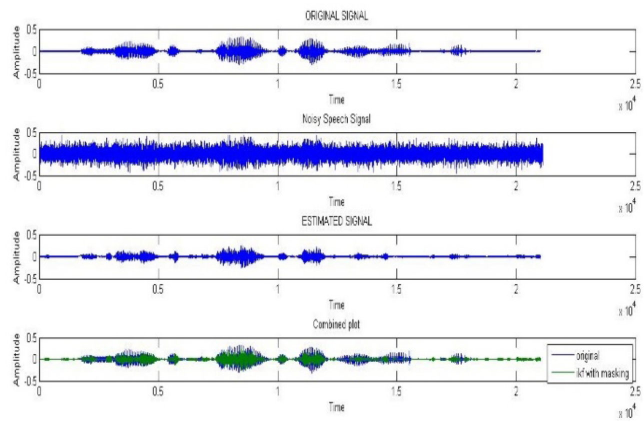

Figure 9. Kalman Filter Method output speech wave forms.

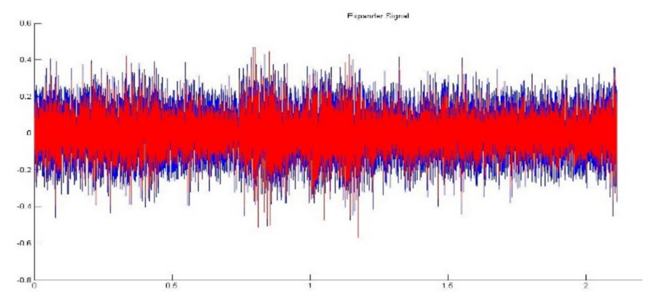

Figure 10. Noisy speech vs digital expander output

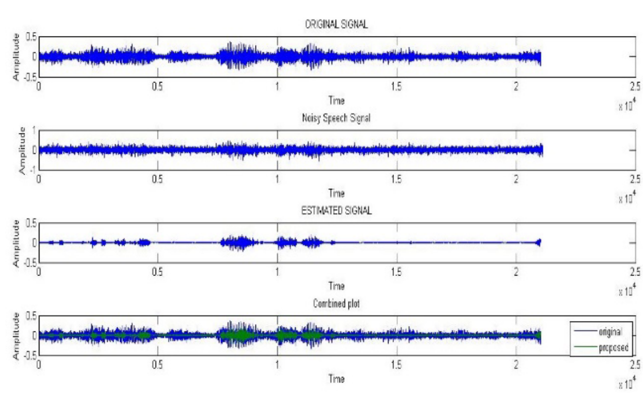

Figure 11. Proposed method (kalman with digital expander) speech wave forms.

Table 1. Input-output SNR comparison

Noise Level(in dB)/ Method		Output SNR				Proposed method PESQ Scores						
		Spectral Subtract	Wiener filter	Kalman filter	Proposed	LLR	SNR SEG	WSS	PESQ	C	B	O
Clean2		−7.1276	−6.2947	−3.3244	−1.1276	0.8791	−8.1220	40.6133	2.9364	3.3528	2.0508	2.9020
Station (dB)	15	−7.2568	−6.4760	−3.3021	−1.6284	0.8951	−8.2685	46.7672	2.9556	3.1715	1.9117	2.7047
	10	−7.8476	−6.8592	−3.4796	−1.2903	0.9079	−8.5755	49.0362	2.8450	3.0712	1.8236	2.5932
	5	−8.7458	−7.9490	−3.6577	−1.0106	0.9874	−8.9637	51.0818	2.7430	2.8492	1.6883	2.3755
	0	−9.9899	−9.9317	−4.4322	−1.8839	1.1375	−9.3760	57.9182	2.5426	2.5727	1.5665	2.1700
Restaurant (dB)	15	−7.1147	−6.2895	−3.3514	−1.3830	0.8457	−8.1558	43.5604	2.9105	3.2843	1.9675	2.7966
	10	−8.0258	−6.9964	−3.4496	−1.5633	0.8709	−8.4047	52.0910	2.9476	3.0833	1.8143	2.5928
	5	−8.5992	−8.5015	−3.5729	−1.2033	0.9742	−8.8856	54.6263	2.7144	2.8739	1.7025	2.4149
	0	−9.9689	−9.8391	−4.2656	−2.1383	1.1849	−9.1889	65.1016	2.4069	2.2567	1.3675	1.8252
Car (dB)	15	−7.2586	−7.0340	−3.4666	−1.5316	0.8632	−8.4473	44.8024	2.8703	3.2308	1.9212	2.7465
	10	−7.9856	−7.1374	−3.2491	−1.5001	0.9279	−8.5196	49.2566	2.8387	3.0485	1.8226	2.5763
	5	−8.8856	−8.2521	−3.5112	−1.4297	1.0408	−8.9284	54.0921	2.3797	2.7892	1.6870	2.3566
	0	−9.8957	−9.4048	−3.9676	−1.5051	1.2033	−9.3317	61.5375	2.2446	2.4736	1.5449	2.1126
Babble (dB)	15	−6.9856	−6.8325	−3.3749	−1.2405	0.8562	−8.3369	45.8447	2.9246	3.2614	1.9468	2.7865
	10	−7.5689	−7.4887	−3.3439	−1.0649	0.8841	−8.7751	49.0078	2.3475	3.0974	1.8124	2.6075
	5	−8.5896	−8.0256	−3.8456	−1.3569	0.9856	−8.7958	55.6524	2.2458	2.8564	1.7589	2.3568
	0	−9.7541	−9.1832	−4.2664	−1.6324	0.1689	−8.8893	64.5714	1.9099	2.4607	1.5349	2.0810

Table 2. Comparison of different methods with proposed method and PESQ scores

Noise/ Method	Spectral Subtraction	Wiener filter	Kalman filter	Kalman with digital expander
Clean speech	−7.2601	−6.6115	−3.2059	−1.6322
0dB	−10.6141	−9.8875	−4.3595	−2.3242
5dB	−8.5450	−7.9572	−3.6105	−2.0545
10dB	−7.6450	−7.0027	−3.4408	−2.0417
15dB	−7.2926	−6.6079	−3.2998	−1.6600

ratio values of all the above stated method. Table 2 list the results of composite performance measures, organized according to the segment SNR, LLR, PESQ, and C (signal), B (background distortion), O (output). And also Objective measure quality test is conducted and the results are placed below Table 2. Performance of proposed method is analyzed in different Noisy Environments like station, babble, car and restaurant. It is observed that the proposed method giving better SNR values and its performance is superior for both non-stationary and stationary speech signals.

5. Conclusion

In the present study, based on Adaptive Kalman filter and digital effective techniques, an improved method of Iterative Kalman filter with Digital Expander is proposed. In this paper we discussed drawbacks of speech enhancement with spectral subtraction and wiener filter methods. Even though conventional Kalman filter approach is giving better results than spectral and wiener but its response is poor at high SNR. In this paper we implemented iterative Kalman filter with digital expander which overcome the disadvantages of early two methods. All these methods are simulated and SNR values of respective methods are compared. Performance of Proposed method is analyzed in different Noisy Environment like babble, car, station, and restaurant. We found that the proposed method giving better SNR values and its performance is comparatively superior for both stationary and non-stationary signals.

6. References

1. Boll S. Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans Acoust, Speech, and Sig Proc. 1979; 27(2):113–20.
2. Benesty J, Makino S, Chen J editors. Speech Enhancement. Berlin: Springer; 2005.

3. Prasad G, Surendar S. A review of different approaches of spectral subtraction algorithms for speech. Enhancement Current Research in Engineering, Science and Technology (CREST) Journals. 2013; 1(2):57–64.
4. Spriet A, Moonen M, Wouters J. Spatially pre-processed speech distortion weighted multi-channel Wiener filtering for noise reduction. Signal Processing. 2004; 84(12):2367–87.
5. Alam MJD. Perceptual improvement of Wiener filtering employing a post-filter. Digital Signal Processing. 2010; 21(1):54–65.
6. Paliwal KK, Basu A. A speech enhancement method based on Kalman filtering. Proceedings of ICASSP'; Dallas, TX, USA. 1987; 7(5):510–24.
7. Ganesan V, Manoharan S. Surround noise cancellation and speech enhancement using sub band filtering and spectral subtraction. Indian Journal of Science and Technology. 2015 Dec; 8(33). DOI: 10.17485/ijst/2015/v8i33/77774.
8. Kallakunta RK, Bhavani L, Spandana P, Rishitha K, Ashitha G. Removal of real world noise in speech: Comparision of various parameters using kalman and h-infinity filter algorithms. Indian Journal of Science and Technology. 2016 Aug; 9(30).
9. Padmavathi G, Shanmugavel S. Performance analysis of cooperative spectrum sensing technique for low SNR regime over fading channels for cognitive radio networks. Indian Journal of Science and Technology. 2015 Jul; 8(16). DOI: 10.17485/ijst/2015/v8i16/64746.
10. Ramu SS, Umadevi S. Realization of low power FIR digital filter using modified DA-based architecture. Indian Journal of Science and Technology. 2015 Aug; 8(18). DOI: 10.17485/ijst/2015/v8i19/77161.
11. Hong D-H, Choi J-K, Kim G-J, Kim H-G. A study on the size of bore on Signal to Noise Ratio (SNR) in Magnetic Resonance Imaging (MRI). Indian Journal of Science and Technology. 2016 May; 9(20). DOI: 10.17485/ijst/2016/v9i20/94683.