International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045

ENHANCEMENT OF SPEECH INTELLIGIBILITY AND QUALITY IN HEARING AID USING FAST ADAPTIVE KALMAN FILTER ALGORITHM

R. Ramya Dharshini [1], R. Senthamizh Selvi [2], G.R. Suresh [3], S. Kanaga Suba Raja [4]

[1]Communication systems, Easwari Engineering College, Chennai,(India)
[2,3]Dept. of Electronics and Communication Engg., Easwari Engineering College, Chennai, (India)
[4]Dept. of Computer Science Engg., Easwari Engineering College, Chennai,(India)

ABSTRACT

A new signal-processing approach is proposed to enhance speech signal for listeners with hearing loss. The quality of such speech signals is improved using speech-enhancement techniques. Thereby, many filtering techniques such as conventional, Fast Adaptive, etc are used. This approach focuses on improving the quality and intelligibility of the speech signals using Fast Adaptive Kalman Filter (FAKF) algorithm. Therefore, the aim of the noise reduction algorithms is to estimate the clean speech signal from the noisy recordings in order to improve enhanced signal. The conventional Kalman filter is one of the time domain speech-enhancement techniques. It requires calculating parameters by AR (auto-regressive) model, and inverse matrix operation, is non-adaptive. FAKF is based on masking characteristics of human auditory system. It constantly updates the first value of state vector which automatically amend the estimation of environmental noise by observation data. Speech enhancement system (FAKF) consists of ability to update the estimated noise constantly. Based on the threshold level the current speech frame is determined, thereby reducing the execution time. Hence the complexity of matrix operations is reduced. From the simulation results, it has been observed that FAKF is more efficient in obtaining the clean speech signal and also provides best results for stationary noises. With the implementation of this speech enhancement algorithm, the objective measures shows the excellent quality even at 15dB input noise level resulting in output SNR of 41.6dB.

Keywords: ***Speech Enhancement; Fast Adaptive Kalman Filter (FAKF); Signal to Noise Ratio (SNR).***

I. INTRODUCTION

The goal of Speech enhancement is to empower speech quality by using several algorithms. It is one of the significant topics to enhance the performance of the systems with noise in speech signal processing. It has many applications like hearing aids, forensic applications, cellular environments, front-ends for speech recognition system, telecommunication signal enhancement, military, etc. In communication systems noise and distortions are the main limiting factors. Hence to sweep over these, their modeling and removal have been at the core of the theory and practice of communications and signal processing. Various techniques are modeled for this purpose to improve the speech signal-to noise ratio which depends on the performances in quality and intelligibility of the processed speech signal.

In hearing aids, the Kalman filter based noise suppression is used, as it depends on spatial information to classify the required noise and speech models. Basically for a Kalman filter to estimate the speech and noise model, the parameters is of fundamental importance. Identification of exact type of noise is slightly difficult and it is the effective application of the Kalman filtering technique. So we need to implement a real time adaptive algorithm to estimate the ambient noise. The Kalman filtering deals with random processes described using state-space modeling, it generate signals that can be measured and processed utilizing time recursive estimation formulas. The noise suppressed signal may deteriorate the quality of the speech signal based on accuracy of parameters estimated dependent on the auto regressive (AR) model. The Fast Adaptive Kalman Filtering algorithm can be categorized as follows:

1. Conventional (Matrix) Kalman Filtering algorithm.
2. Modified fast adaptive Kalman filtering algorithm.

A conventional Kalman filter is an optimal recursive data processing algorithm. One of the general aspects of optimality is that the Kalman filter incorporates all the informations provided to it. The different methods were proposed by [1]-[5]. In this process, the calculation of LPC (linear prediction coding) coefficient and inverse matrix increase the complexity of the filtering algorithm.

Assuming the system models is linear and acceptable up to some Gaussian additive noise. Hence with known variance constantly updates the first values of state vector which provides the elimination of the matrix operations. Therefore to reduce the complexity of the process of estimating noise, the modified fast adaptive kalman filtering is implemented. This algorithm adapts for any type of environmental noise and works better with stationary type of background noises [6]. It has been observed that the intelligibility of speech signal is much better in the proposed method than with the conventional method.

II. METHODOLOGY USED

A.Conventional Kalman Filter

Using a Kalman filter for speech enhancement asks for a state space model. An appropriate one that is often used in [13], assumes that the speech signal s(n) as well as the interference signal i(n) can be adequately modeled by Autoregressive (AR) processes of order p and q respectively. The excitation signals ζ(n) and η(n) are assumed to be independent zero mean white Gaussian noise with variance σ_ζ^2 and σ_η^2 respectively. A corresponding state space model with the state vector x(n) =$[s(n-p+1) \cdots s(n) \quad i(n-q+1) \cdots i(n)]^T$ can be given as

$$x(n)=A(n-1)\,x(n-1) + B\,u(n) \tag{1}$$

$$y(n)= C\,x(n) + v(n) \tag{2}$$

Where v(n) is the white, Gaussian measurement error with variance σ_v^2 and the input u(n) = $[\zeta(n)\eta(n)]^T$.

International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045

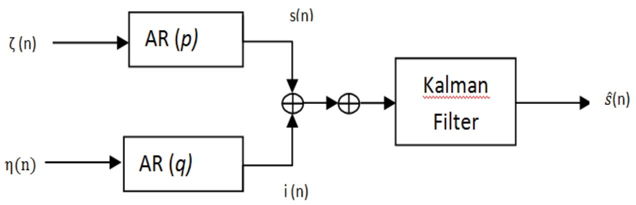


Figure 1: The speech / interference model

Note that the transition matrix computation is time invariant, while in fact, the parameters a_k and b_k may change at every time step n in [7] and [8]. This ability of the Kalman filter to deal with a time variant signal and speech model is essential for the use of the Kalman filter instead of a Wiener filter. Since in real world applications the input u(n) is unknown, one will consider it to be zero. Based on this implication one will have an uncertainty in the state vector x(n). The covariance matrix $Q_W(n)$ of the corresponding state error can be calculated as follows:

$$Q_W(n) = BE\{u(n)u^T(n)\}\, B^T = B\begin{bmatrix}\sigma_\zeta^2(n) & 0\\ 0 & \sigma_\eta^2(n)\end{bmatrix}B^T \tag{3}$$

Based on this state space model, a Kalman filter can be used to estimate the state vector x(n) based on the noisy measurements y(k) (k up to n). This estimate $\hat{x}$(n) as given in [9]. K(n) is the Kalman gain vector and I is the identity matrix of order p+q.

$$K(n) = \frac{\hat{P}(n|n-1)C^T}{C\hat{P}(n|n-1)C^T+Q_v(n)} \tag{4}$$

The estimated speech signal $\hat{s}$(n) can be found at the pth position of the estimated state vector $\hat{x}$(n|n). Note that because of the special structure of the vector x(n), one will estimate not only s(n) but also s(n−1)···s(n−p + 1). Since these estimates are all based on measurements y(k) with k up to n, they correspond to fixed-lag estimates [7].

B.Fast Adaptive Kalman Filter

In FAKF the process noise and measurement noise can be estimated on-line according to the measured value and filtered value, with tracking changes of noise in real time to amend the filtering parameters, thereby improving the filter effect. This is obtained by setting a reasonable threshold in adaptive method which is used to evaluate the current speech frame is noise or not. It mainly consists of two steps:

(a) Update the variance of environmental noise by [11],

$$R_v(n) = (1-d)\times R_v(n) + d\times R_u(n) \tag{5}$$

In (5) d is the loss factor and it is given by

$$d = \frac{1-b}{(1-b^{t+1})} \tag{6}$$

b is a constant and it is assumed as 0.99.

International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045

Here before updating the variance of environmental noise, the variance of current speech frame $R_u(n)$ is compared with the threshold U (7). If $R_u(n)$ is less than or equal to U, the current speech frame can be considered as noise and then the algorithm will reestimate the noise variance.

$$U = (1 - d) \times U + d \times R_u(n) \tag{7}$$

(b) Obtaining the SNR

When the noise is very large in (7), the updating threshold cannot be used directly thereby the SNRs are to be calculated by determining the variance of pure speech signal, variance of input degraded speech signal, and variance of background noise. According to [12], the two SNRs are calculated, one for current speech frame $SNR_I(n)$ and another for whole speech signal $SNR_o(n)$ and are compared. The speech frame is noise, when $SNR_I(n)$ is $\leq SNR_o(n)$. If $SNR_I(n)$ is $> SNR_o(n)$ the noise estimation will be attenuated to avoid damaging of speech signals. According to [13], noise attenuation can be expressed as

$$R_v = \frac{R_v(n)}{1.2} \tag{8}$$

Updating Amplitude Threshold

Along with the above two steps, an extra condition is set by determining an amplitude threshold for the noisy speech signal. This condition is sufficient to have the updating of background noise constantly. Here an amplitude threshold Z of 0.008 is set for the noise samples taken for simulation. Here all the sample values less than 0.008 has been considered as pure noise and then been processed such that their enhanced amplitude is 0. The K(n) is the Kalman gain vector is calculated as follows

$$K(n) = \frac{R_S(n)}{R_S(n) + R_v(n)} \tag{9}$$

$$S(n) = K(n)Xy(n) \tag{10}$$

Where $R_s(n)$ =Variance of noisy speech, $R_v(n)$ =Variance of noise, $\delta_r^2(n)$ =Varianceof speech frame, *K(n)* = Kalman gain, *S(n)* = Enhanced speech.

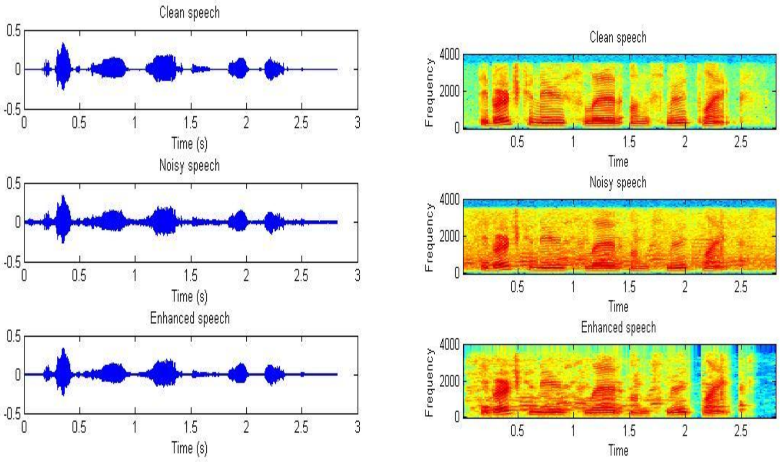
III. RESULT AND DISCUSSION

A. Objective Evaluation

The significant standardization efforts have been made by the International Telecommunications Union (ITU-T) to standardize both intrusive and nonintrusive algorithms using NH listeners and CIs users. On the other hand, only a handful of algorithms have been proposed that are specifically tuned to assistive listening devices. In the following sections, the choice of measures used was guided only by the applicability to the task in HA, but also by the availability of publicly available source code i.e., code that could be licensed at a reasonable cost.

International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045

The performance evaluation of the proposed system is tested by using NOIZUES speech corpus's database. This database contains IEEE sentences produced by 3 male and 3 female speakers and was corrupted by 8 different real time noises at various levels of SNR at the input level to the Hearing Aid. Using AURORA database as input for noisy signals, it also include the recordings from different places: car, restaurant, babble (crowd of people), exhibition hall, street, airport, train, station, and train. Such stationary noise signals were added to the clean speech at SNRs of 0, 5, 10, and 15dB. The clean signal suppressed by different noises is given as input to the Hearing Aid at various Signal to Noise Ratio (SNR) levels.



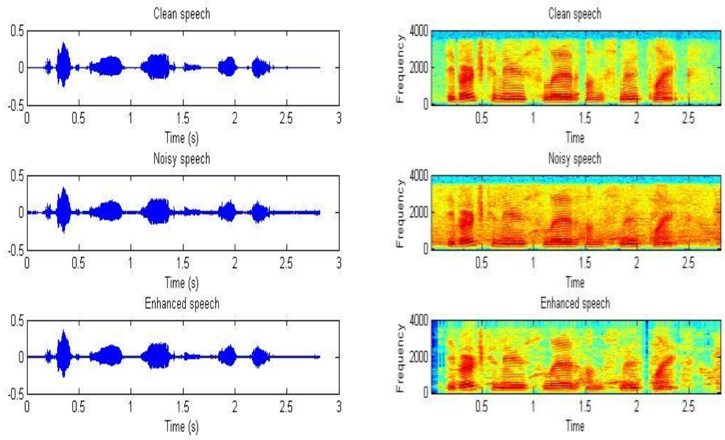
2(a) **2(b)**

Fig 2 Waveform and Spectrogram of clean speech, noisy signal, FAKF output and Enhanced Speech of 10dB restaurant noise

Figure 2 represents the Waveforms and Spectrograms of FAKF approach of 10dB restaurant noise. Figure 2(a) shows the waveform of input Clean Speech Signal, Noisy Signal and the Enhanced Speech Signal respectively. The waveform of Noisy signal shows the harmonic part but still there is a tradeoff between residual noise suppression and spectral distortion, which degrades the speech signal. The FAKF process is done to filter part of noisy signals in HA. The reduction in noisy part of the signal can be seen from the enhanced waveform.

Figure 2(b) shows the Spectrogram of the input Clean Speech Signal, Noisy Signal and the Enhanced Speech Signal respectively. The spectrogram of Noisy signal clearly displays the energy level of the harmonic part of signal corrupted by noise. It generally represents the corrupted signal with the restaurant noise of 10dB SNR. For this corrupted signal of 10dB input SNR is run through the FAKF algorithm, the output SNR is **30.7314dB.** It can be seen clearly from the spectrogram of Enhanced Speech Signal that the harmonic part of the signal has high energy level. This harmonic pattern represents that the noisy signal are reduced drastically.

International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045



3(a) 3(b)

Fig 3 Waveform and Spectrogram of clean speech, noisy signal, FAKF output and Enhanced Speech of 15dB restaurant noise

Figure 3 represents the Waveforms and Spectrograms of FAKF approach of 15dB restaurant noise. Figure 3(a) shows the waveforms and spectrograms of input Clean Speech Signal, Noisy Signal and Enhanced Speech Signal respectively. The spectrogram of Noisy signal shows the visible harmonic part is highly corrupted by noise of 15dB. The waveform of the Enhanced speech signal represents the processed signal of the Fast Adaptive Kalman Filtering algorithm. The enhanced waveform represents the drastically reduction of noisy part in the signal.

By comparing with the existing systems, FAKF shows better visible harmonic parts of the processed speech signal. Using FAKF algorithm, the output SNR is **41.6105dB** for restaurant noise with 15dB. It also reduces the processing time as well as the computational complexity of system.

B. Algorithmic Parameters

The intrusive and nonintrusive algorithm metric are being used in NH listeners and CIs users speech signal processing prediction. The intrusive intelligibility metric consists of two indices, i.e., SNR (Speech Signal-to noise ratio) and PESQ (Perceptual Evaluation Speech Quality).

1. SNR Estimation

Signal-to-Noise Ratio (SNR) is one of the oldest and widely used objective measures. It requires both the distorted signal and undistorted (clean) speech samples for the mathematical calculation. SNR can be calculated as follows

$$SNR = 10log_{10}\frac{\sum_{n=1}^{N} x^2(n)}{\sum_{n=1}^{N}\{x(n)-\hat{x}(n)\}\wedge 2} \quad (11)$$

Where, *x(n)* is the clean speech, *xˆ(n)* the distorted speech, and *N* the number of samples. This classical definition of SNR is known to be not well related to the speech quality for a wide range of distortions. Thus, the classical SNR exists with several variations which show much higher correlation with subjective quality. The classical SNR typically does not correlate with speech quality even though speech is not a stationary signal; it generally averages the ratio over the entire signal. Speech energy fluctuates over time hence the portions in which the speech energy is large and noise is relatively inaudible; should not be washed out by other portions where speech energy is small and noise can be heard over speech. Thus, SNR was calculated in short frames and averaged. This measure is called the segmental SNR, and can be defined as

$$SNR_{seg} = 10log_{10} \frac{10}{M} \sum \frac{\sum_{n=1}^{N} x^2(n)}{\sum_{n=1}^{N} \{x(n) - \hat{x}(n)\}\wedge 2} \quad (12)$$

The fwSNRseg can be defined as follows

$$fwSNR_{seg} = \frac{10}{M} \sum_{m=0}^{M-1} \frac{\sum_{j=0}^{K-1} W(j,m) log_{10} \frac{X(j,m)^2}{\{X(j,m) - \hat{X}(j,m)\}\wedge 2}}{\sum_{j=0}^{K-1} W(j,m)} \quad (13)$$

Where *W(j,m)*, weight on j^{th} sub band in the m^{th} frame, *K* is the number of sub bands, *X(j,m)* is the spectrum magnitude of the j^{th} sub band in the m^{th} frame, and ˆ*X(j,m)* its distorted spectrum magnitude.

TABLE I PERFORMANCE OF FAKF- OUTPUT SNR VERSES INPUT SNR AT DIFFERENT NOISE LEVEL

Noise Type	Input SNR	Output SNR FAKF
Street Noise	10dB	20.4481
	15dB	38.3008
Car Noise	10dB	15.4855
	15dB	34.1267
Restaurant Noise	10dB	30.7314
	15dB	**41.6105**
Train Noise	10 dB	21.6158
	15dB	34.5290
Babble Noise	10 dB	27.1963
	15dB	39.9932

Hence the developed system provides satisfactory performance for various SNR levels of different noises. Table I shows the SNR in dB for different noises in database.

IV. CONCLUSION

A Speech Enhancement method which is more adaptable for stationary noise using Fast Adaptive Kalman Filter is developed. It reduces the matrix operation by using coefficient factor and also it gives better human auditory characteristics. The simulation results show that FAKF is more efficient in obtaining the clean speech signal.

The evaluation parameter such as SNR shows better results at various noise levels. Using FAKF the maximum SNR value obtained is **41.6105dB** for the input noise level of 15dB along with less running time. The above result shows that proposed system outperforms the existing system. It is concluded that the proposed method can be easily realizable as well as provides good noise suppression without sacrificing the quality of the speech signal.

REFERENCES

[1] ZHANG Xiu-zhen, FU Xu-hui, WANG Xia, "Improved kalman filter method for speech Improvend enhancement", Computer Applications, Vol.28, enhancement.Computer pp.363-365, Dec 2008.

[2] Nari tanabe, Toshiniro Furukawa, Shigeo Ts,ujji, "Fast noise Suppression algorithm with Kalman filter Theory" , Second International Symposium on Universal Communication, pp.411-415, 2008.

[3] Nri Tanabe, Toshiniro Furukawa, Hideaki Matsue and Shigeo Tsujji, "Kalman Filter for Robust Noise Suppression in White and Colored Noise" IEEE International Symposiumon Circiuts and Systems 2008.1172-1175.

[4] WU Chung-ling, HAN Chong-zhao, "Square-root Quadrature Filter", Acta Electronica Sinics,Vol.37, No.5, pp.987-992, May.2009.

[5] SU Wan-Xin, HUANG Chun-mei LIU Peiwei, MA Ming-long. Application of adaptive Kalman filter in intial alignment of inertial navigation system. Journal of Chinese Inertial Technology Vol.18, No.1, pp. 44-47,Feb.2010.

[6] Prabhavathi CN, Ravikumar KM, " Acoustic noise classification and characterisation using statistical properties" ,IJETAE, 2014;4 6:638.

[7] John B. Moore, "Discrete-Time Fixed-Lag Smoothing Algorithms", Pergamon Press, 1973.

[8] Ramun Schmid, "Kalman LOCO", HSR Hochschule fuer Technik Rapperswil, 2010 http://www.medialab.ch/EUSIPCO2010/.

[9] Monson H. Hayes, "Statistical Digital Signal Processing and Modeling", John Wiley & Sons, 1996.

[10] P. R. Bellanger, "Estimation of noise covariance matrices for a linear time varying stochastic process", Automatica, Vol. 10, pp. 267-275 (1974).

[11] Mai Q, He D, Hou Y, Huang Z. A fast adaptive Kalman filtering algorithm for speech enhancement. In: IEEE International Conference on Automatic Science and Engineering Trieste, Italy, 2011.

[12] Gao Y, Zhang J. Kalman filter with wavelet based unknown measurement noise estimation and its application for information fusion. Acta Electronica Sinica 2007;35 1:108–111.

[13] R. K. Mehra, "Approaches to adaptive filtering", IEEE Trans. on Automatic Control, Vol. 17, pp. 693-698 (1972).

[14] B. Carrew and P. Bellanger, "Identification of optimum filter steady state gain for systems with unknown noise parameters", IEEE Trans. Automatic Control, Vol. 18, N.6, pp. 582-587 (1973).

[15] R. K. Mehra, "On the identification of variance and adaptive Kalman filtering" , IEEE Trans. on Automatic Control, Vol. 15, N.2, pp. 175-184 (1970).

[16] P. Sorqvast, P. Handel, B. Ottersten, "Kalman filtering for low distortion speech enhancement in mobile communication" in IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997.

[17] V. Grancharov, J. Samuelsson, B. Kleijn, "On Causal Algorithms for Speech Enhancement" in IEEE Transactions on Audio, Speech, and Language Processing, 2006.

[18] Sage, A.P. and Husa,G.W. (1969a).Adaptive filtering with unknown prior statistics. In Proceedings of joint automatic control conference, Boulder, 760–769.

[19] Sage, A.P. and Husa, G.W. (1969b). Algorithms for sequential adaptive estimation of prior statistics. In Adaptive Processes (8th) Decision and Control, 1969 IEEE Symposium on, 61 –70.

[20] Bavdekar, V.A., Deshpande, A.P. and Patwardhan, S.C. (2011). Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter. Journal of Process Control, 21(4), 585–601.