International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-4, February 2020

Removal of Pink Noise from Corrupted Speech Signal using Kalman Filter

Mohammed Abrar Ahmed, Malaya Kumar Hota

Abstract:** Speech enhancement has been a major challenge in the field of Signal processing. The process of filtering the noise component from the speech signal has achieved many milestones since the early 20th century. Beside many theories Linear prediction coding is one of the best methods for speech, audio signal processing which uses the algorithm of predicting the current estimates based on the past states of an LTI system. Linear prediction is usually used in Speech recognition, Speech enhancement. One of such Kalman filter was introduced and described in 1960 by Rudolf Kalman, which uses the concept of linear quadratic estimation. Kalman filtering is effectively being used in the practical applications like navigation of ships or aircraft, designing motion planning algorithms, in communication area. Kalman filters use the autoregression model of speech for the recursive equations of Kalman filter used in state space model of filter for state estimation. In this paper, we have used Kalman filter to eliminate the pink noise from the corrupted speech signal. Pink noise is very common in electronic devices and occurs in almost all devices. The Speech corrupted with pink noise has been obtained from SpEAR database. We have used MATLAB software for the simulation purpose. Finally, Spectrograms of signals are plotted for a better visual understanding of filtered results.

***Keywords:** Kalman filtering, Linear prediction coding MATLAB, Noise removal, Pink noise, Speech enhancement.*

I. INTRODUCTION

Speech is the major means of communication in the human civilization. Speech enhancement is the process of suppressing the noise from the noisy speech signal. Speech enhancement implies improvement of quality of speech signals by exploitation of several processing techniques. Noise gets added due to numerous reasons. Noise can be categorized as natural noise like sound from thunder, wind, rain or man-made noises are train noise, factory noise, airplane noise, machine noise, babble noise. The foremost influencing noise for a speech signal which interrupts the speech most commonly is additive white gaussian noise which gets added in the channel. A speech enhancer is placed at the receiver's end to filter the noise added in the channel. Corruption of speech signals is unavoidable in many cases and the quality should be improved effectively.

Revised Manuscript Received on February 28, 2020.
***** Correspondence Author
Mohammed Abrar Ahmed*, ECE department, Vellore Institute of Technology, Vellore, India. Email: mohammedabrar.ahmed2017@vitstudent .ac.in
Dr. Malaya Kumar Hota, Professor, Department of communication Engineering, School of Electronics Engineering, Vellore Institute of Technology, Vellore, India. Email: malayakumar.h@vit.ac.in

For example, in sensible case of communication between the captain of the ship to coast guards due to wind, ocean noise if captain receives noisy command signal it would create a major confusion trying to understand what the message is. So, in these cases of practical interest filtering of noise is necessary. Kalman filter is largely MMSE estimator that predicts the unknown states of system [1].

There are quite a few algorithms proposed for enhancement of speech such as Spectral subtraction which subtracts an estimate of the average noise from the noisy signal spectrum [2]. But these algorithms are not suitable for our objective due to distortion created.

Kalman filtering [3] was proposed as a new method for linear filtering it has wide applications and speech enhancement is one of them.

To implement Kalman filtering we need some statistical data i.e. process noise covariance denoted by 'Q' and measurement noise covariance 'R'.

The proper choice of these two parameters enhances the output very much in reducing noise. Using these parameters Kalman filter calculates estimate of current state and monitors Kalman gain for improvement of speech.

The noisy speech signal is differentiated by frames and categorized as a silent frame, noisy frame [8] and algorithm is applied for the same accordingly.

Researchers have proposed various methods of speech enhancement techniques since 1900's. But speech enhancement using Kalman filtering was introduced in 1960 by R. Kalman [3].

Since then Kalman filters were used in speech recognition, Robotics, Power systems field. The Kalman filter works only for Linear systems but for Non-linear systems an Extended Kalman filter is applied. In [4] the researcher applied Kalman filter in speech enhancement for the very first time. The results were compared to Wiener filtering process [5] where Kalman filter extremely stands out in predicting the future estimates of the LTI system with a good accuracy. Since then various methods have been proposed like iterative Kalman filtering [6].

The whole process of Kalman filter in speech enhancement was focused very recently in [1] where filter tuning was achieved very accurate and an self model-order determination system is proposed and using Kalman filter equations based on this order future states are estimated and noisy speech signal is filtered and using Correlation functions segmential SNR are obtained which are compared to various noises like white, train, babble. Many observations have been made in increasing SNR of Kalman filtered output and many researches are going on at present day to make more efficient Kalman filter.

II. METHODOLOGY

A. Linear Prediction Coefficients:

As we discussed earlier speech enhancement uses the autoregression model of speech. It is defined as x(k) i.e. present or current sample or state depends on the past with some linear coefficients and a process noise component that gets added.

It can be mathematically represented as equation(1.1)

$$x(k) = -\sum_{j=1}^{n} a_j x(k-j) + u(k) \quad ...(1.1)$$

Where a_j is linear prediction coefficient and x(k-j) is the past reference state value and u(k) is the process noise and 'n' is the order. Since we know that Kalman filters use autoregressive model that implies that we are assuming the discrete sample x(k) follows autoregression process. As we have to find the AR coefficients, we use Yule walker equations to solve for the coefficients [7] which basically is getting LPC from the Cross- correlation function. We know that correlation is defined as the expectation of present and past values i.e. E[x(k)x(k-1)]. On rewriting equation(1.1) and substituting a_0=1 we get (1.2)

$$\sum_{j=1}^{n} a_j x(k-j) = u(k) \quad ...(1.2)$$

On multiplying both sides of equation(1.2) by x(k-*r*) and taking expectation we get

$$\sum_{j=1}^{n} a_j Rxx(r-j) = -Rxx(r) \quad ...(1.3)$$

Where r ranges till n and Using this n×1 LPC matrix the prediction of coefficients a1, a2,--- a_n are recorded.

B. Kalman filtering concept:

Since the filter is a steady space model a n×1 state vector-matrix X(k) is to be obtained and equation 1.1 is rewritten in steady space model which is explained in [1] as equation 1.4.

$$X(k)= ØX(k-1)+Gu(k) \quad …(1.4)$$

We define Ø as the n × n state transition matrix which has linear prediction coefficients obtained from equation 1.3 and G is input matrix and u(k) is noisy input signal. Now Kalman filter algorithm is applied with initial data as mean value of state vector $\hat{X}(0|0) = m0$ and its corresponding covariance matrix $P_{(0|0)}=P_0$. The input to the filter is taken as the previous state and its corresponding covariance.

The total algorithm of Kalman filter is represented in Fig.1 and there are two main steps those are propagation and update through which filter estimates the future states of signal and again these states act as past states for next future estimate.

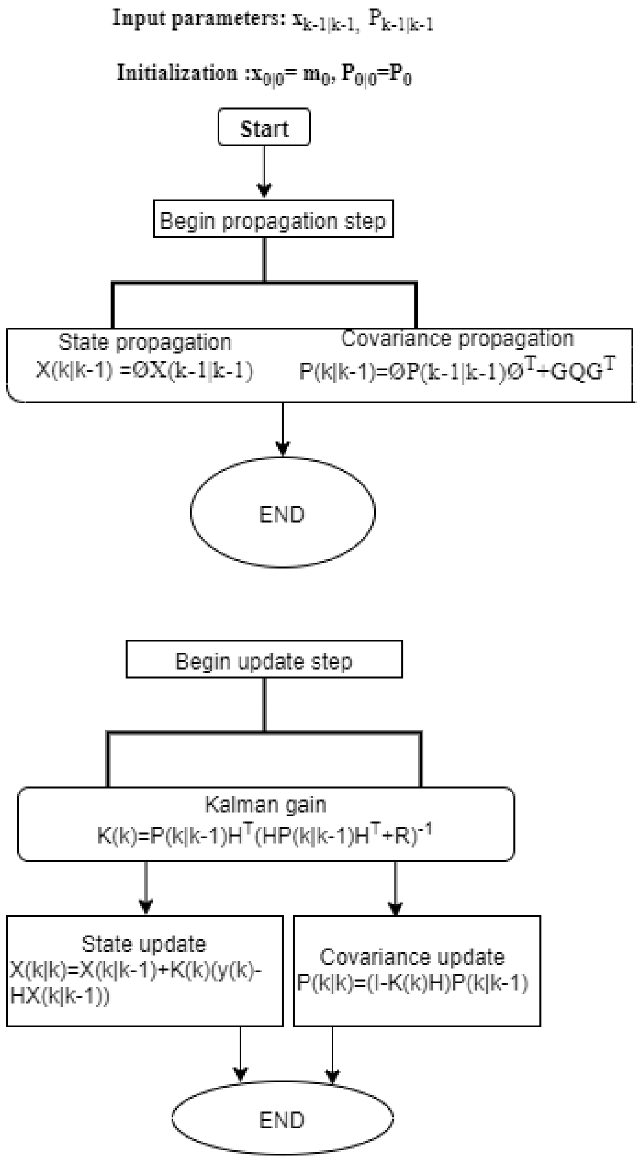

Fig.1 Kalman Filter Algorithm

Where H=[0 0 . . . . 1] and the Kalman gain K(k) is chosen specified that it minimizes the P(k|k) and as P(k|k-1) decreases, K(k) reduces.

C. Estimation of optimum noise covariances:

The most important filter tuning parameters 'R' and 'Q' are to be calculated.

1. Measurement Noise covariance 'R':

To classify voiced and silent frames, the spectral flatness is calculated i.e. the ratio between Geometric mean and Arithmetic mean of the power spectrum ,this methodology is clearly explained in [8]. Flatness is 1 for white noise.

Using the information, R can be determined by following the given process

a) The input speech signal is broken to 80milliseconds each with 10milliseconds overlapping.
b) To every frame, Autocorrelation function and Power spectral density are calculated.
c) PSD range is set between [100 Hz,2000Hz] human frequency range.
d) Flatness is calculated and it is normalized to [0,1].

e) Threshold (th) = (1/sqrt(2)) is chosen, any frame with Spectral flatness below 'th' can be classified as voice, Above 'th' is regarded as Silent frame.

f) Measurement noise variance is determined as the biggest of the variances of all silent frames.

2. Process Noise Covariance 'Q':

To achieve this, we use two parameters i.e. Sensitivity and robustness metrics, J1 and J2 respectively. The two performance metrics and controlling parameter n_q are explained and defined elaborately in [9]. Suppose the process noise variance, σ_u^2 for a frame is to be denoted as Q_f. For each frame of speech, a nominal value of $Q_f=Q_{f\text{-nom}}$ is taken an initial calculation and again varied as $Q_{f\text{-nom}} \times 10^n$. Therefore $n_q = n\times\log_{10} Q_f$. For every value of n corresponding Q_f, J_1, J_2 are determined. Q_c is the value of Q_f at the intersection of J_1 and J_2.

D. Kalman gain:

Voiced frames which have noisy signal component should have high Kalman gain and silent frame which have only noisy signal low Kalman gain is to be given. The gain adjustment is done by using the selection of $Q=Q_c$ for voiced frame and <Qc.
For silent frames. This ensures the voice frame has to utilize high Kalman gain and low Kalman gain for silent frames.

E. Properties of Pink Noise that gets added:

Pink Noise is one of the noise that causes disturbance in the speech signal [10]. It is also called as 1/f noise because of its power spectral density is inversely proportional to frequency. Pink noise is most common occurring signals in biological systems [11]. The contrast between white noise and pink noise is that pink noise has equal amount of noise energy in each octave. Pink noise can also be classified as flicker noise which is one of the main type in Additive noise.
The pink noise has its name because of pink appearance of visible light in this power spectrum. This type of noise is very common in electronic devices and occurs in almost all devices. It decreases the efficiency of output desired by the user. Pink noise is described as fractional noise with α=1 where as for α=2 it is called as brown noise[12]. The 1/f fluctuations are widely found in nature and it is mathematically given by equation 1.15

$$S_x(f) = k/f^{\alpha} \quad \ldots(1.15)$$

Pink noise generators are commercially available and it has the spectrum as shown in the Fig2 [13] where Intensity (dB) and frequency is plotted in Y and X axes respectively in a bode plot.
Pink noise can be easily differentiated with white noise by human ear and it can be modelled to our requirements which makes it very suitable for audio, signal processing applications. Researchers have found that pink noise occurs in heart-beat rhythms, water tides, impulses of neuron, semiconductors, resistivity in electronics.

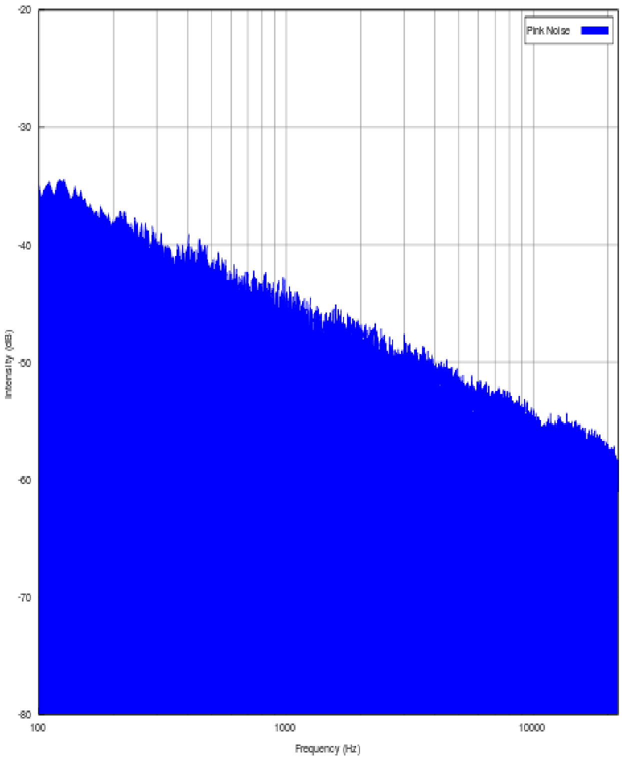

Fig2. Spectrum of a pink noise approximation on a log-log plot. Power density falls off at 10 dB/decade of frequency

III. IMPLEMENTATION OF ALGORITHM

The methodology applied to filter corrupted speech signal is implemented in this section

1. The sample speech signals are downloaded from SpEAR database [14]. The files to be downloaded are clean speech signal and speech signal corrupted with pink noise of 5dB SNR as .wav files.
2. The noisy speech signal is divided into 80milliseconds frames with an overlap of 10 milliseconds.
3. The frames are differentiated into Voice and Silent using flatness parameter. 'R' is obtained as the biggest value of variances of all silent frames.
4. Filter order is fixed to 13 -35 range. We have taken as n=15
5. For each frame of noisy signal, n^{th} order LPC coefficients are calculated as proposed in equation 1.3
6. From these LPC coefficients a (n × n) State transition matrix is designed as given in equation 1.4
7. For the process noise covariance, $Q_f = Q_c$ is taken as the point of intersection of J1 and J2 curves.
8. Kalman filter algorithm in Fig.1 is applied for each frame and if the order of last and current frame is different, then P(k|k) are adjusted.
9. Perform Iterative Kalman filtering, without any filter tuning and LPCs are calculated from the latter state estimates X(k|k).
10. Sum up all the posteriori state estimates obtained after filtering to get the final output.

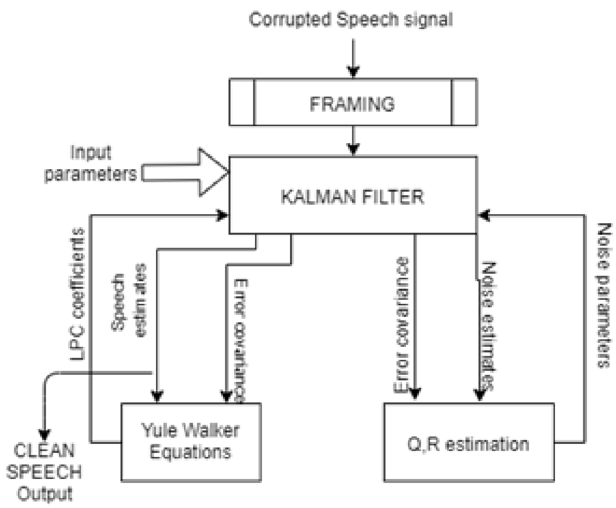

Fig.3 Flowchart of the process of Kalman filter

The complete flow chart process of kalman filter is represented in the Fig.3

IV. RESULT AND DISCUSSION

In this work we have considered speech signal from SpEAR database. Then the speech signal is corrupted with Pink noise of SNR 5dB. After that we have used Kalman filter to remove the noise from the corrupted signal. Fig 4 shows the plot of original speech, corrupted speech and cleaned speech signal. It can be observed that pink noise was effectively removed. The filtered speech can be heard using "wavplay" function in MATLAB.

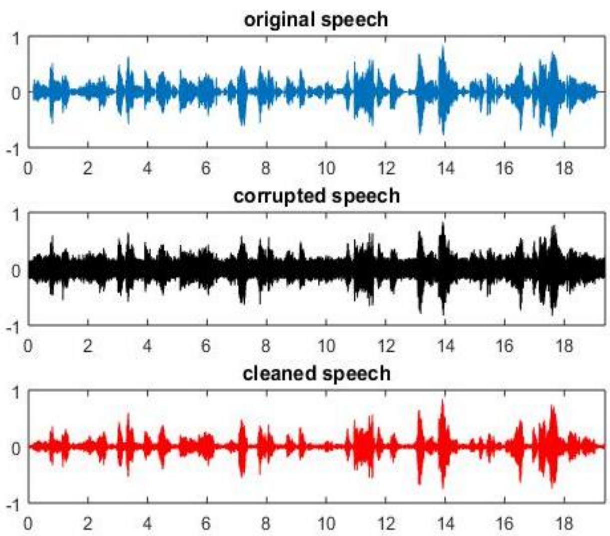

Fig.4 Plots of Speech signal before and after filtering

As we can observe from Fig.4 that pink noise is masking the original/clean speech signal very much. In relative comparison with other types of additive noises pink noise has the capability to corrupt the speech signal in a broad way. In this case Kalman filter provides best results to retrieve the original signal.

Conjointly Spectrogram of three signal are plotted as Fig.5 for a clear estimate of what extent noise is filtered. Analyzing the spectrograms, it is observed that Pink noise corrupted the speech signal heavily as seen in spectrum of noisy signal in Fig.4 and after filtering is done by Kalman filter the Pink noise was cleaned to great extent which gives us more reliable output.

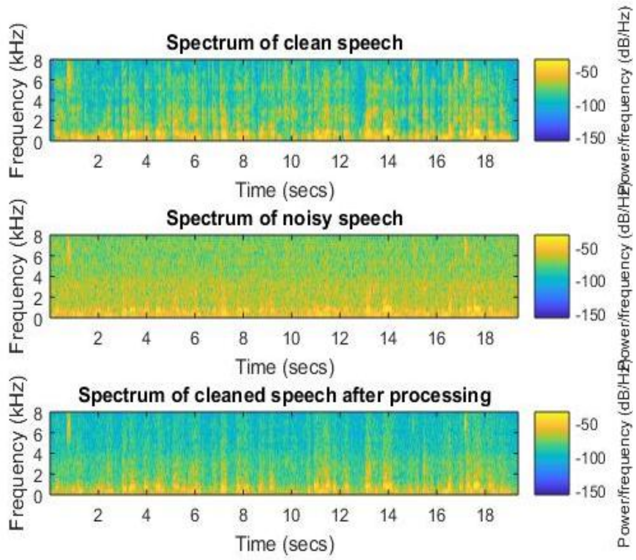

Fig.5 Spectrograms of clean, noisy and filtered speech

V. CONCLUSION

Speech enhancement employing a Kalman filter technique is performed in this paper when the speech signal is corrupted due to additive pink noise which is generated in electronics device. First we have considered a clean speech signal and pink noise was added to it. By employing Kalman filter we found that pink noise was removed to greater extend. The simulation results show that the Kalman filtering technique produced an wonderful estimate of future clean samples in an iterative way.

With the increasing iterations that is the passage of time, Kalman filter gives us more and more precise estimate of speech signal from the input signal by filtering out Pink noise. Conjointly with the increasing iterations, it is observed from the spectrogram that the noise decreases continually. Kalman filter and autocorrelation model were used and Power Spectral Density is used for distinguishing between voiced and noisy signal using which a clear results are obtained. The future work can be analyzing the function in physical world and real time situations where SNR are very high.

REFERENCES

1. O. Das, B. Goswami, R. Ghosh, "Application of the tuned Kalman filter in speech enhancement", *IEEE First International Conference on Control Measurement and Instrumentation (CMI)*, pp. 62-66, Feb 2016.
2. Saeed V. Vaseghi, "Spectral Subtraction", Advanced Signal Processing and Digital Noise Reduction, John Wiley & Sons Ltd. and B.G. Teubner, pp.214-260, 1996
3. Kalman, R. E. (1960). "A New Approach to Linear Filtering and Prediction Problems". *Journal of Basic Engineering*. 82: 35. doi:10.1115/1.3662552
4. B.S Atal. Speech analysis and synthesis by linear prediction of the speech wave. Journal of the Acoustical Society of America, 47(1):65, 1970.
5. S. So and K.K. Paliwal. Suppressing the influence of additive noise on the kalman gain for low residual noise speech enhancement. Speech Communication, 53:355–378, 2011.

Retrieval Number: C8578019320/2020©BEIESP
DOI: 10.35940/ijitee.C8578.029420
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

6. J.D. Gibson, B. Koo, and S.D. Gray. Filtering of colored noise for speech enhancement and coding. IEEE Trans. Signal Process., 39(8):1732–1742, 199
7. Gidon Eshel. The yule walker equations for the ar coefficients. 2003.
8. https://ccrma.stanford.edu/~orchi/Documents/thesis_KF.pdf
9. M.Saha, R. Ghosh, B. Goswami, "Robustness and sensitivity metrics for tuning the extended Kalman filter", IEEE Trans. Instrum. Meas, vol. 63, no. 4, pp. 964-971, 2014.
10. D. Isar, P. Gajitzki, "Pink noise generation using wavelets", Proceedings of ISETc 2016 Timisoara, 2016.
11. Szendro, P (2001). "Pink-Noise Behaviour of Biosystems". *European Biophysics Journal*. **30** (3): 227–231. doi:10.1007/s002490100143
12. Mandelbrot, Benoit B.; Wallis, James R. (1969). "Computer Experiments with Fractional Gaussian Noises: Part 3, Mathematical Appendix". *Water Resources Research*. **5** (1): 260–267.
13. Warrakkk [CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0)]
14. https://ecs.utdallas.edu/loizou/speech/noizeus/ Speech Enhancement Assessment Resource (SpEAR) Database. http://ee.ogi.edu/NSEL/. Beta Release v1.0. CSLU, Oregon Graduate Institute of Science and Technology. E. Wan, A. Nelson, and Rick Peterson

AUTHORS PROFILE

Mohammed Abrar Ahmed is currently 3rd year student of Electronics and communication engineering at Vellore Institute of Technology (VIT), Vellore, Tamil Nadu, India. His main research interest includes Digital communication, Information Theory and encoding, Digital signal processing, Internet of things, Computer Science.

Dr. Malaya Kumar Hota is the Professor in the department of Communication Engineering, School of Electronics Engineering at Vellore Institute of Technology (VIT), Vellore, Tamilnadu, India. He received his M.Tech. in Electronics Engineering from Visvesvaraya National Institute of technology, Nagpur, India, in 2002 and Ph.D. in Electronics and Communication Engineering from Motilal Nehru National Institute of Technology, Allahabad, India, in 2011. He has authored or co-authored about twenty-two publications. His biography has been included in Marquis Who's Who in Science and Engineering, and also in Marquis Who's Who in the World. His main research interest is in digital signal processing, genomic signal processing, biomedical signal processing and speech processing.