A Robust Speech Enhancement By Using Adaptive Kalman Filtering Algorithm

G. SANDEEP
M.Tech Student Scholar, DECS,
Dept of Electronics and Communication Engineering,
Nalanda Institute of Engineering and technology,
Sattenapalli (M); Guntur (Dt); A.P, India.

K.SAMPRASAD
M.Tech , Asst Professor
Dept of Electronics and Communication Engineering,
Nalanda Institute of Engineering and technology,
Sattenapalli (M); Guntur (Dt); A.P, India

Abstract-
Speech enhancement aims to improve speech quality by using various algorithms. The objective of enhancement is improvement in intelligibility and/or overall perceptual quality of degraded speech signal using audio signal processing techniques. Enhancing of speech degraded by noise, or noise reduction, is the most important field of speech enhancement. In this paper a Robust speech enhancement method for noisy speech signals presented in speech signals, which is based on improved Kalman filtering. By using Kalman filtering arise some drawbacks to overcome to modified the conventional Kalman filter algorithm. conventional Kalman filter algorithm needs to calculate the parameters of AR(auto=regressive model),and perform a lot of matrix operations, which is generally called as non- adaptive .In this paper we eliminate the matrix operations and reduces the computational complexity and we design a coefficient factor for adaptive filtering, to automatically amend the estimation of environmental noise by the observation data. Experimental results shows that the Proposed technique effective for speech enhancement compare to conventional Kalman filter.

Keywords- **Kalman Filter, Speech Enhancement.**

I. INTRODUCTION

The Kalman filter was created by Rudolf E. Kalman in 1960, though Peter Swerling actually developed a similar algorithm earlier. The first papers describing it were papers by Swerling (1958), Kalman (1960) and Kalman and Bucy (1961). It was developed as a recursive solution to the discrete-data linear filtering problem. Stanley Schmidt is generally credited with the first implementation of a Kalman filter. The background noise is a dominant source of errors in speech recognition systems. Noise reduction for speech signals has therefore application in entry procedures of those systems[1]. The Kalman filter is known in signal processing for its efficient structure. There are many studies of using of Kalman filtering for noise reduction in speech signals. Speech signals are modeled as stationary AR process. Modeling and filtering noisy speech signals in the sub band domain . Since the power spectral densities (PSD's) of sub band speech signals are flatter than their full band signals, low-order AR models are satisfactory and only lower-order Kalman filters will be required. In the next focus on first-order modeling. In this case the Kalman filter involves only scalar operations and computational amount is saved. For identification of AR parameters use the second Kalman filter which cooperates with the first one. The Kalman filter in this situation converges more rapidly than an adaptive filter with NLMS algorithm .It means increasing segmental SNR results in speech enhancement, thus visibly improving the quality of the speech signal[2].

The purpose of this note is to give a practical introduction into Kalman filtering and one of its byproducts, the Ensemble Kalman Filter The filter may be summarized as follows: It is a set of equations to estimate the state of some process, with the possibility of assimilating new information as it arrives. The filter is optimal in a sense that it incorporates all recently acquired information in the best possible way. The information arrives typically when we measure the state variables of the process provide an overview of the optimal linear estimator, the Kalman filter. This will be conducted at a very elementary level but will provide insights into the underlying concepts. As we progress through this overview, contemplate the ideas being presented: try to conceive of graphic images to portray the concepts involved (such as time propagation of density functions), and to generate a logical structure for the component pieces that are brought together to solve the estimation problem.

A Kalman filter is simply an optimal recursive data processing algorithm . There are many ways of defining optimal , dependent upon the criteria chosen to evaluate performance. It will be shown that, under the assumptions to be made in the next section, the Kalman filter is optimal with respect to virtually any criterion that makes sense. One aspect of this optimality is that the Kalman filter incorporates all information that can be provided to it. To overcome the drawback of conventional Kalman filtering for speech enhancement, we propose a Robust algorithm of Kalman filtering. This algorithm only constantly updates the first value of state vector X(n), which eliminates the matrix operations and reduces the time complexity of the algorithm. Actually, it is difficult to know what environmental noise exactly is. And it affects the application of the Kalman filtering algorithm[3]. So we need a real-time adaptive algorithm to estimate the ambient noise. We add the forgetting factor which has been mentioned by [4] and [5] to amend the estimation of environmental noise by the observation data automatically, so the algorithm can catch the real noise. Simulation results show that, compared with the conventional Kalman filtering algorithm, the robust algorithm of Kalman filtering is more effective.

II. IMPROVED KALMAN FILTERING ALGORITHM

A. Conventional Kalman Filtering Method:

Speech driven by white noise is All-pole linear output from the recursive process. Under the short-time stable supposition, a pure speech can establish L step AR model by

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

(1)

In (1) $a_i(n)$ is the LPC coefficient, ω(n) is the white Gaussian noise which the mean is zero and the variance is. In the real environment, the speech signal s(n) is degraded by an additive observation noise v(n). Its mean is zero, and its variance is δ_v^2.This noise isn't related to s(n) . A noisy speech signal y(n) is given by

y(n) = s(n) + v(n) (2)

In this paper, it is assumed that the variance δ_v^2 is known, but in practice we need to estimate it by the "silent segment" included in the y(n). (1) and (2) can be expressed as the state equation and the observation equation which are given by

[State equation]

$$x(n)=F(n)x(n-1)+G\omega(n)$$

(3)

[Observation equation]

$$y(n)=Hx(n)+v(n)$$

(4)

$$F(n)=\begin{bmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 1\\ a_L(n) & a_{L-1}(n) & a_{L-2}(n) & \cdots & a_1(n)\end{bmatrix}$$

(5)

F(n) is the L X L transition matrix expressed as G is the input vector and H is the observation vector. It is easy to see that the conventional Kalman filtering is using the LPC coefficient to estimate the observations of the speech signal. This part spends half the time of the whole algorithm.

In [2] the transition matrix F and the observation matrix H are modified. They has defined as

$$F=H=\begin{bmatrix}0 & 0 & 0 & \cdots & 0\\ 1 & 0 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1 & 0\end{bmatrix}$$

(6)

It also has defined the L×1 state vector X(n) =[s(n) ....s(n-L+1) s(n-L+2)] , the L×1input vector Q(n = [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]

$$X(n)=F\times X(n-1)+Q(n)$$

(7)

[Observation equation]

$$Y(n)=H\times X(n)+R(n)$$

(8)

Then the recursion equation of Kalman filtering algorithm is given in Table I. In this case the noise variance δ_v^2 is known.

TABLE I

THE CONVENTIONAL METHOD PROCEDURE

[Initialization]

$X(0\,|\,0)=0$, $P(0\,|\,0)=I$

$R_v(n)=\delta_v^2$, $G=[1 \quad 0 \quad \cdots \quad 0]$

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

[iteration]

$$P(n/n-1)=F\times P(n-1/n-1)\times F^T+G\times R_s(n)\times G^T \quad (9)$$

$$K(n)=P(n/n-1)\times G^T/(G\times P(n/n-1)\times G^T+R_v(n)) \quad (10)$$

$$X(n/n-1)=F\times X(n-1/n-1) \quad (11)$$

$$X(n/n)=X(n/n-1)+K\times(y(n)-G\times X(n/n-1)) \quad (12)$$

$$P(n/n)=(I-K(n)\times G)\times P(n/n-1) \quad (13)$$

This algorithm abrogates the computation of the LPC coefficient. The number of calls for the filtering equations is equal to the number of sampling point's n of the speech signals, so the algorithm's time complexity is O (Ln).

B. Improved Filtering Method:

By the recursive equations of the conventional Kalman filtering algorithm, we can find that (5) - (13) contain a large number of matrix operations. Especially, the inverse matrix operations lead to an increase in the algorithm's complexity. If we can reduce the dimension of matrix or eliminate matrix operations, we can greatly reduce the complexity of the algorithm. In Table I, we can find that (11) constantly pans down the value of state vector X(n), and then (12) constantly updates the first value s(n) of X(n). However, during the whole filtering process, only the value of s(n) is useful. So we can use the calculation of s(n) instead of the calculation of the vector in order to avoid the matrix inversion[4]. Furthermore, computational complexity of the algorithm can be reduced to O(n). The recursive equations of the improved filtering method are shown in Table II.

TABLE II
AN IMPROVED FILTERING METHOD PROCEDURE

[Initialization]	
$s(0)=0$, $R_v=\delta_v^2$,	
$R_s(n)=E(y(n)\times y(n))-\delta_v^2$	
[iteration]	
$K(n)=R_s(n)/(R_s(n)+R_v)$	(14)
$s(n)=K(n)\times y(n)$	(15)

C. The Adaptive Filtering Algorithm:

Due to the noise changes with the surrounding environment, it is necessary to constantly update the estimation of noise. So we can get a more accurate expression of noise. Here we further improve the Kalman filter algorithm, so that it can adapt to any changes in environmental noise and become a fast adaptive Kalman filtering algorithm for speech enhancement. The key of the fast adaptive algorithm is that it can constantly update the estimation of background noise. We can set a reasonable threshold to determine whether the current speech frame is noise or not. It consists of two steps: one is updating the variance of environmental noise, $R_v(n)$ the other is updating the threshold U[5].

*1)*Updating the variance of environmental noise by

$$R_v(n)=(1-d)\times R_v(n)+d\times R_u(n) \quad (16)$$

In (16), d is the loss factor that can limit the length of the filtering memory, and enhance the role of new observations under the current estimates. According to [4] its formula is

$$d=\frac{1-b}{1-b^{t+1}} \quad (17)$$

(B is a constant between 0.95 and 0.99. In this paper, it is 0.99) Before the implementation of (16), we will use the variance of the current speech frame $R_u(n)$ to compare with the threshold U which has been updated in the previous iteration. 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 re-estimate the noise variance. In this paper, $R_u(n)$ can't replace directly $R_u(n)$, because we do not know the exact variance of background noise. In order to reduce the error, we use d.

2) Updating the threshold by

$$U=(1-d)\times U+d\times R_u(n)$$
(18)

In (15), d is used again to reduce the error. However, there will be a large error when the noise is large, because the updating threshold U is not restricted by the limitation $R_u(n)$<U. It is only affected by $R_u(n)$ So, we must add another limitation before implementation (18). In order to rule out the speech frames which their SNR (Signal-to-noise rate) is high enough, it is defined that δ_r^2 the variance of pure speech signals δ_x^2 is the variance of the input noise speech signals, and δ_v^2 is the variance of background noise.

We calculate two SNRs and compare between them. According to [6], one for the current speech frames is

$$SNR_1(n)=10\times\log_{10}\left(\frac{\delta_r^2(n)-\delta_v^2(n)}{\delta_v^2(n)}\right)$$
(19)

Another for whole speech signal is

$$SNR_0(n)=10\times\log_{10}\left(\frac{\delta_r^2-\delta_v^2(n)}{\delta_v^2(n)}\right)$$
(20)

In (19) and (20), n is the number of speech frames, and δ_v^2 has been updated in order to achieve a higher accuracy. The speech frame is noise when $SNR_1(n)$is less than or equal to $SNR_0(n)$, or $SNR_0(n)$is less than zero, and then these frames will be follow the second limitation ($R_u(n) \leq U$) [6]. However, if $SNR_1(n)$ is larger than $SNR_0(n)$, the noise estimation will be attenuated to avoid damaging the speech signals.

According to [7], this attenuation can be expressed as

$$R_v(n) = R_v(n)/1.2$$

The implementation process for the whole algorithm can be seen in Table III.

III. MATLAB/ SIMULATION RESULTS

A. The Comparison between the Conventional and the Fast Filtering Method The simulations are done under the following conditions:

1) The pure speech signals were recorded in an anechoic chamber with 16 kHz sampling frequency and digitized.

2) The background noises are an additive white Gaussian noise which is produced by MATLAB.

TABLE III
THE ADAPTIVE METHOD PROCEDURE

[Initialization]
$s(0) = 0$
$R_v(1) = \delta_v^2(1)$
(variance of the first speech frame)
[iteration]
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、 $U = (1-d) \times U + d \times R_u(n)$
Else
3、 $R_v(n) = R_v(n)/1.2$
End
4、 $R_s(n) = E(y(n) \times y(n)) - R_v(n)$
5、 $K(n) = R_s(n)/(R_s(n) + R_v(n))$
6、 $s(n) = K(n) \times y(n)$

function awgn. The noise variance δ_v^2 is assumed to be known, and the SNR of the noise signal SNR in, is defined by

$$SNR_{in} = 10\log_{10}[\{\frac{1}{n+1}\sum_{i=0}^{n} d^2(i)\}/\delta_v^2][dB]$$

(21)

where i is the total number of samples for the speech signal. We adopt two patterns of the noisy speech as the signal samples for the simulations. One is the speech signal corrupted with a background noise, In this section, we will compare the simulation result in the 3 different phases: (a) signal wave, (b) performance, (c) running time for the fast filtering method and the conventional Kalman filtering method.

(a) To compare the filtering efficiency in the view of signal wave, the results are shown in Fig.1 and Fig.2. Each figure contains the wave of pure speech signal, and noise speech signal, filtering the result by the conventional method and filtering the result by the improved method under the condition of SNR 10[dB] in . We can see that the improved method gives quite the same results as the conventional method [7]-[8].

(b) Compare the SNR of them in the view of performance:

$$SNR_{out} = 10\log_{10}[\sum_{i=0}^{n} d^2(i) / \sum_{i=0}^{n} \{d(i) - d(i)\}^2][dB]$$

(21)

means the dimension of the transition matrix;

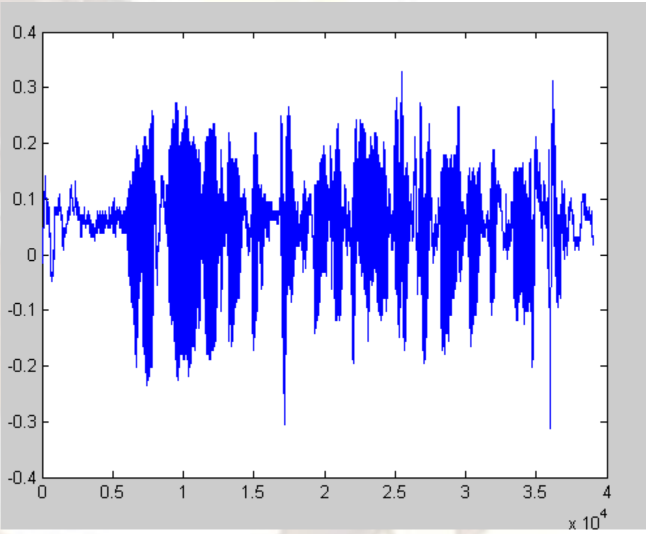

Fig.1 Pure Speech signal

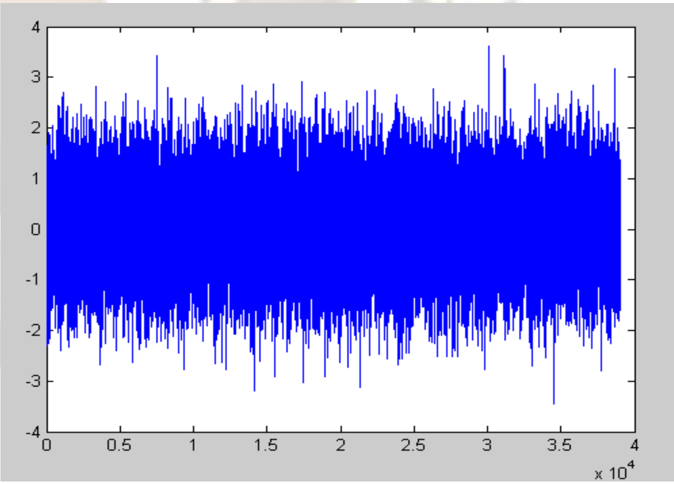

Fig.2 The noise speech

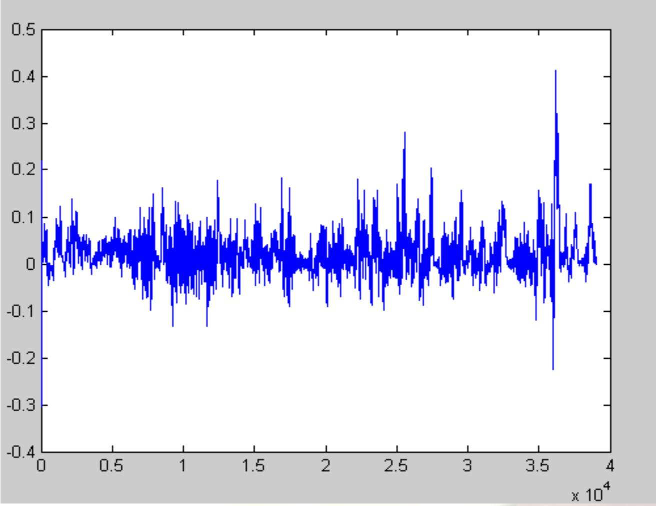

Fig.3 The conventional Kalman filter method

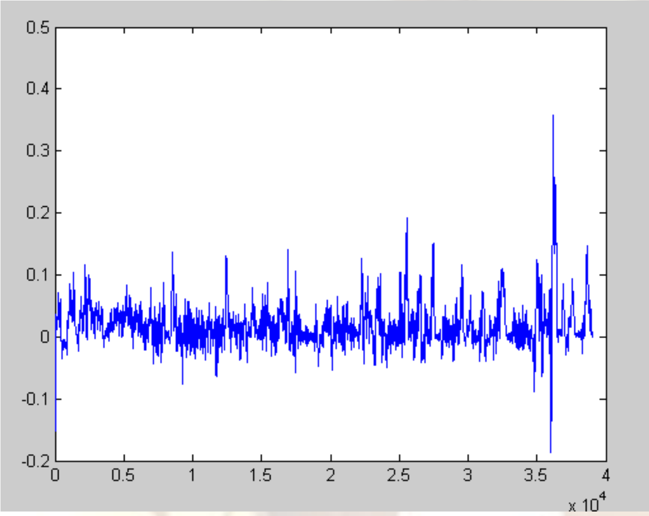

Fig.4 Improved version of speech signal by using proposed method

(c) Compare the running time of the conventional with the improved method. Table I and the improved filtering method in Table II. Both of them are running under this condition:

V CONCLUSION

This paper has presented a Robust Kalman filtering algorithm for speech enhancement by eliminating the matrix operation and designing a coefficient factor. It has been shown by numerical results and subjective evaluation results that the proposed algorithm is fairly effective. Especially, the proposed method contains two-step multiplications in each procedure so that it requires less running time, and the SNR out of this proposed method is higher when the speech signals are degraded by the colored noise. It is concluded that this proposed algorithm is simpler and can realize the good noise suppression despite the reduction of the computational complexity without sacrificing the quality of the speech signal. In the further study, we will improve the adaptive algorithm based on this paper to make it a more accurate assessment of environmental noise. On the other hand, the algorithm will be applied to the embedded-speech-recognition system at the hardware level, so that it can improve the robustness of the system.

References

[1] *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 Aug, 2011 pp. 327 – 332.*

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

[3] *Nari Tanabe, Toshiniro Furukawa , Shigeo Tsujii. Fast noise Suppression Algorithm with Kalman Filter Theory. Second International Symposium on Universal Communication, 2008.411-415.*

[4] *Nari Tanabe, Toshiniro Furukawa ,Hideaki Matsue and Shigeo Tsujii. Kalman Filter for Robust Noise Suppression in White and Colored Noise. IEEE International Symposium on Circuits and Systems, 2008.1172-1175.*

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

[6] *SU Wan-xin, HUANG Chun-mei, LIU Pei-wei, MA Ming-long. Application of adaptive Kalman filter technique in initial alignment of inertial navigation system. Journal of Chinese Inertial Technology,Vol.18, No.1, pp.44-47, Feb.2010.*

[7] *GAO Yu, ZHANG Jian-qiu. Kalman Filter with Wavelet-Based Unknown Measurement Noise Estimation and Its Application for Information Fusion. Acta Electronica Sinica, Vol.35, No.1, pp.108-111, Jan.2007.*

[8] *XIE Hua. Adaptive Speech Enhancement Base on Discrete Cosine Transform in High Noise Environment. Harbin Engineering University, 2006. 332*