Speech Enhancement Using Fast Adaptive Kalman Filtering Algorithm Along With Weighting Filter

Shaik Sami[1], P.Padmaja [2]

[1]Department of Electronics and Communication Engineering,
Sree Vidyanikethan Engineering College, Tirupati – 517 102, A. P., India

[2] Department of Electronics and Communication Engineering,
Sree Vidyanikethan Engineering College, Tirupati – 517 102, A. P., India

Abstract: ***The speech signal which is degraded by noise is improved by the technique called speech enhancement .For this purpose we use many filters such as conventional ,Fast Adaptive, weighted filter,etc. conventional Kalman filter require calculating the parameters of AR (auto-regressive) model, and inverse matrix operation, which is non-adaptive. In this paper we proposed adaptive Kalman filter along with perceptual weighting filter which is used to eliminate the matrix operations, to reduce the calculating time and complexity. Perceptual weighting filter is based on masking characteristics of human auditory system and constantly updating the first value of state vector Z(n) in adaptive filtering which automatically amend the estimation of environmental noise by observation data. The simulation results show that the performance is improved compared to Conventional method.***

Keywords: Conventional Kalman filter, Speech enhancement, adaptive Kalman filter, Perceptual weighting filter.

1. Introduction

Speech enhancement has been a hot research area in recent years with the fast development of multimedia communications and other application. The presence of background noise in speech significantly reduces the intelligibility of speech. Noise reduction or speech enhancement algorithms are used to suppress such background noise and improve the perceptual quality and intelligibility of speech. Removing various types of noise is difficult due to the random nature of the noise and the inherent complexities of the speech. Noise reduction techniques usually have a trade-off between the amount of noise removal and speech distortions introduced due to processing of the speech signal. Several techniques have been proposed for this purpose in the area of speech Enhancement, like spectral subtraction approach, wiener filter, Kalman filter, weighted filter. The performance of these techniques depends on the quality and intelligibility of the processed speech signal. The improvement in the speech signal to noise ratio is the target of most techniques.

Kalman Filter:

The Kalman filter is a recursive prediction filter that is based on the use of state space techniques and recursive algorithms. It estimates the state of a dynamic system. This dynamic system can be disturbed by some noise, mostly assumed as White noise. To improve the estimated state the Kalman filter uses measurements that are related to the state but disturbed as well. Thus the Kalman filter consists of two steps:

1. The prediction
2. The correction

In the first step the state is predicted with the dynamic model. In the second step it is corrected with the observation model, so that the error covariance of the estimator is minimized. In this sense it is an optimal estimator. This procedure is repeated for each time step with the state of the previous time step as initial value. Therefore the Kalman filter is called a recursive filter. Other methods along with conventional methods need to calculate LPC in AR (auto-regressive) model at first and noise reduction later. Simple Kalman filtering algorithm without calculating LPC coefficients in (3) and (4) is shown, but still this algorithm consists of matrix inversion operation and redundant data in large number which is non –adaptive.

Adaptive Kalman filtering algorithm along with perceptual weighting filter was proposed to enhance the speech better than conventional method. First value of the state vector Z(n) was constantly updated in adaptive algorithm and human auditory characteristics is provided by perceptual weighting filter. We need real-time adaptive algorithm to estimate the ambient noise because we don't know what is environmental noise so we add the forgetting factor in (4) and (5) which amends the estimation of environmental noise by observation data automatically, so that it catches the real noise.

2. Kalman Filtering Algorithm

2.1 Conventional Kalman Filtering Method:

There were many types of noises such as White noise, color noise and etc. Here we are considering that speech signal was driven by White noise is all pole linear output from the recursive process. A pure speech signal can be established in q step AR model (predicts an output of a system based on the previous output) given by

$S(n) = \sum_{i=1}^{q} a_i(n) \times s(n-i) + w(n)$ ……….(1)

In (1), $a_i(n)$: LPC coefficient of AR model, w(n): White Gaussian noise which the mean is zero and the variance is δ_u^2 , S(n): Pure speech signal.

Additive observation noise v(n) (i.e. environmental noise) degrade the pure speech signal S(n) whose mean is zero, and

its variance is δ_v^2 .Then noisy speech signal y(n) was formed i.e. given by
y(n)=S(n)+v(n)………...(2)

In this paper, we assumed that the variance δ_v^2 is known, silent segmentation was done to avoid complexity. (1) And (2) can be expressed as the state equation and the observation equation which are given by

[State equation]
m(n)=F(n)m(n-1)+G w(n)…….(3)

[Observation equation]
y(n)=H m(n)+v(n)……(4)
H: Observation vector.

F(n) is the q×q transition matrix expressed as

$$F(n) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \ldots..(5)$$

F (n) =bunch of speech signals.

LPC coefficients were used in conventional Kalman filter for easy estimation and observation of the speech signal. Half the time was spend in computation of above algorithm.

The transition matrix F and the observation matrix H are modified. They has defined as

$$F=H=\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 1 & \cdots & 0 \end{bmatrix} \ldots\ldots\ldots\ldots\ldots.. (6)$$

When we consider one speech signal for enhancement from bunch of signals then it is also defined as q×1 state vector

Z(n)=[S(n)……S(n-q+1) S(n-q+2)], the q×1input vector Q(n)=[S(n) 0………0],and the 1×q observation vector R(n)=[1, v(n)……v(n-q+2)]. Equations (3) and (4) can be rewritten into matrix operations by

[State equation]
Z (n) =F×M (n-1) +Q (n)…………... (7)

[Observation equation]
Y (n) =H×Z (n) +R (n)………… (8)

State equation consisted of the speech signal, and an observation equation consisted of speech signal and additive noise (3). In each iteration of Kalman filter it updates the estimation of the state vector of a system based upon the new observation information. Recursive estimation of Kalman filtering algorithm is showed below and assumed that noise variance δ_v^2 is known and this algorithm abrogates the computation of the LPC coefficient.

The Conventional Method Procedure

[Initialization]
Z(0/0)=0, P(0/0)=I(identity matrix)

$B_v(n)= \delta_v^2$ G=[1 0………0],

$$B_s(n)[i,j]=\begin{cases} E(Y(n) \times Y(n)) - \delta v^2 (i,j = 1) \\ 0 \quad \text{otherwise} \end{cases}$$

[iteration]
$P(n/n-1)=F\times P(n-1/n-1)\times F^T+G\times B_s(n)\times G^T$…(9)
$K(n)=P(n/n-1)\times G^T / G\times P(n/n-1)\times G^T+B_v(n)$.(10)
Z(n/n-1)=F× Z(n-1/n-1) …………..(11)
Z(n/n)=Z(n/n-1)+K(y(n)-G×Z(n/n-1))……(12)
P(n/n)=(I-K(n)×G) ×P(n/n-1)…………….(13)
S(n)=K(n) ×y(n) ……….(14)

2.2 Adaptive Kalman Filtering Algorithm

2.2.1 Classical Adaptive Filters:
For non-zero errors the update algorithm changes the filter parameters such that system output equals with desired output in adaptive filters. The general adaptive –filter configuration is shown in Fig.1.

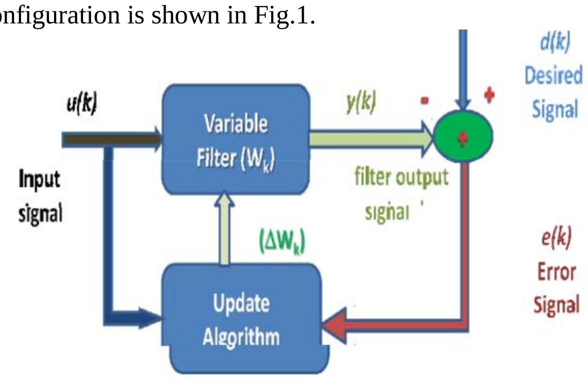


Figure 1: General adaptive Filter Configuration Algorithm

In surrounding environment, noise changes all the time so it is necessary to update noise constantly. Changes in the environmental noise can be adapted by an adaptive Kalman filtering algorithm by constantly updating the background noise.

Process noise and measurement noise can be estimated in fast Adaptive Kalman filtering algorithm on-line according to the measured value and filtered value, tracking changes of noise in real time to amend the filter parameters, and improve the filtering effect. We can set a reasonable threshold in adaptive method which is used to determine whether current speech frame is noise or not. It mainly consists of two steps:

a: Updating the variance of the environmental noise $B_v(n)$,
b: Updating the threshold U.

2.2.2 Updating the variance of the environmental noise by:

$B_v(n)=(1-d)\times B_v(n)+d\times B_u(n)$…………(15)
d is the loss factor which limits the length of the filtering memory and under the current estimation it enhances the role of new observation. Making new data play a major role in the estimation, and leaving the old data forgotten gradually. According to the (4) its formula is
$d=1-b/1-b^{t+1}$……………. (16)
Where b is the forgetting factor (0<b<1), usually ranged from 0.95 to 0.99 here in this paper the value of b is considered 0.99.

Current speech frame $B_u(n)$ is compared with threshold U before implementation of equ (15) . If $B_u(n) \leq U$ then current speech frame can be considered as noise, and the algorithm will re-estimate the noise variance. We don't know background noise variance so $B_u(n)$ can't replace $B_v(n)$ directly. We use d to reduce the error.

2.2.3 Updating the threshold by:

$U = (1-d) \times U + d \times B_u(n)$…………(17)
When errors are large noise will be large, because the updating threshold U is not restricted by the limitation $B_u(n) \leq U$ and it is only affected by $B_u(n)$ Before implementation of equation (17) we add another limitation to increase the SNR(signal-to-noise rate)of the speech frames , it is defined that

δ^2_r: variance of the pure speech signals,
δ^2_x: variance of the input noise speech signals,
δ^2_v: variance of background noise.

We mainly calculate two SNRs and compare between them. According to (6),

1. $SNR_1(n)$: SNR for the current speech frame
   $SNR_1(n) = 10 \times \log_{10}((\delta^2_r(n) - \delta^2_v(n))/\delta^2_v(n))$…(18)
2. $SNR_0(n)$: SNR for the whole speech frame
   $SNR_0(n) = 10 \times \log_{10}((\delta^2_r - \delta^2_v(n))/\delta^2_v(n))$….(19)

In [18] and [19], n is the number speech frames, higher accuracy can be achieved by updating δ^2_v .If $SNR_1(n) \leq SNR_0(n)$, or $SNR_0(n)<0$, then speech frame is noise and then these frames will follow the second limitation ($B_u(n) \leq U$) . 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 $B_v(n) = B_v(n)/1.2$……………….(20)

2.2.4 Perceptual Weighting Filter Algorithm:

Perceptual Weighting filter procedure often results in improvement in the speech performance and it is based on linear prediction (LP) co-efficients that represent the short-term correlation in the speech signal. Weighting filters are widely used in the measurement of electrical noise on telephone circuits, and in the assessment of noise as perceived through the acoustic response of different types of instruments.

$W(Z) = \frac{A(Z)}{A(\frac{Z}{\gamma})} = (1-\sum_{i=1}^{p} a_i Z^{-i}) / (1-\sum_{i=1}^{p} a_i \gamma^{i} Z^{-i})$……(21)

Where A (Z) represents the p^{th}-order LP analysis filter and a_i is LP coefficient. To compute the filter coefficient for this filter, linear predictive analysis is used in (8). Also , γ is a perceptually weighting factor which does not alter the center formant frequency ,but just broadens the bandwidth of ,the formants. Specifically , frequency broadening δ_f given by
$\delta_f = (f_s/\pi)\ln\gamma$…………………..(22)
f_s : sampling frequency in hertz.

For that reason, the weighting filter deemphasizes the format structure while emphasizing the format valleys of the speech signal. This results in a larger matching error in the region of the formants, where spectral masking makes the auditory systems less sensitive to quantization Error. The most suitable value of γ is subjectively selected by listening tests, and for 8KHZ sampling, γ is adopted as 0.9 here.

3. Simulation Results

The comparison between the conventional, the fast filtering Method and perceptual Weighting filter are done using MATLAB

The Performance Evaluation of Adaptive Method Along with Perceptual Weighting Filter:

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.

Table 1 shows the results of the filtering efficiency SNR_{out} under the condition SNR_{in}=6.50[dB] for the female signal and SNR_{in}=3.71[dB] for the male signal. Below table shows that the SNR_{out} of the adaptive method is higher than the non-adaptive method when the speech signals is degraded by the white noise. it is clear that adaptive method can achieve higher performance noise suppression capability than non-adaptive method when the speech signal is degraded by white noise.

Table 1: SNR_{out} Result for the Noisy Speech Signal with White Noise

SNR_{in}[dB]		*SNR_{out}[dB]*		
		Non-adaptive	*Adaptive*	*Perceptual weighting*
Female	6.50	8.89	13.29	19.40
Male	3.71	4.49	5.39	9.04

Table 2: Different filtering methods comparisons of MSE for male and female speech signal

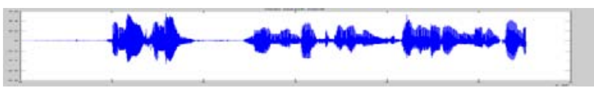
Speech Signal	*Kalman Filter*	*Adaptive Kalman Filter*	*Perceptual Weighting Filter*
Male	0.431	0.045	0.002
Female	0.324	0.032	0.001

Table 3: Different filtering methods comparisons of CPU time for male and female speech signal

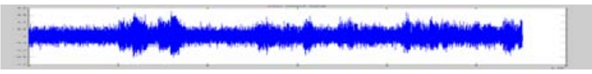
Speech Signal	*Kalman Filter*	*Adaptive Kalman filter*	*Perceptual Weighting Filter*
Male	9.602 sec	5.701 sec	3.562 sec
Female	8.490 sec	3.324 sec	2.826 sec

Tables 1, Tables 2, Tables 3 shows that the proposed method is simpler and can achieve a better filter efficiency despite greatly reducing running time without sacrificing quality of the speech signal.

The Filtering Results for the Male Speech with Noise:

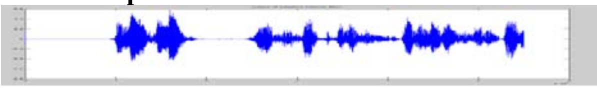


Clean Speech Signal



Noisy Speech Signal

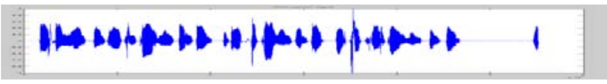
Output of Conventional Kalman Filter



Output of Adaptive Kalman Filter

Output of Weighted Filter

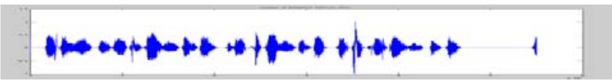
The Filtering Results for the Female Speech With Noise:



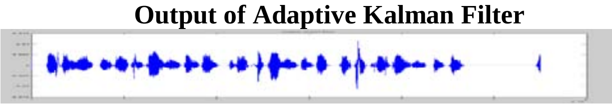
Clean Speech Signal

Noisy Speech Signal

Output of Conventional Kalman Filter



Output of Adaptive Kalman Filter



Output of Weighted Filter

4. Conclusion

This paper has presented "A new technique for speech enhancement using fast adaptive Kalman filtering algorithm along with perceptual weighting filter" which reduces the matrix operation by using coefficient factor, and also provides human auditory characteristics. By numerical results and subjective evaluation results has shown that proposed algorithm was fairly effective.

By using two state multiplications in each procedure of proposed method by that it requires less running time and the SNR_{out} of this proposed method is higher when the speech signals are degraded by the white noise. It is concluded that proposed method was 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.

References

[1] Quanshen Mai, Dongzhi He, Yibin Hou, Zhangqin Huang, A Fast Adaptive Kalman Filtering Algorithm For Speech Enhancement, Aug.2011.

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

[3] 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.

[4] Nari Tanabe,Toshiniro Furukawa, Shigeo Tsujii. Fast noise Suppression Algorithm with Kalman Filter Theory. Second International Symposium on Universal Communication, 2008, 411-415.

[5] WU Chung-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 Environmental. Harbin Engineering University, 2006.

Author Profile

Shaik Sami received the B.Tech degrees in Electronics and Communication Engineering from Narayana Engineering College in 2011, respectively.During 2007-2011; she is currently pursuing Master degree, M.Tech (CMS) in Sree Vidyanikethan Engineering College, Tirupathi.