ABUAD Journal of Engineering and Applied Sciences (AJEAS)

ISSN: 2992-5584

Volume 2, Issue 2, 91-96

Experimental Assessment of Spectral Subtraction and Kalman Filtering Algorithm on Electricity Generator Noise Reduction in Wireless Communication System

Esther Toyin OLAWOLE[1], Damilare Oluwole AKANDE[2], Zacchaeus Kayode ADEYEMO[2], Ruth Adebimpe AJAYI [1], Folashade Olamide ARIBA[3], Olalekan Femi ADEBAYO[1]

[1]Computer Engineering Department, University of Ilorin, Ilorin, Kwara State, Nigeria.
[2]Electronic and Electrical Engineering Department, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria.
[3]Department of Electrical and Information Engineering, Landmark University, Omu-Aran, Kwara State, Nigeria.

olawole.et@unilorin.edu.ng, doakande@lautech.edu.ng, zkadeyemo@lautech.edu.ng, ajayi.ra@unilorin.edu.ng, ariba.folashade@lmu.edu.ng, adebayo.of@unilorin.edu.ng

*Correspondence: olawole.et@unilorin.edu.ng; Tel.: +2347032060554

Date Submitted: 27/11/2024
Date Accepted: 15/12/2024
Date Published: 31/12/2024

***Abstract**: Noise corrupts, slows down, and reduces the clarity or accuracy of communication. It prevents an undistorted signal or message from being transmitted over wireless communication systems. To reduce noise in wireless communication channels, methods like spectral subtraction and Kalman filter can be used. Spectral subtraction uses its subtractive ability to remove noise in a noisy speech signal, while Kalman filter provides estimates of some unknown noise variables given the measurements observed over time. Therefore, this paper proposes an experimental assessment of both Spectral subtraction (SS) and Kalman Filtering (KF) algorithm for an electricity generator noise-corrupted speech signal over a wireless communication system at the receiver's end. The proposed assessment was carried out using noisy speech signals obtained from a conventional mobile phone and the evaluation was carried out using generator noisy speech signals recorded in a workshop where a generator is been used while, the experimental assessment was performed using MATLAB software. The noise attenuation techniques evaluation was analysed using Mean Square Error (MSE), Signal-to-Noise Ratio (SNR), Perceptual Evaluation of Speech Quality (PESQ), and Short-Time Objective Intelligibility (STOI). The analysis revealed that Kalman filter performed better than spectral subtraction in reducing generator noise in wireless communication systems.*

Keywords: Spectral subtraction, Kalman filter, signal-to-noise ratio, perceptual evaluation of speech quality, short-time objective intelligibility.

1. INTRODUCTION

Noise in wireless communication occurs when undesirable signals disturb a transmitting signal over a wireless communication system. This prevents proper reception and signal interference as a result of network noise which reduces the speed at which data is transmitted within the network, this in turn affects the quality of the audio or video produced (Olawole *et al.*, 2024). Noise, especially electricity generator noise decreases coverage, capacity and limits the efficiency of wireless communication systems.

The common types of noise in a wireless communication system are internal and external noise. Internal noise emanates from the communication devices. It is also called intrinsic noise; it groups a lower limit on dimensions and it is available in all electronic measuring systems. Examples of internal noise are thermal noise and shot noise. External noise also known as extrinsic noise is induced from an external source and causes unsatisfactory behaviour in a circuit (Kyon *et al.*, 2013). This source of noise is from another circuit on the same circuit board (often referred to as cross talk), or it may be external to the equipment. The external source may also result from conduction, capacitive coupling, magnetic coupling, or radiation (Vasilescu, 2005).

Noise can be attenuated in wireless communication systems either by using an integrated analogue-to-digital converter to reduce the main frequency interference or by filtering the signal. There exists a lot of noise attenuating filters in wireless communication, filters like the Kalman filter, radio frequency filters and spectral subtraction techniques.

There have been several conventional works on noise attenuation, a couple of which are revised as follows: Chowdhury *et al.* (2018) worked on noise reduction from speech signal using modified spectral subtraction technique. Varieties of environmental sources of noise and distortion on the quality of the speech signal in a communication system was explored, also the effects of these interfering sounds on speech applications and introduced a technique for reducing their influence and enhancing the speech signal. The experimental result showed a remarkable improvement in SNR for the generalized spectral subtraction and it was observed that the result was satisfactory when white noises are added, an average improvement in SNR for the color noise was recorded. The generalized spectral subtraction method was shown to improve the speech quality and to improve SNR. However, this method was not compared with any other noise reduction algorithm.

Megha and Shridhar (2019) worked on speech enhancement using multiband spectral subtraction with cross spectral component reduction, the authors used two algorithms to enhance the quality of speech signal, and the first algorithm was modified multiband spectral subtraction used to reduce additive noise in stationary noise. The second method was implemented to reduce the cross spectral components when the noise signal is correlated to some extent with the noisy speech signal. The two methods were combined to enhance the noisy speech signal. Subjective listening test and spectrogram analysis were used for the assessment of the speech quality. Result showed that the speech quality improved when using the modified spectral subtraction compared with the multiband spectral subtraction. However, this method was only used for stationary noise.

Mohammed and Malaya (2020) worked on removal of pink noise from corrupted speech signal using Kalman Filter (KF). The authors work on removal of pink noise common in electronic devices that obstructs communication exchange. KF was used to eliminate pink noise from the corrupted speech signal, a SpEAR database (SpEAR is a database of in-ear speech signals recorded in varying conditions of the audio-phonation loop). The result obtained using spectrogram showed that the noise decreases continually. However, this approach was not applied to a real time situation where SNR is very high. In view of the drawbacks and strength observed from the reviewed works, this paper proposes to analyse, evaluate and compare the performance of both SS and KF on electricity generator noisy speech signal over wireless communication devices, this paper proposes to compare the performance of SS and KF techniques in enhancing the quality of propagated speech signal.

The contributions of this paper are as follows:

1) The noisy speech signal used in this paper comprises an amalgamation of external noise obtained from an electricity generator (non-stationary) and conveyed over the wireless channel (internal noise);
2) Attenuation of the unrestrained noise in the noisy speech signal, both Spectral subtraction and Kalman filter was developed to remove the external and internal noise; and
3) Finally, an evaluation comparison of the performance of both SS and KF was done using results evaluated using SNR and MSE. Also, a test for the quality and the intelligibility was performed using both SS and KF respectively.

2. THEORETICAL ANALYSIS

Spectral subtraction is a technique which the input samples are converted from time domain to frequency domain. This is achieved when Discrete Fourier Transform (DFT) is applied to windowed frames in other to separate the magnitude and phase of the noisy speech signal (Hamid, 2018). The noise power estimation and the analysis of spectral weighting take place on magnitude of the speech signal (Sadiq *et al.*, 2013). Once the noise spectral estimate is subtracted from the noisy speech spectrum magnitude, it is then converted back to the original time domain enhanced speech signal via Inverse Fourier Transform (IDFT) (Verteleskaya, 2010; Venkateswarlu and Simak, 2011). From Figure 1, each noisy speech signal frame m_i which comprises of both internal and external noisy speech signal is processed per frame.

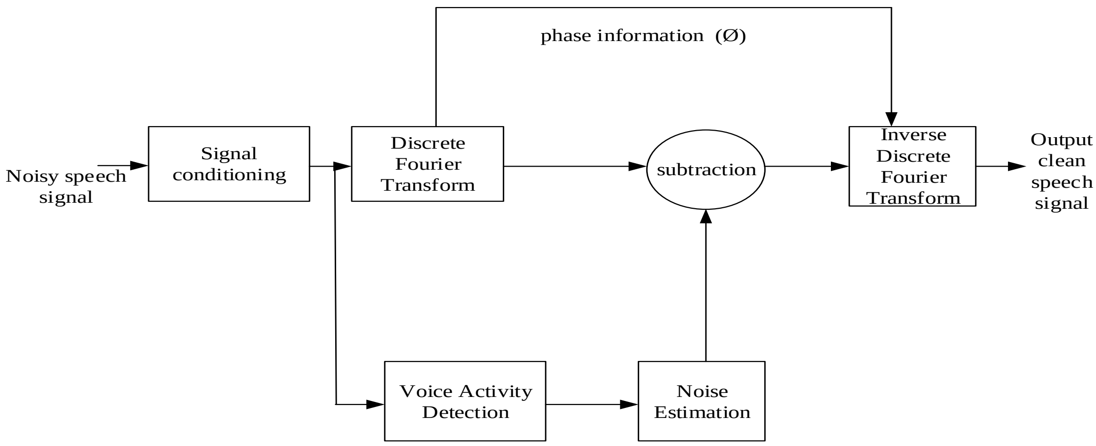

Figure 1: Spectral subtraction

The Discrete Fourier Transform (DFT) block converts noisy speech signals in each frame to its frequency domain to obtain the spectral magnitudes and phase (Upadhyay and Karmakar, 2015). The non-speech segments or frames from

https://doi.org/10.53982/ajeas.2024.0202.13-j

the Voice Activity Detection (VAD) block (binary classifier) are used to determine the noise estimate, the noise estimate or magnitude from the noise estimation block is subtracted from the original stream of noisy speech spectral magnitudes. The difference is then combined with the phase obtained from the DFT block to generate the enhanced speech spectral magnitudes, which are then processed by the Inverse Discrete Fourier Transform (IDFT) to obtain the enhanced speech signal in time domain. The additive noisy speech $m(i)$ which has been corrupted by noise is basically mathematically expressed as shown in Equation (1).

$$m(i) = g(i) + x(i) \tag{1}$$

where: - $g(i)$ = clean speech signal and $x(i)$ = noise signal, all in the discrete-time domain. What spectral subtraction attempts to do is to estimate $g(i)$ from $m(i)$. If the noise process is represented by its power spectrum estimate $|\hat{x}(w)|^2$, that of the noisy speech is $m(w)|^2$, the power spectrum of the clean speech estimate $|\hat{g}(w)|^2$ can be written as shown in Equation (2).

$$|\hat{g}(w)|^2 = |\hat{m}(w)|^2 - |\hat{x}(w)|^2 \tag{2}$$

Since the power spectrum of two uncorrelated signals is additive. The clean speech phase $\theta g(w)$ is estimated directly from the noisy speech signal phase $\theta m(w)$ as shown in Equation (3).

$$\theta Y(w) = \theta G(w) \tag{3}$$

Kalman filter was named after its inventor, Rudolf E. Kalman in 1960 (Kalman, 1960). KF was originally used in aircraft, spacecraft, and other astrological related signal analysis (Kalman, 1960). Over the last two decades, KF based on speech signal enhancement has been a very energetic part of research (Peksa, 2020). KF is an arithmetical process that operates through a prediction and correction method, it combines all the accessible extracted data, coupled with the understanding of the system and the measurement strategy, in order to manufacture an estimation of the desired variables in such a manner that the error is statistically minimized (Peksa, 2020). Paliwal and others in the year 1987 introduced the use of KF for speech enhancement (Paliwal and Basu, 1987; Roy *et al*., 2021).

The block diagram in Figure 2 explains the working principle of Kalman filter when used for speech enhancement; it works in a two-step which are, the prediction step and the observation step as expressed in (Peksa, 2020; Rao and Kumar, 2016). For the prediction step process, the filter shows an estimate of the current state variables, coupled with various uncertainties. Immediately the output of the next measurement is observed, the estimates are updated with the use of weighted average, more weight are given to the estimate with higher certainty (Leonardro *et al.*, 2018). K.F algorithm in speech enhancement is based on state space approach in which a state equation models the dynamics of the signal generation process and an observation equation models the noisy and distorted observation signal (Orchisama *et al.*, 2016; Olawole *et al.*, 2022).

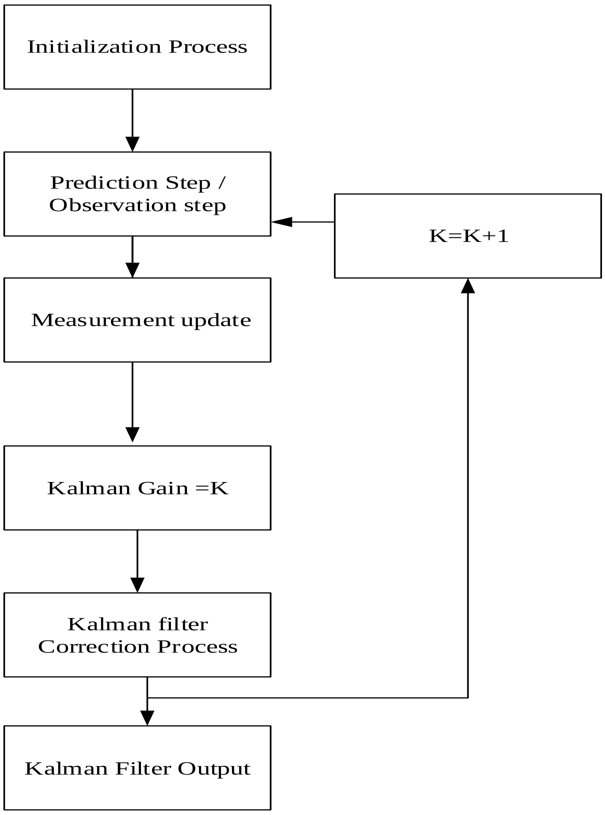

Figure 2: Kalman filter

By initializing an electricity generator noisy speech signal $m(x)$ with noisy observation $n(x)$, an equation describing the state process model and observation model is expressed by (Tamilselvin and Aarthy, 2017) as in Equations (4) and (5).

$$m(x) = Hm(x-1) + Bu(x) + e(x) \tag{4}$$

$$f(x) = Gy(x) + n(x) \tag{5}$$

where: $m(x)$ is a p-dimensional state parameter at time x, $f(x)$ is the M-dimensional observation vector.

H gives a $p \times p$ dimensional state matrix which relates to the state of the process at time x and $x-1$, G is a $p \times p$ control matrix, $u(x)$ is the p-dimensional control input, $e(x)$ is the p-dimensional uncorrelated input excitation (Olawole *et al.*, 2022).

Hence, by applying Kalman filter to signal processing or speech enhancement, exterior control unit $'B'$ is disused and the KF equation is re-expressed by (Tamilselvin and Aarthy, 2017) as shown in Equations (6) and (7).

$$m(x) = Hy(x-1) + e(x) \tag{6}$$

$$f(x) = Gy(x) + n(x) \tag{7}$$

H gives a $p \times p$ dimensional state transition matrix as expressed in Equation (8).

$$H = \begin{bmatrix} 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 1 & \cdot & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 1 \\ b_p & b_{p-1} & b_{p-2} & \cdot & \cdot & \cdot & b_2 & b_1 \end{bmatrix} \tag{8}$$

Therefore, the sum of the diagonal of the $p \times p$ dimensional state transition matrix in equation 9 represent the mean squared error. Hence, the mean square error is reduced by reducing $y(x)$, which will in turn reduce H. by differentiating $n(x)$ with respect to G, Equation 5 will become Equation (9):

$$f(x) = \mathrm{G}^T m(x) + n(x) \tag{9}$$

where: $f(x)$ is the noisy speech signal, $\mathrm{G}^T m(x)$ is the clean speech signal, $n(x)$ is noise.

3. MATERIALS AND METHODS

This research was initiated by obtaining noisy speech sample recorded over a mobile phone. The obtained speech samples were that of spoken phrases by both a female and male speaker in a factory where generator is been used. The obtained noisy speech signals have pulses at the beginning and end of the samples. The period of the pulses creates room for background noise region at the beginning and end of the words. The recorded noisy speech samples are stored in a wave file format which is generally a window-friendly format. The obtained noisy speech data were then subjected to digital data conditioning or pre-processing suitable for each technique. The conditioning involves the sampling of the noisy speech signal, framing, and then windowing. The noisy speech samples where simulated using Spectral Subtraction and Kalman filtering algorithm.

Framing is one of the important stages of pre-processing, and was executed after the noisy speech signal was obtained. This was performed on the signal in order to evaluate the characteristics of the noisy speech signals in bits. This helps to obtain the actual characteristics of the noise in the silent regions of the speech (Loizou, 2013). Framing was performed on the recorded noisy speech signals and the framing process was obtained by taking a frame width or period to be 20 msec and 10 msec shift. The 10 msec frame width gave room for 50% overlap and ensured that all samples were accounted for within at least two blocks. The speech signals are quasi-stationary between 5 msec and 100 msec, hence, the frame width/period chosen were not standards, but were chosen based on the period of the speech length (Oday, 2018). An average speech length of 20 sec was used in this work so as to be able to record all environmental noise present at the data collection location.

Windowing technique is a very important stage of the pre-processing stage. It involves a function known as windowing function. In this paper, hamming window was used, this is because the method was used to resolve sinusoid that are close together in frequency. In order to minimize spectral distortions when framing the speech signal, each frame was multiplied with a hamming window by taking the product of the window and input signal to give the output of windowed signals. The output from the pre-processing stage is now used as the input of Kalman filtering and Spectral subtraction. Also, the result was analysed using MSE, LSD, STOI and PESQ.

4. RESULTS AND DISCUSSION

Basically, Mean Squared Error (MSE) in speech enhancement relates how close a model is to arrive at an audible speech signal. Hence, the lower the MSE, the cleaner the speech signal. MSE in noise reduction is achieved by finding a means to reduce the noise in a noisy speech signal in order to minimize the difference between the noisy speech signal and the enhanced speech signal. Buttressing the fact that, the filtering technique with the least MSE has a better performance. Figure 3 presents a graph comparing the performance of the developed filtering technique and the other techniques understudied using MSE. The MSE values for spectral subtraction and Kalman filter are 0.008643 and 0.003841, while the corresponding values at 4 dB the MSE for spectral subtraction, Kalman filter were 0.000540 and 0.000240 respectively. At SNR of 12 dB, the MSE values of Spectral subtraction and Kalman filter were 6E-05 and 6E-05. Also, at SNR of 16 dB, the MSE value obtained were 3.38E-05 and 1.5E-05. While the corresponding values at SNR of 20 dB were 2.16E-05 and 1.19E-05 for spectral subtraction and Kalman filter respectively. Although, there exist an overall reduction in MSE as the SNR increases. However, Kalman filter experiences a higher reduction in MSE when compared to Spectral subtraction

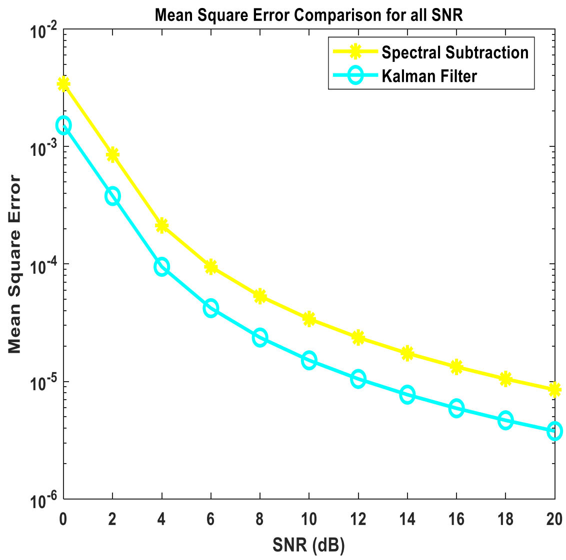

Figure 3: MSE against SNR

The smaller the LSD at each SNR, the better the quality of the enhanced speech signal as observed in the result. The LSD values at 0 dB for spectral subtraction and Kalman filter were 1.245775 and 1.110392 respectively, while at 4 dB the corresponding values were 0.773656 and 0.643606 respectively. Also, at SNR at 8 dB, the LSD values obtained were 0.559652 and 0.453631 for SS and KF, while at SNR of 12 dB, the corresponding values obtained were 0.453631 and 0.363658 respectively. At SNR value of 16 dB the LSD obtained were 0.388141 and 0.309002, while at SNR value of 20 dB, the LSD values for spectral subtraction and Kalman filter were 0.342854 and 0.270419 respectively. From the result presented in Figure 4, it is observed that with an increase in SNR value, the LSD decreases gradually for the two filters.

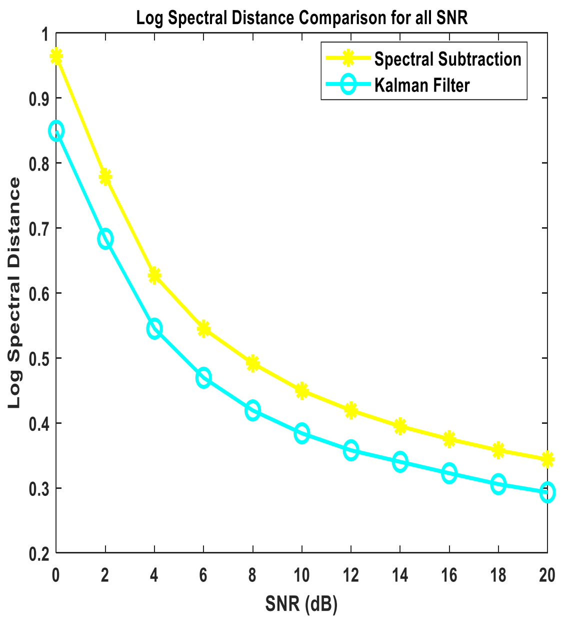

Figure 4: LSD against SNR

Figure 5 represents a bar chart of STOI against SNR. The result shows that the intelligibility of the enhanced speech signals for both spectral subtraction and Kalman filter increases with increase in SNR. However, Kalman filter technique shows better performance in intelligibility over spectral subtraction noise filtering technique. This is a signal that the intellect of the enhanced speech is well conserved by applying the Kalman filter technique. Figure 6 represents a bar chart of PESQ as against SNR. The PESQ of the individual existing techniques with the developed hybridized spectral-Kalman filtering technique at different SNR is presented in Figure 6. The closer the PESQ values to 4.5 at each SNR, the better the quality of the enhanced speech signal as observed in the result. The PESQ values at 0 dB for spectral subtraction and Kalman filter were 2.2 and 2.25 respectively, while at 4 dB the corresponding values were 2.23 and 2.30. Also, at SNR at 8 dB, the PESQ values obtained were 2.42 and 2.50 for SS and KF, while at SNR of 12 dB, the corresponding values obtained were 2.43 and 2.53 respectively. At SNR value of 16 dB the PESQ obtained were 2.50 and 2.52, while at SNR value of 20 dB, the PESQ values for spectral subtraction and Kalman filter were 2.52 and 2.57 respectively. From the result presented in Figure 6, it is observed that with an increase in SNR value, the PESQ values increases gradually for all the filters considered. However, the rate of increment recorded using Kalman filter is more than that of spectral subtraction.

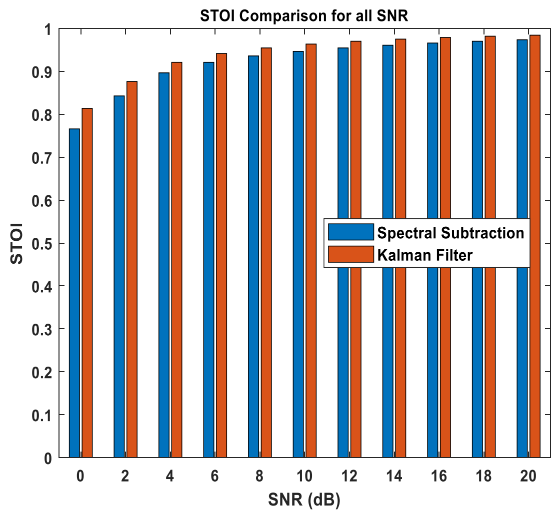

Figure 5: STOI against SNR

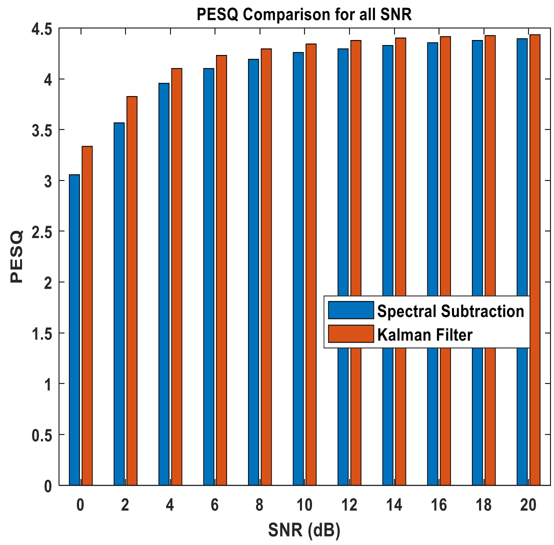

Figure 6: PESQ against SNR

5. CONCLUSION

The MSE was determined at different SNR by forming a state space transition matrix. The MSE is then formed from the trace of the matrix for KF. Also, LSD was determined by finding the ratio of the log of the waveform obtained from the noisy speech signal to that obtained from the enhanced speech signal. While, the PESQ for speech quality and STOI to determine the speech intelligibility were obtained. After the empirical analysis, Kalman filtering technique is judged to be more efficient than spectral subtraction technique. It also realizes a good noise suppression by reducing computational complexity without sacrificing the quality of speech signal.

ACKNOWLEDGEMENT

The completion of this research paper would not have been possible without the support and guidance of Prof Z. K Adeyemo, and Dr. D. O Akande. Their dedication and overwhelming attitude towards helping their students is solely responsible for completing this research paper. The inspiration and perceptive feedback were influential in achieving this task. Special thanks to Dr Mrs Ariba and Engr. O. F Adebayo for data analysis and technical assistance. Their astonishing assistance have impressively enhanced the quality of this paper.

REFERENCES

[1] Chowdhury, S. M., Mahmudul, H., and Khalil, A. (2018). Noise Reduction from Speech Signals using Modified Spectral Subtraction Technique, *European Journal of Engineering Research and Science. EJERS,* **3**(7): 78-80. 11(2): 82-85.

[2] Hamid, O. K. (2018). Frame blocking and windowing in speech signal. Journal of Information, Communication and Intelligence System, 4 (5): 87-94.

[3] Kyon, D. H., Won-lee, L., Myung-sook, K., and Myung, J. B., (2013). Hi Pass Pink Noise: Its Acoustic Features and Standards, *International Journal of Multimedia and Ubiquitous,* 8(6): 229-236.

[4] Kalman, R. E., (1960). A New Approach to Linear Filtering and Prediction Problems, *Journal of Basic Engineering.*

[5] Leonardro, A. Culberto, A. A., Edwardo S. V. and Roberto, D. C. (2018). Comparative study between the discrete frequency Kalman filtering and discrete time Kalman filtering with application in noise reduction in speech signals. Journal of Electrical and Computer Engineering. 1-5.

[6] Loizou, P. C. (2013). Speech enhancement: theory and practice. 2nd edition. CRC Press, Inc.: Boca Raton, FL, USA.

[7] Megha, S. K., and Shridhar, S. K., (2019). Speech Enhancement using Multiband Spectral Subtraction with Cross Spectral Component Reduction, *International Journal of Engineering Research and Technology (IJERT).* **7**(10):1-4.

[8] Mohammed, A. A., and Malaya, K. H., (2020) Removal of Pink Noise from Corrupted Speech Signal using Kalman Filter, *International Journal of Innovative Technology and Exploring Engineering (IJITEE).* **9**(4):2034-2038.

[9] Odey, K. (2018). Frame blocking and windowing in speech signal, *Journal of information, communication and Intelligence system, vol.,* 4(5), pp. 87-94.

[10] Olawole E. T, Akande, D. O. and Adeyemo Z. K. (2022). Speech Enhancement in Wireless Communication System using Hybrid Spectral-Kalman Filter, *International Journal of Electrical and Electronics Telecommunication*, vol.22.

[11] Olawole E. T., Akande, D. O., Adeyemo Z. K., Ojo F. K. and Ojo S. I. (2024). Effect of Modulation Domain Coupled Kalman Spectral Filter on Speech Enhancement over Wireless Voiced Communication System. *Nigerian journal of technological development*, 21(3), PP 20-26.

[12] Orchisama, D., Goswami, B. and Ghosh. R. (2016). Application of the tuned Kalman Filter in speech enhancement. In 2016 IEEE First International Conference on Control, Measurement and Instrumentation (CMI), pages 62-66.

[13] Paliwal K.K., and Anjan B. (1987), “A speech enhancement method based on Kalman filtering”. In ICASSP, number 6.3, pages 1-4.

[14] Peksa, J. (2020). Prediction framework with Kalman filter algorithm. *Institute of Information Technology, Riga Technical University,* vol. 11(358), pp.2-8.

[15] Rao, G. M. and U. S. Kumar (2016). Speech enhancement using iterative Kalman filter with time and frequency mask in different noisy environment. International Journal of Signal Processing, Image Processing and Pattern Recognition. 9(9): 317-324.

[16] Roy, S. K. and K. K. Paliwal (2021). Robustness and sensitivity tuning of the Kalman filter for speech enhancement. *Signals*, (2) 434–455.

[17] Sadiq, M.T. Sabbir, N. and Kulesza, W.J. (2013). Spectral Subtraction for Speech Enhancement in Modulation Domain. *International Journal of Computer Science Issues JCSI.* **10**(4): 56-70.

[18] Tamilselvin, G.M and Aarthy, P. (2017). Online tuning of fuzzy logic controller using Kalman algorithm for critical system. *Department of EIE, Barinari Amman Institute of Technology, Thyamangalam, India, vol.* 15(5): pp. 492-503.

[19] Upadhyay, N. and Karmakar, A. (2015). Speech enhancement using spectral subtraction-type algorithms: a comparison and simulation study. *Eleventh International Multi-Conference on Information Processing*-2015 (IMCIP 2015), 54: 574-584.

[20] Verteleskaya, E. (2010). Spectral subtractive type speech enhancement. In: *Information and Communication Technologies and Services, Vol. 8.*

[21] Venkateswarlu, S. C., Prasad, K. S and Reddy, A. S. (2011). Improve speech enhancement using Weiner filtering. In Global Journal of Computer Science and Technology Volume 11 Issue 7.

[22] Vasilescu, W. (2005). Electronic Noise and Interference Signal. *Signals and Communication Technology.* Springer, 255-313.