Fast Adaptive Kalman Filter for Online Speech Enhancement Using Signal Subspace Algorithm

[1]*B Pannalal, Assistant Professor, Mail ID:bpannalal@gmail.com*

[2]*G Shankar Rao, Assistant Professor ,Mail ID:shankarraogadidesi@gmail.com*

[3]*P V Sarath Chand, Assistant Professor,HOD ,Mail ID:chandsarath70@gmail.com*

[4]*Dr.K.Venkata Ramana, Associate Professor , Mail ID:ramanaites@gmail.com*

Department of CSE Engineering.
Pallavi Engineering College Hyderabad, Telangana 501505.

Abstract:

Voice enhancement is a method for improving the quality of a speech signal that has been affected by noise. We employ a variety of filters, including standard, Fast Adaptive, and Weighted, to accomplish this goal. Conventional Kalman filters need the calculation of AR (auto-regressive) model parameters, and the inverse matrix operation, which is non-adaptive, in order to be implemented. Adaptive Kalman filter and perceptual weighting filter are introduced in this study as a way to remove matrix operations and therefore decrease computation time and complexity. The initial value of the state vector Z(n) in adaptive filtering is continually updated to reflect changes in the human auditory system's masking features, which is how the perceptual weighting filter gets its start. This strategy outperforms conventional methods, according to the simulation findings.

Keywords:Speech enhancement, adaptive Kalman filter, and perceptual weighting filter are all examples of adaptive Kalman filters

Introduction

Speech augmentation has been a popular study topic in recent years due to the rapid growth of multimedia communications and related applications. Speech becomes substantially less understandable when there is excessive background noise. Noise reduction or speech enhancement algorithms are used to reduce background noise and increase the quality and intelligibility of speech by reducing background noise. The randomness of the noise and the intrinsic complexity of the speech make it difficult to remove different forms of noise from the recording. The level of noise reduction and the amount of speech distortion created by processing the speech stream are frequently trade-offs in noise reduction approaches. In the field of voice enhancement, many strategies have been developed, such as spectral subtraction, wiener filter, Kalman filter, and weighted filter. Quality and intelligibility of the processed speech signal are critical to these approaches' effectiveness. Most strategies are aimed at improving the voice signal to noise ratio. Using the Kalman Filter Using state space approaches and recursive algorithms, the Kalman filter makes recursive predictions. An estimation of the state of a dynamic system. White noise is the most common kind of noise that may affect this dynamic system. Measurements that are connected to the state but also disrupted are used to enhance the Kalman filter's estimation of the state. This is how the Kalman filter works: 1. Predicted outcomes In addition to that, The initial stage is to use the dynamic model to anticipate the current condition. In order to reduce the estimator's error covariance, it is adjusted using the observation model in the second phase. In this respect, it's the best estimate there could be. For each time step, the preceding time step's state is used as the starting point. As a result, a recursive filter is used to describe the Kalman filter. It is necessary to first compute LPC in an AR (auto-regressive) model before reducing the noise. Kalman filtering without generating LPC coefficients in (3) and (4) is illustrated, although this approach still consists of matrix inversion and redundant data that is non-adaptive. It was hypothesised that an adaptive Kalman filtering technique, coupled with a perceptual weighting filter, may improve the quality of voice. Perceptual weighting filter was used to produce the first value of the state vector Z(n) in the adaptive algorithm, which was continually updated. An adaptive method that can be used to estimate ambient noise is needed, and since we don't know what environmental noise is, this approach includes the forgetting factor in (4) and (5) so that it may be updated based on observations to capture the actual noise.

2. Kalman Filtering Algorithm

2.1 Conventional Kalman Filtering Method:

White noise, colour noise, and other sorts of noise were all present. We assume that the White noise that drove the speech signal is a linear output of the recursive process here. It is possible to create a pure voice signal by using the q step AR model (which anticipates a system output based on the previous output).

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

AR model LPC coefficient, white Gaussian noise (w(n)), and pure speech signal (S(n)) are all used in (1). Its standard deviation is v 2. As a result, a noisy speech signal y(n) was generated, as shown by the equation y(n)=S(n)+v(n).... Silent segmentation was used in this work since we believed that the variance v 2 was known. The state equation and the observation equation may be stated in terms of (1) and (2), respectively.

[Observation equation]

$$Y(n)=H\times Z(n)+R(n)\ldots\ldots\ldots\ldots(8)$$

There was a state equation and an observation equation that included the voice signal (3). The estimate of a system's state vector is updated with each iteration of the Kalman filter based on fresh observation data. Calculating the LPC coefficient is no longer necessary since the recursive estimate of the Kalman filtering technique assumes that the noise variance v 2 is known.

The Conventional Method Procedure

$B_v(n)=\delta^2_v \quad G=[1 \;\; 0\ldots\ldots\ldots 0]$,

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

[iteration]

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

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

$Z(n/n-1)=F\times Z(n-1/n-1) \;\ldots\ldots\ldots\ldots(11)$

$Z(n/n)=Z(n/n-1)+K(y(n)-G\times Z(n/n-1))\ldots\ldots(12)$

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

$S(n)=K(n)\times y(n)\ldots\ldots\ldots(14)$

Adaptive Kalman Filtering Algorithm

2.2.1 Classical Adaptive Filters:

System output equals intended output if there is a non-zero error in adaptive filters' filter parameters. Figure 1 depicts a generalised adaptive –filter design.

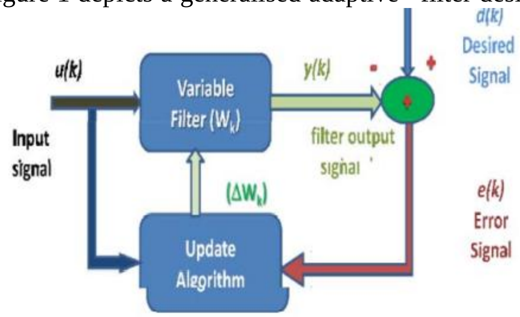


Figure 1: General adaptive Filter Configuration Algorithm

Noise in the surrounding environment fluctuates frequently, necessitating a continuous noise update. An adaptive Kalman filtering technique, which updates the background noise in real time, may adapt to changes in ambient noise. Measurement noise may be tracked in real time and adjusted to enhance the filtering impact using an online adaptive Kalman filtering technique that uses the measured and filtered values. In the adaptive technique, we may establish a suitable threshold for determining if the current speech frame is a noise. There are two key steps: Adjusting the ambient noise Bv(n) variance and the threshold U.

2.2.2 Updating the variance of the environmental noise by:

Bv(n)=(1-d) Bv(n)+d Bu(n)…………(15) According to current estimates, d is a loss factor that limits the duration of the filtering memory and increases the importance of fresh observations. New data should take precedence over old data in the estimate process. D=1-b/1-bt+1................ (16) where the b is the forgetting factor (0b). A comparison of the current speech frame Bu(n) with threshold U is performed before using Equation (15). Current speech frames may be deemed noise if Bu(n) U, and the system will re-estimate noise variance if this is the case. Bu(n) can't replace Bv(n) directly since we don't know the variance of the background noise. To minimise the amount of mistake, we use the d function.

Updating the threshold by:

To put it another way, U is (1-d) U+d Bu(n)………… (17) Because the threshold for updating is high when mistakes are significant, noise will also be high. The restriction has no effect on U.

Bu(n) is the sole factor that affects U. (n) Prior to implementing equation (17), we add an additional limitation to increase the SNR (signal-to-noise rate) of speech frames, it's defined as follows: the SNR (signal-to-noise rate) is defined as SNR (signal-to-noise rate).

We primarily use two SNRs to compare and contrast. First, SNR1(n) is the SNR for this current speech frame, as stated in (6). Assuming that SNR1(n) is ten times the logarithm of the ratio of the signal-to-noise ratio, this means that SNR1(n) = 10 log10 (18) 2. SNR0(n): SNR for the whole speech frame (n). There are two ways to calculate SNR0(n). The first is to use the logarithm of the difference between the two values. (19) A better degree of precision may be reached by revising the value of 2 v in the examples in [18] and [19]. If SNR1(n) SNR0(n), or if SNR0 (n)

2.2.4 Perceptual Weighting Filter Algorithm:

Perceptual The linear prediction (LP) co-efficients that describe the short-term correlation in the speech signal are used in the weighting filter operation. For example, a common use of weighting filters is to quantify electrical noise on telephone lines, and to evaluate the acoustic response of various instruments to noise levels.

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

A (Z) is the pth-order LP analysis filter, and ai is the LP coefficient for ai. Linear predictive analysis is utilised to get the filter coefficient for this filter (8). A perceptually weighted component that does not modify the central formant frequency, but rather widens the bandwidth of the formants, is also present. Fs: sampling rate in hertz; frequency broadening f provided by f=(fs/)ln......................(22) Because of this, the weighting filter de-emphasizes the structure of the format while highlighting the speech's format troughs. As a result, the weighting filter accentuates the speech's format troughs rather than its format structure.

Simulation Results

MATLAB is used to compare the traditional, the fast, and the perceptual Weighting filters.

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

In order to simulate real-world speech, we used two different types of loud speech patterns. One is a background noise-corrupted female voice signal, while the other is a background noise-corrupted male speech signal. In Table 1, we demonstrate the filtering efficiency SNRout for the conditions SNRin=6.50[dB] for the female signal and SNRin=3.71[dB] for the male signal, respectively. For example, when voice signals are degraded by white noise, adaptive methods have greater SNRout than non-adaptive methods. When the voice signal is corrupted by white noise, it is obvious that adaptive methods can provide better noise suppression performance than non-adaptive methods.

Table 1: SNRout 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

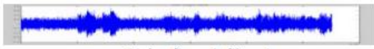
Speech Signal	*Kalman Filter*	*Adaptive Kalman Filter*	*Perceptual Weighting Filter*
Male	0.431	0.045	0.002
Female	0.324	0.032	0.001

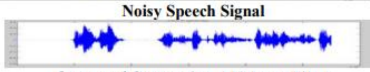
Table 3: Different filtering methods comparisons of CPU time for male and female speech signal

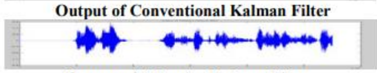
Speech Signal	*Kalman Filter*	*Adaptive Kalman filter*	*Perceptual Weighting Filter*
Male	9.602 sec	5.701 sec	3.562 sec
Female	8.490 sec	3.324 sec	2.826 sec

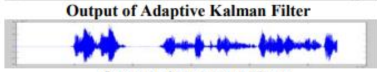
It is shown in Tables 1–3 that the suggested technique is simpler and can obtain a superior filter efficiency while considerably reduced operating time without losing the quality of voice signal quality.

The Filtering Results for the Male Speech with Noise:


Noisy Speech Signal

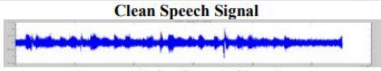

Output of Conventional Kalman Filter

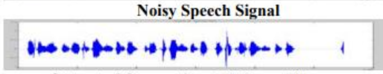

Output of Adaptive Kalman Filter


Output of Weighted Filter

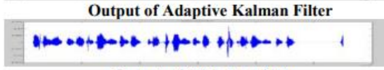
The Filtering Results for the Female Speech With Noise:


Clean Speech Signal


Noisy Speech Signal


Output of Conventional Kalman Filter


Output of Adaptive Kalman Filter


Output of Weighted Filter

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