Noise Free Speech Enhancement based on Fast Adaptive Kalman Filtering Algorithm

N. S. Banale
MBES College of Engineering,
Ambajogai.

S. K. Sudhansu
MBES College of Engineering,
Ambajogai

S. M. Jagde, Ph.D
College of Engineering,
Osmanabad.

ABSTRACT

Speech enhancement is the process of eliminating noise and increasing the quality of a speech signal, which is contaminated with other kinds of distortions. Conventional Kalman filtering is known as an effective speech enhancement technique, in which speech signal is usually modeled as autoregressive (AR) model and perform a lot of matrix operations. In this paper we proposed a fast adaptive algorithm in presence of environment noise which eliminates the matrix operations and reduces the calculating time by only constantly updating the first value of state vector X(n). To evaluate the system performance we employed the calculation of SNR. Simulation results show that the fast adaptive algorithm using Kalman filtering is effective for speech enhancement.

General Terms

Kalman filter algorithms, Speech enhancement.

Keywords

Kalman filter algorithm, SNR (Signal to Noise Ratio), LPC.

1. INTRODUCTION

The goal of a speech enhancement algorithm is to reduce or eliminate background noise without distorting the speech signal. Kalman filtering is known as an effective speech enhancement technique, in which speech signal is usually modeled as autoregressive (AR) process with the LPC (linear prediction coding) coefficient [13]. Many filtering techniques have been proposed by [1]-[5] to study the noise problem in the speech signal. In this process, the calculations of LPC (linear prediction coding) coefficient and inverse matrix greatly increase the computational complexity of the filtering algorithm. Simple Kalman filtering techniques are proposed by [3] and [4] without calculating LPC in AR model, but they still has a large number of matrix inverse operations and redundant data. To overcome the drawback of the conventional Kalman filtering a fast adaptive Kalman filtering is proposed. This algorithm constantly updates the first value of state vector, which eliminates the matrix operations and reduces the time complexity of the algorithm. Simulation results show that, the fast adaptive algorithm of Kalman filtering is more effective than the conventional and perceptual Kalman filtering algorithm. At the same time, it reduced its running time without sacrificing quality of the speech signal.

2. KALMAN FILTERING ALGORITHM FOR SPEECH ENHANCEMENT

2.1 Conventional Kalman Filtering Algorithm

Speech can be represented by an autoregressive (AR) process which is essentially the output of an all-pole linear system driven by white noise sequence. Thus speech signal at n-th time instant, s(n) is given by:

$$s(n) = \sum_{i=1}^{L} \alpha_i(n) \times s(n-i) + \omega(n) \qquad (1)$$

Equation (1) can be expressed as the state equation which is given by

[State equation]

$$X(n) = F\,X(n-1) + Gu(n) \qquad (2)$$

Where

1. The sequences $u(n)$ is the Gaussian white noise whose mean is zero and the variances is δ^2_u .
2. $X(n)$ is the L× 1 state vector

$$X^T(n) = [s(n-L+1)\ldots s(n-1), s(n)] \qquad (3)$$

3. $F(n)$ is the L × L transition matrix

$$F = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \vdots & \cdots & 1 \\ a_L(n) & a_{L-1}(n) & a_{L-2}(n) & \cdots & a_{1(n)} \end{bmatrix} \qquad (4)$$

4. G is the L×1 input vector which is defined as follows:

$$G = [0\ 0\ .\ .\ .\ .\ .\ 1] \qquad (5)$$

When only the noise corrupted signal y(n) is available, the observation process can be written in the following form:

$$y(n) = s(n) + v(n) \qquad (6)$$

This equation can be written in the matrix form as follows:

$$y(n) = H\,x(n) + v(n) \qquad (7)$$

Where

1. The sequences $v(n)$ is Gaussian white noise whose mean is zero and the variances is δ^2_v
2. $x(n)$ is the state vector already defined by equation (4)
3. H is the observation matrix given by:

$$H = G^T = [0\ 0\ .\ .\ .\ .\ 0\ 1] \qquad (8)$$

The conventional Kalman filtering is using the LPC coefficient to estimate the observations of the speech signal [6]. This part require half the time of the whole algorithm. In [2] the transition matrix F and the observation matrix H are modified. They has defined as

$$F=H=\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \vdots & \cdots & 1 \\ a_L(n) & a_{L-1}(n) & \cdots & \cdots & a_{1(n)} \end{bmatrix} \quad (9)$$

The L×1input vector $Q(n)^T$= [s(n) 0 …0] and the 1×□L observation vector R(n)=[1,v(n), …,v(n-L+2)]. Finally, (3) and (4) can be rewritten into the matrix equations□ by□□□□□

[State equation]

X(n)= F×X(n-1)+Q(n) (10)

[Observation equation]□

Y(n)= H ×X(n) + R(n)

The Kalman filtering algorithm provides the updating state vector estimator equations

[Initialization]
$X(0 \mid 0) = 0$, $P(0 \mid 0) = I$
$R_v(n) = \delta_v^2$, G = [0 0 …1]

$$R_s(n)[i,j]=\begin{cases} E\big(Y(n)\times Y(n)\big)-\delta_v^2 & (i,j=1) \\ 0 & (others) \end{cases}$$

[Iteration]

$e(n) = y(n) - H \times X(n/n-1)$ (12)

$P(n/n-1) = F \times P(n-1/n-1) \times F^T + G \times R_s(n) \times G^T$ (13)

$K(n) = P(n-1/n-1) \times H^T \times [G \times P(n/n-1) G^T + R_s(n)]^{-1}$ (14)

$X(n/n-1) = F \times X(n-1/n-1)$ (15)

$X(n/n) = X(n/n-1) + K(n) \times e(n)$ (16)

$P(n/n) = [I - K(n)H] \times P(n/n-1)$ (17)

Where

1. $X(n/n-1)$ is the minimum mean-square estimate of the state vector X(n) given the past $n-1$ observatio$n$s y(1) . . . , $y(n-1)$
2. $X(n/n-1) = X(n) - X(n/n-1)$ is the predicted state-error vector
3. $P(n/n-1) = E[x(n/n-1)X^T(n/n-1)]$ is the predicted state- error correlation matrix
4. $X(n/n)$ is the filtered estimate of the state vector x(n)
5. $X(n/n) = X(n) - X(n/n)$ is the filtered state-error vector
6. P(n/n) = $E[X(n/n) \times X^T(n/n)]$ is the filtered state error correlatio$n$ matrix
7. e(n) is the innovation sequence
8. K(n) is the Kalman gain

Here the number of iterations is equal to the number of speech signal sampling points. Hence there exists time complexity.

2.2 Perceptual Kalman Filtering Algorithm

From the equations of the conventional Kalman filtering, one can find large number of matrix operations which leads to an increase in the computational complexity. Hence the complexity of the algorithm is reduced to maximum extent. To avoid the matrix inverse operation in the improved filtering method, calculation of s(n) is alone done. The estimated speech signal can be retrieved from the state vector estimator

s(n) =H×X(n/n)

The recursive equations of the perceptual Kalman filtering are defined as:

[Initialization]

s(0)=0 , $R_v(n) = \delta_v^2$, R_s(n)= E($y(n) \times y(n)$) – $\delta_v{}^2$

[Iterations]

$K(n) = R_s(n) / (R_s(n) \times R_v)$ (18)

$s(n) = K(n) \times y(n)$ (19)

2.3 Fast adaptive Kalman filtering algorithm

There are always noise changes with the surrounding environment. To design the fast adaptive Kalman filtering algorithm, we need to know information about environmental noise for that it is necessary to constantly update the estimation of noise. Here in the fast adaptive Kalman filtering algorithm, it can constantly update the estimation of background noise and update the threshold U. So it consists of two steps.

1) Updating of variance of noise is obtained by
$R_v(n) = (1-d) \times R_v(n) + d \times R_u(n)$ (20)
Where d is the loss factor which limits the memory of the filtering. As proposed by [3]'d' is defined as:

$$d=\frac{1-b}{1-b^{t+1}} \quad (21)$$

Where b is a constant whose value ranges in between 0.95 and 0.99. The loss factor d is used to reduce the error.

2) Updating of the threshold is known from
$U = (1-d) \times U + d \times R_u(n)$ (22)

Here the loss factor d is used which can reduce the error. We always compare the $R_u(n)$ [variance of the current speech frame] with the threshold U which is updated above. We calculate the SNR of current speech frame and the whole speech signal and compare them.

$$SNR_1(n) = 10log_{10}\left(\frac{\delta_r^2(n) - \delta_v^2(n)}{\delta_v^2(n)}\right) \quad (23)$$

$$SNR_0(n) = 10log_{10}\left(\frac{\delta_x^2(n) - \delta_v^2(n)}{\delta_v^2(n)}\right) \quad (24)$$

Where δ_r^2 □is the variance of pure speech signals δ_x^2□is the variance of the input noise speech signals, δ_v^2□is the variance of background noise and n is the number of frames of the speech signal.

When the current speech frame $\mathrm{SNR}_1(n)$ is less than or equal to whole speech frame $\mathrm{SNR}_0\,(n)$ or SNR_0 (n) is less than zero then the speech frame is noisy and it can follow $R_u(n)$<=U. However, when $\mathrm{SNR}_1\,(n)$ is larger than $\mathrm{SNR}_0\,(n)$ then the noise estimation will be attenuated to avoid damaging the speech signals. According to [2], this attenuation can be expressed as,

$$R_v(n) = R_v(n)/1.2 \qquad (25)$$

The whole algorithm for A Fast Adaptive Kalman Filter is as follows:

```
[Initialization] s(0) =0, $R_v(1) = \delta_v^2(1)$
[Iterations]
If $SNR_1(n)$ <=$SNR_0(n)$ || $SNR_0(n)$ <0 then
    If $R_u(n) \leq U$ then
        1) $R_v(n)= (1-d)\times R_v(n)+d\times R_u(n)$
    End
        2) $d=(1-b)/(1-b^2)$
        3) $U=(1-d)\times U+d\times R_u(n)$
    Else
        4) $R_v(n)= R_v(n)/1.2$
    End
        5) $R_s(n)= E(y(n)\times y(n))- R_v(n)$
        6) $K(n)= R_s(n)/ (R_s(n)+R_v(n))$
        7) $s(n) = K(n) \times y(n)$.
```

3. MATLAB SIMULATION

The comparison between the Conventional, Perceptual and Fast Adaptive Kalman Filtering Algorithm are done using MATLAB. We adopted two patterns of the noisy speech as the signal samples for the simulations. One is the female speech corrupted with a background noise, and the other is the male speech signal corrupted with a background noise. Figure shoes the simulation samples of female and male speech and the different filtering techniques results. In this section, we will compare the simulation result in the 2 different phases: (a) accuracy and (b) running time. The variance of the noise is assumed to be known and to compare the accuracy of the filtering methods the SNR is calculated. Table 1 shows fast adaptive Kalman filter achieves high performance filtering accuracy.

Table 1. Signal To Noise Ratio for male and female speech

	Signal To Noise Ratio		
	Conventional	Perceptual	Fast Adaptive
FEMALE	9.066	14.207	20.77
MALE	0.762	3.079	3.8462

Table 2 shows the comparison of the running time of the fast adaptive Kalman filter with the conventional and the perceptual filter. It is clear that the improved filtering method can achieve higher performance noise suppression capability and reduce its' running time without sacrificing the high quality of the speech signal.

Table 2. Running Time for male and female speech

	Running Time(Sec)		
	Conventional	Perceptual	Fast Adaptive
FEMALE	0.183	2.67	0.000974
MALE	0.141	2.34	0.000761

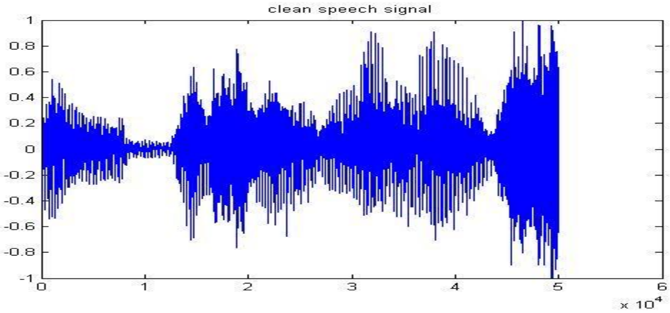

Fig 1: Clean male speech signal

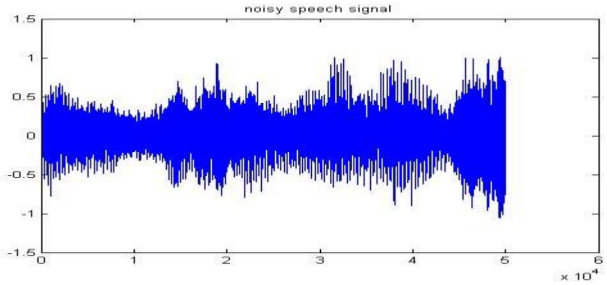

Fig 2: Noisy male speech signal

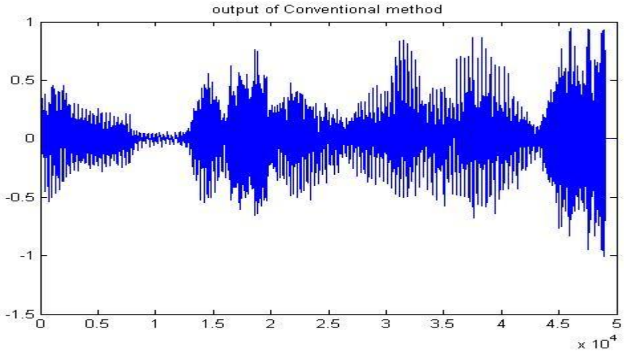

Fig 3: Output of conventional Kalman filter

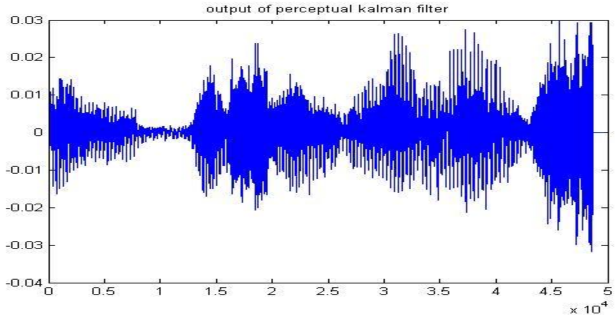

Fig.4. Output of perceptual Kalman filter

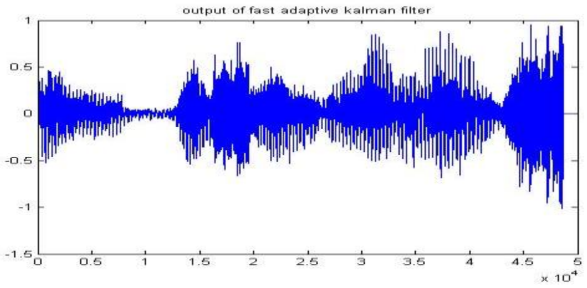

Fig 5: Output of fast adaptive Kalman filter

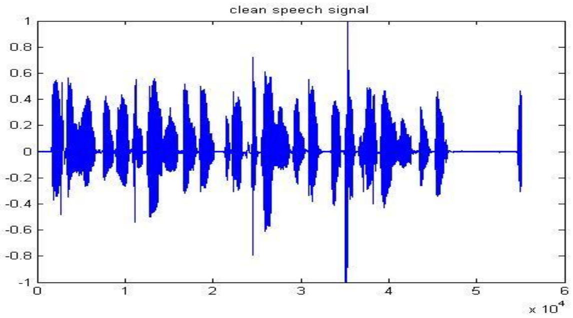

Fig 6: Clean female speech signal

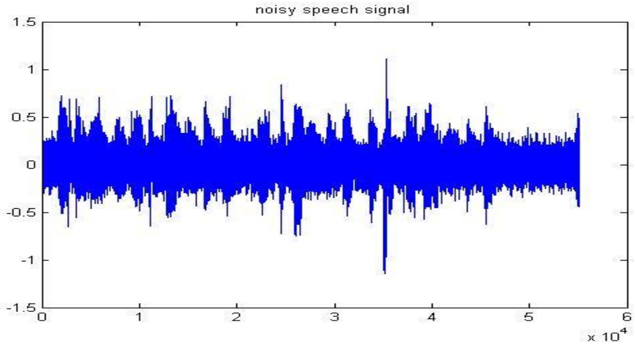

Fig 7: Noisy female speech signal

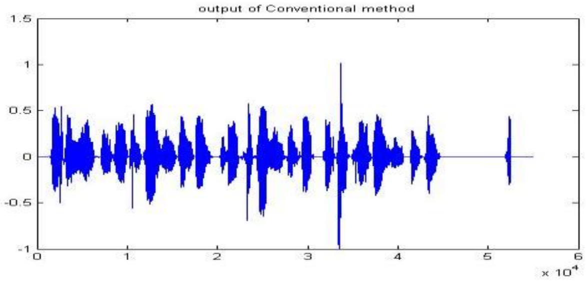

Fig 8: output of conventional Kalman filter

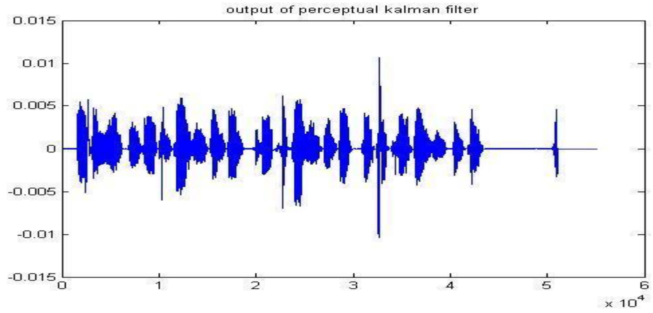

Fig 9: Output of perceptual Kalman filter

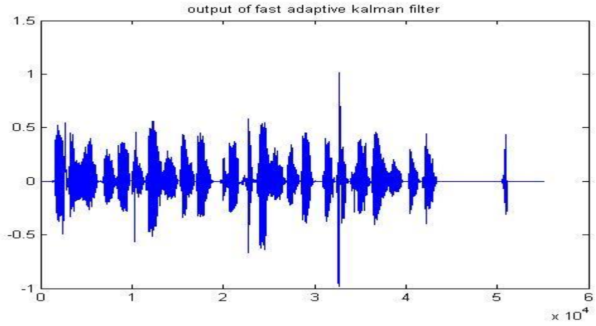

Fig 10: Output of fast adaptive Kalman filter

3. CONCLUSION

This paper has presented a noisy speech enhancement based on fast adaptive Kalman filtering algorithm. It is concluded that this proposed algorithm is simpler and can realize the good noise suppression by reducing the computational complexity without sacrificing the quality of the speech signal. In the further study, it can be improving the adaptive algorithm and will be applied to the embedded-speech-recognition system at the hardware level.

4. REFERENCES

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

[2] Quanshen Mai, Dongzhi He, Yibin Hou, Zhangqin Huang, "A Fast Adaptive Kalman Filtering Algorithm For Speech Enhancement", *IEEE International Conference on Automation Science And Engineering*, pp 327-332, 2011.

[3] WU Chun-ling, HAN Chong-zhao Square-Root Quadrature Filter. Acta Electronica Sinica, Vol.37, No.5, pp.987-992, May.2009.

[4] Marcel Gabrea, "Kalman Filter-Based Single Microphone Noise Canceller," *International Workshop on Acoustic Echo and Noise Control, Sept. 2003, Kyoto.*

[5] Adly A. Girgis, David G. Hart, Member, IEEE, "Implementation Of Kalman And Adaptive Kalman Filtering Algorithms For Digital Distance Protection On A Vector Signal Processor", *IEEE Transactions on Power Delivery*, Vol.4, No.1, pp 141-156, 1989.

[6] Fang Deng, Jie Chen, and Chen , " Adaptive unscented Kalman Filter for parameter and state estimate of nonlinear high speed objects", *Journal of Systems*

Engineering and Electronics, Vol. 24, No. 4, 2013, pp.655–665.

[7] Ali Almagbile, Jinling Wang, and Weidong Ding," Evaluating the Performances of Adaptive Kalman Filter Methods in GPS/INS Integration", *Journal of Global Positioning Systems*, Vol.9, No.1, 33-40, 2010.

[8] K. K. Paliwal and A. Basu, "A Speech Enhancement Method Based on Kalman Filtering", *in Proc. ICASSP'87*, pp. 177–180.

[9] Michel Verhaegen , Paul Van Dooren, Member, IEEE, "Numerical Aspects Of Different Kalman Filter Implementations", *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, VOL. AC-31.No.10 ,pp 907-917, 1986.

[10] Maurice G. Bellanger, "Adaptive Digital Filters", Marcel Dekker, pp 416-419, 1987.

[11] Douglas, S.C. "Introduction to Adaptive Filters", *Digital Signal Processing*, CRC Press LLC, 1999.

[12] Angus P. Andrews, Mohinder S. Grewal , "Kalman Filtering Theory and Practice Using MATLAB" , John Wiley & Sons , pp 131-148, 2008