IOSR Journal of Engineering (IOSRJEN)
e-ISSN: 2250-3021, p-ISSN: 2278-8719, www.iosrjen.org
Volume 2, Issue 11 (November2012), PP 34-38

Speech Enhancement Using Adaptive Kalman Filter Combined With Perceptual Weighting Filter

[1]M.Balasubrahmanyam, [2]V.SunnyDayal
[1]M.Tech,[2]Assistant Professor, Deportment of E.C.EGudlavalleru Engineering College, Gudlavalleru

Abstract:The speech enhancement is one of the important techniques used to improve the quality of a speech signal i.e. degraded by noise. Speech enhancement using conventional kalman filter require calculating the parameters of AR (auto-regressive) model, and performing a lot of matrix operations, which is non-adaptive. In this paper the proposed method i.e. adaptive kalman filter combined with perceptual weighting filter used to eliminate the matrix operations, reduces the calculating time and reduces the complexity. However the perceptual characteristics of the speech signal depends upon the perceptual characteristics of human ear. To further improve the performance of speech enhancement system, adaptive kalman filter combined with perceptual weighting filter, which is based on masking characteristics of human auditory system. Comparison of previous method it is very effective for speech enhancement.

Keywords: conventional Kalman filter, Speech enhancement, adaptive Kalman filter, Perceptual weighting filter, SNR.

I. Introduction

Kalman filteris a mathematical method named after Rudolf E. Kalman 1960, through Peter Swerling actually developed a similar algorithm earlier. Kalman(1960) and Kalman and buky(1961). It was developed as a recursive solution to the discrete-data linear filtering 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.

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 provided to it. The different methods were proposed by [2]-[6]. The many of the methods needs to estimates the parameters of AR model at first, and then perform the noise reduction usingKalman filter algorithm. In this process, the calculation of LPC(linear prediction coding) coefficient and inverse matrix increase the complexity of the filtering algorithm. [3] and [4]have been given a simple Kalman filtering algorithm without calculating LPC coefficient in the AR model, but the algorithm still contains a large number of redundant data and matrix inverse operations. This algorithm is non-adaptive algorithm.

To overcome the drawback of conventional Kalman filter for speech enhancement, propose a new method i.e adaptive Kalman filter combined with perceptual weighting filter. This algorithm constantly updates the first values of state vector Z(n), which eliminate the matrix operations and reduce the complexity of the algorithm. Identification of exact type of environmental noise is difficult and it effects the application of the Kalman filtering algorithm. So we need a real time adaptive algorithm to estimate the ambient noise. The perceptual weighting filter is used to provide human auditory characteristics. Simulation results show that, compared with the conventional Kalman filtering algorithm, the adaptive algorithm of Kalman filtering is more effective. At the same time, it reduced its running time without sacrificing quality of the speech signal.
And also simulation results show that, compared with the conventional kalman filter and adaptive kalman filtering algorithm, perceptual weighting filtering is more efficient in the case of human auditory systems.

II. Kalman Filtering Algorithm

A. Conventional Kalman Filtering Method

A clean speech signal s(n) can be defined as a p-th (AR) autoregressive process and n- th of the noisy speech signal y(n) is expressed as
$s(n) = \sum_{i=1}^{p} a_i\,(n)s(n-i) + w(n)$-(1)
In (1), a_i is the i-th AR coefficient, w(n) is the white Gaussian noise which the mean is zero and the variance is known.
y(n)=s(n)+v(n) -(2)
In (2) v(n) is the additive observation noise, its mean is zero and variance is known.

In this paper, it is assumed that the varianceis known, but in practice we need to estimate it by the initial 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 is
x(n)=F(n)x(n-1)+Gw(n) -(3)
Observation equation is
y(n)=Hx(n)+v(n) - (4)
F(n) is the p by p transition matrix expressed as

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

Where G is the input vector and H is the observation vector.
By using the LPC coefficient in the conventional Kalman filter is to estimate the observations of the speech signal, this process is easy. This part spends half the time of the total algorithm.
The transition matrix F and the observation matrix H are modified. They has defined as

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

It is also defined as the p×1 state vector Z(n)=[s(n)…..s(n-p+1) s(n-p+2)], the p×1 input vector Q(n)=[s(n) 0 …. 0], and the 1×p observation vector R(n)=[1, v(n)…..v(n-p+2)].
Finally, (3) and (4) can be written into the matrix operations by
State equation is
Z(n)=F×X(n-1)+Q(n) - (7)
Observation equation is
Y(n)=H×Z(n)+R(n) - (8)
State equation consisted of the speech signal, and an observation equation consisted of the speech signal and additive noise [3].
The purpose of each iteration of a Kalman filter is to update the estimation of the state vector of a system (and the covariance of that vector) based upon the information in a new observation.
The recursive estimation of Kalmanfiltering algorithm is shown below

$\mathrm{Z}(0|0) = 0, M(0|0) = I$ -(9)

$T_V(n) = \delta_v^2$ -(10)

$$T_S(n)[i,j] = \begin{cases} E\big(Y(n) * Y(n)\big) - \delta_v^2 \ (\mathrm{I, j} = 1) \\ 0 \qquad \text{otherwise} \end{cases}$$

- (11)

[iteration]
M(n|n-1)=F* M(n-1|n-1)*F^T+G*$T_S(n)$*G^T -(12)
K(n)=M(n|n-1)*$G^T|G * \mathrm{M(n|n-1)} * G^T + T_V(n)$
-(13)
Z(n|n-1)=F*Z(n-1|n-1) -(14)
Z(n|n)=Z(n|n-1)+K*(y(n)-G*Z(n|n-1)) -(15)
M(n|n)=(I-K(n)*G)*M(n|n-1) -(16)
S(n)=K(n)*y(n) -(17)
In the above case the noise variance δ_v^2 is known. This algorithm abrogates the computation of the LPC coefficient.

B. AdaptiveKalmanFilteringAlgorithm

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. An adaptive Kalman filtering algorithm for speech enhancement can adapt to any changesin environmental noise, and also it can constantly update the estimation of background noise.

Everyone known Klaman filtering algorithm is very well. Adaptive kalman filtering algorithm can estimate system process noise and measurement noise 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.

In this adaptive kalman filter, we can set a reasonable threshold, it is used to determine whether the current speech frame is noise or not. It consists of mainly two steps: one is updating the variance of the environmental noise $T_v(n)$, and the second one is updating the threshold U.

1) Updating the variance of the environmental noise by

$T_v(n)=(1-d)\times T_v(n)+d\times T_u(n)$ -(18)

In above equation 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. Making new data play amajor role in the estimation, and leaving the old data forgotten gradually. According to the [7]……..? its formula is

$d=1-b/1-b^{t+1}$ -(19)

b is the forgetting factor($0<b<1$), usually ranged from 0.95 to 0.99. In this paper the value of b is considered 0.99.

before implementation of (18), we will compare between the variance of the current speech frame $T_u(n)$ and threshold U which has been updated in the previous iteration. If $T_u(n)$ 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,$T_u(n)$ can't replace $T_V(n)$ directly. In order to reduce the error, we used.

2) Updating and threshold by

$U=(1-d)*U+ d* T_u(n)$ -(20)

In (17), 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 $T_u(n)<U$. It is only affected by $T_u(n)$. So we must add another limitation before implementation of (20). In order to rule out of speech frames which their SNR (Signal-to-noise rate) is high enough, it is defined that δ_r^2 is 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 frame is

$SNR_1(n)=10\times\log_{10}\left(\frac{\delta_r^2(n)-\delta_v^2(n)}{\delta_v^2(n)}\right)$ -(21)

Another for the whole speech signal is

$SNR_0(n)=10\times\log_{10}\left(\frac{\delta_r^2-\delta_v^2(n)}{\delta_v^2(n)}\right)$ -(22)

In(21) and (22), n is the number of speech frames, and δ_v^2 has been updated I 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 ($T_u(n)\leq U$). However, if $SNR_1(n)$ is larger than $SNR_0(n)$, the noise estimation will be attenuated to avoid damaging the speech signals.

C. Perceptual Weighting Filter Algorithm

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.

Usually, the perceptual weighting procedure often Results in speech coder performance. A commonly used weighting filter is based on the linear prediction coefficients that represent the short-term correlation in the speech signal [8]. A representative perceptual weighting filter$W(z)=\frac{A(z)}{A(\frac{z}{\gamma})}=\frac{1-\sum_{i=1}^{p}a_iZ^{-i}}{1-\sum_{i=1}^{p}a_i\gamma^iZ^{-i}}$ is given by WhereA(z) represents the pth-order LP analysis filters and $a_i$is the LP coefficient. To compute the filter coefficients 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 $\delta_f$given by $\delta f=(fs/\pi)\ln\gamma$.Where fs is the sampling frequency in hertz.

For that reason, the weighting filter deemphasizes the formant structure while emphasizing the formant 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 quantizationerror. The most suitable value of γ is subjectively selected by listening tests, and for 8KHZ sampling, γ is adopted as 0.9here.

III. Matlab Simulation

The comparison between the Conventional, the Fast Filtering Method and perceptual weighting filter

The Performance Evaluation of Adaptive Method:

The conventional and improved filtering method is called non-adaptive method for short and the adaptive algorithm of kalman filtering is called adaptive method for short. We will compare the simulation result in 2 different phases: (a) signal wave, (b) performance for the non adptive method, adaptive and perceptual weighting filter method.

In these simulations, two pure speech signals, which have been added with the colored noise, are selected the male and female speech signals in above section.

(a) Compare the filtering efficiency in the view of signal wave.we can see that the adaptive method can achieve higher performance noise suppression capability than the non-adaptive method while the back ground noise changes all the time.

(b) Below table shows the results of the filtering efficiency SNR_{out}under the condition SNR_{in}=6.71[dB] for the female signal and SNR_{in}=3.40[dB] for the male signal. Belowtable shows that the $SNR_{out}$of the adaptive method is higher than the non-adaptive method when the speech signals is degraded by the colored noise.

TABLE

SnroutRESULT FOR THE NOISY SPEECH SIGNAL WITHCOLORED NOISE

SNR_{in}[dB]		SNR_{out}[dB]		
		Non-adaptive	Adaptive	Perceptual weighting
Female	6.71	9.35	10.57	14.197
Male	3.40	4.90	5.05	9.014

From (a) and (b),it is clear that the adaptive method can achieve higher performance noise suppression capability than the non-adaptive method when the speech signal is degraded by the colored noise.

Different filtering methods comparision of MSE for male and female speech signal

Speech Signal	Kalman Filter	Adaptive Klman Filter	Perceptual Weighting Filter
Male	0.432	0.0451	0.002
Female	0.324	0.0321	0.001

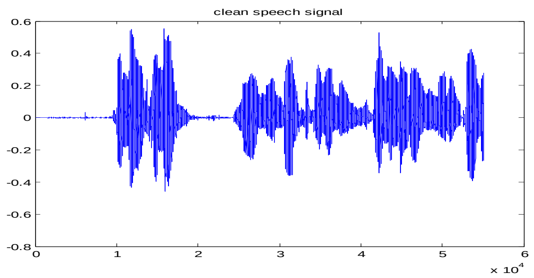
Different filtering methods comparion of CPU time for male and female speech signal

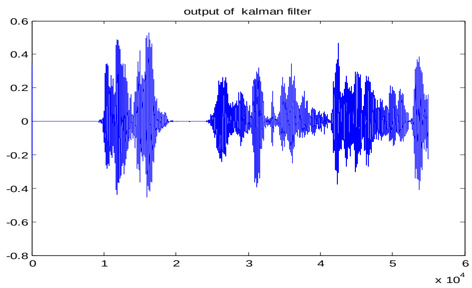
Speech Signal	Kalman Filter	Adaptive Klman Filter	Perceptual Weighting Filter
Male	9.603 sec	5.773 sec	3.801 sec
Female	8.560 sec	3.456 sec	2.956 sec

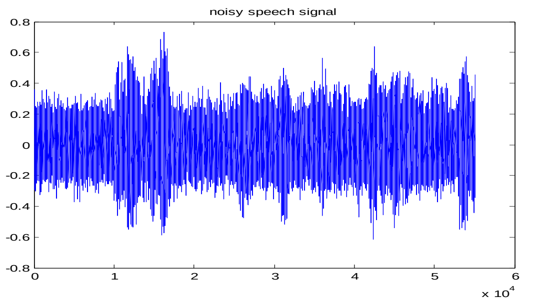
From (a) and (b),it is clear that the adaptive method can achieve higher performance noise suppression capability than the non-adaptive method when the speech signal is degraded by the colored noise.

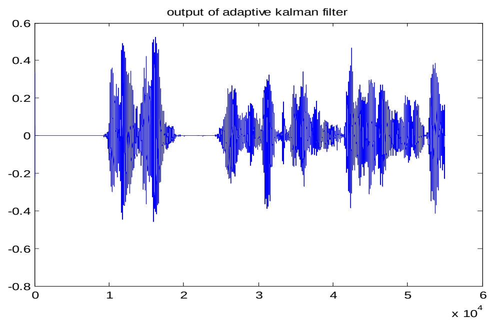
Comparisons of above three tables SNR, MSE, and CPU time. The proposed method is very better.

For all the testing results, the figures and tables shows that the proposed method is simpler and can achieve a better filtering efficiency despite greatly reducing running time without sacrificing quality of the speech signal.









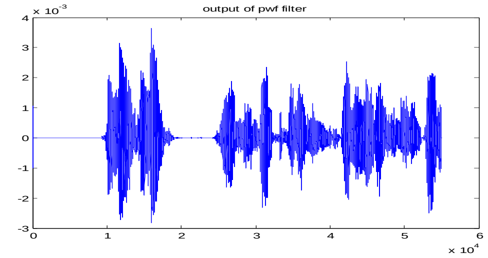
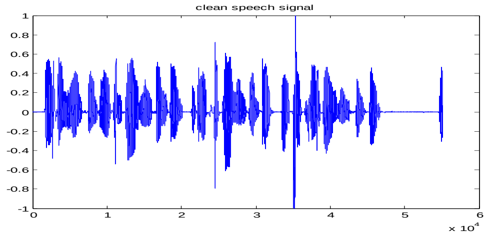
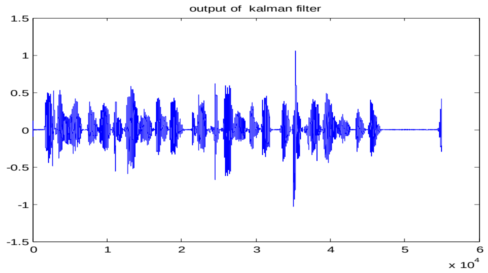


Fig.1. the filtering results for the male speech with noise.





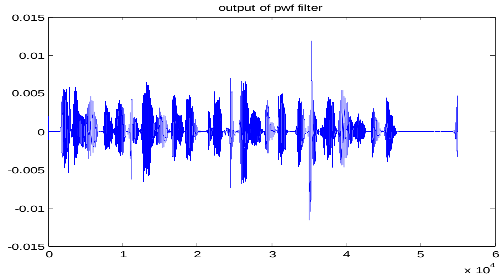
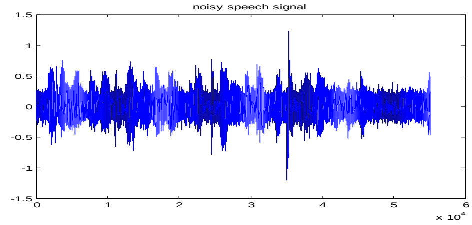
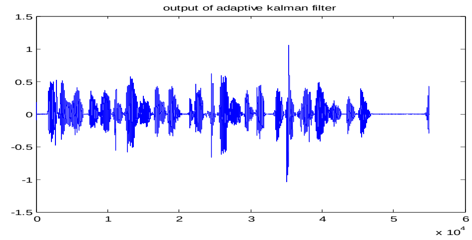


Fig.2. the filtering results for the female speech with noise.





IV. Conclusion

This paper has presented speech enhancement using adaptive kalman filter combined with perceptual weighting filter by eliminating the matrix operation and designing a coefficient factor, and also provide human auditory characteristics. It has been shown by numerical results and subjective evaluation results that the proposed algorithm is fairly effective.

Especially the proposed method of two state multiplications in each procedure so that it requiresLess 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 is 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.

References

[1] Quanshen Mai, Dongzhi He, Yibin Hou, Zhangqin Huang, “Speech Enhancement Using Adapative Kalman Filterining Algorithm”,pp.327-332, August 2011

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

[3] 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.

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

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

[6] 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.

[7] GAO Yu,ZHANG Jain-qui. Kalman filter with Wavelet-Based Unknown Measurement NoiseEstimation 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 Discrere Cosine Transformation in High Noise Environment. Harbin Engineering University,2006.

[9] Joon-Hynk Chang, “Perceptual weighting filter for robust speech modification”, pp.1090-1093,August 2006.