SPEECH ENHANCEMENT USING COMBINATION OF DIGITAL AUDIO EFFECTS WITH KALMAN FILTERS

* JULOORI SANTHOSHI ** P LAXMAN, ***R RAMESH BABU

** MTech student, Dept of ECE, Jagruti Institute of Engineering and Technology, Hyderabad, TS, India.*

*** Assistant Professor, Dept of ECE, Jagruti Institute of Engineering and Technology, Hyderabad, TS, India.*

*** Professor & HOD, Dept of ECE, Jagruti Institute of Engineering and Technology, Hyderabad, TS, India.*

Abstract - The term "Quality of Speech" in Speech Enhancement techniques is associated with Clarity and Intelligibility. Till now due to the variable nature and characteristics of noise with time and process to process, Speech Enhancement is a difficult problem in Noisy environment. In this paper, we proposed a method to improve the quality of speech based on combination of Digital Audio Effects with Improved Adaptive Kalman Filter when only corrupted speech is available. In this approach to enhance the Speech content in the Noisy speech signal, Digital audio effects are used. A Digital Expander generates audio effects which operates on a low signal level and create more likely sound characteristics. And further, noise is removed by Auto Regressive modelled improved adaptive Kalman filter. It is concluded that the proposed method with additive colour noise is found to be better than conventional methods in terms of Signal-to-Noise ratio and intelligibility.

Keywords: *Kalman filter, Digital Audio Effect, conventional, Signal-to-Noise ratio .*

I. INTRODUCTION

Speech plays a vital role in our daily communication and also for human machine interfacing. Therefore, production and perception of speech have become an interesting part of the research since decades. But the quality and intelligibility of the speech are significantly degraded by the presence of background noise, which affects the ability in understanding other's speech, causes error in Human Machine Interfacing, etc. In this digital world, it's really hard for any signal in real-time environment to escape from noise. This hits us really hard when it comes to deliver a message from one place to another and there is a need for cleaning up or enhancing the message signal but at the same time, not giving up any intelligibility of the message (content, not just clarity). Since speech messages have been the mode of communication everywhere, need for speech enhancement is required whenever the signal comes in contact with the real-time environment. Modelling of human speech production process helps in enhancing the speech. But, as speech is a highly non stationary signal, it is difficult to model the human speech production process. Though speech is highly non stationary signal, it is stationary for very short period of time. Based on this fact, Classical speech enhancement techniques are considered for speech segment models for short time, but these short time models do not include the effects of the noise as noise has long term characteristics. On the other hand, such long-term characteristics are naturally taken care of in the autoregressive approach as speech signals are not modelled on a short-time basis but as a whole. The AR model is also known to be good for representing unvoiced speech. However, it is not quite appropriate for voiced speech since voiced speech is often quite periodic in nature. This has motivated us to look into speech models which can satisfactorily describe both voiced and unvoiced speech, and allow for exploitation of the long-term characteristics of noise.

II. RELATED STUDY

Speech enhancement is an area of speech processing where the goal is to improve the intelligibility and/or pleasantness of a speech signal. The most common approach in speech enhancement is noise removal, where we, by estimation of noise characteristics, can cancel noise components and retain only the clean speech signal. The basic problem with this approach is that if we remove those parts of the signal that resemble noise, we are also bounded to remove those parts of the speech signal that resemble noise. In other words, speech enhancement procedures, often inadvertently, also corrupt the speech signal when attempting to remove noise. Algorithms must therefore compromise between effectiveness of noise removal and level of distortion in the speech signal. Current speech processing algorithms can roughly be divided into three domains: spectral subtraction, sub-space analysis and filtering algorithms. Spectral subtraction algorithms operate in the spectral domain by removing, from each spectral band, that amount of energy which corresponds to the noise contribution. While spectral subtraction is effective in estimating the spectral magnitude of the speech signal, the phase of the original signal is not retained, which produces a clearly audible distortion known as "ringing". Sub-space analysis operates in the autocorrelation domain, where the speech and noise components can be assumed to be orthogonal, whereby their contributions can be readily separated. Unfortunately, finding the orthogonal components is computationally expensive. Moreover, the orthognality assumption is difficult to motivate. Finally, filtering algorithms are time-domain methods that attempt to either remove the noise component (Wiener filtering) or estimate the noise and speech components by a filtering approach (Kalman filtering). To fulfil the objective of objective of speech enhancement was initially done by using Kalman Filter, but the results did not meet the requirement. So, we segregated the entire signal into small samples called windows by adopting different windowing techniques like rectangular windowing and Hamming windowing. We iterated the process for few times by updating the autoregressive filter coefficients after every repetition. Even though the process takes long time for a tiny speech signal data, the output can be compared with input for its similarity.

III. AN OVERVIEW OF PROPOSED SYSTEM

The Kalman filter for speech enhancement was proposed by K.K Paliwal and A. Basu by using estimation of speech signal parameters from clean speech before it corrupted by white noise and further extended by using the adaptive kalman filter to the random and coloured noises. Above problem can be overcome by using the digital expander which minimises the distortions. The foundation for parameter identification of auto regressive (AR) model for speech signal which is fluctuating with respect to time is linear prediction coefficient estimation (LPC). It is a standout amongst the most intense discourse examination procedures, and a standout amongst the most valuable strategies for encoding great quality discourse at a low piece rate and gives to a great degree precise assessments of discourse parameters.AR coefficients are updated for every time frame of 25msec duration which is chopped by Hanning window and analysed using the linear prediction analysis method (LPC). The use of Kalman Filter for speech enhancement in the form that is presented here was first introduced by Paliwal (1987).This method however is best suitable for reduction of white noise to comply with Kalman assumption. In deriving Kalman equations it normally assumed that the process noise (the additive noise that is observed in the observation vector) is uncorrelated and has a normal distribution [2, 13-15]. This assumption leads to whiteness character of this noise. There are, however, different methods developed to fit the Kalman approach to colored noises. It is assumed that speech signal is stationary during each frame, that is, the AR model of speech remains the same across the segment.

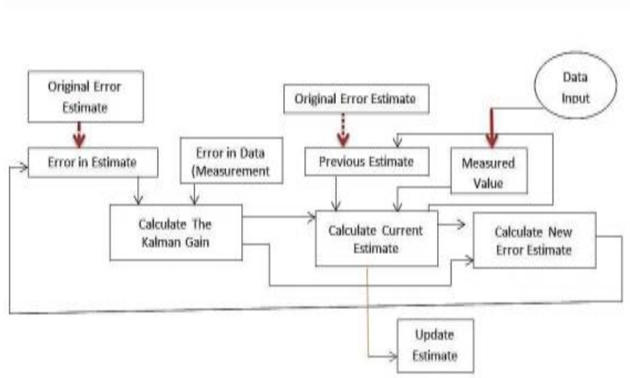

Fig.3.1. Mechanism of Kalman filters in speech enhancement.

The noise removal from noisy speech signal was initially done by using Kalman Filter, but the results did not meet the requirement. So, we segregated the entire signal into small samples called windows by adopting different windowing techniques like rectangular windowing and Hamming windowing. We iterated the process for few times by updating the autoregressive filter coefficients after every repetition. Even though the process takes long time for a tiny speech signal data, the output can be compared with input for its similarity.

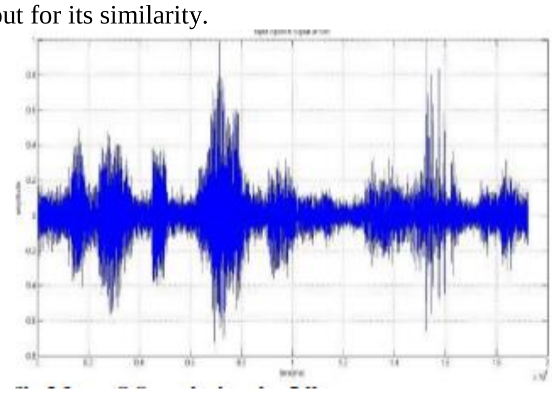

Fig.3.2. Input Speech signal – 5db.

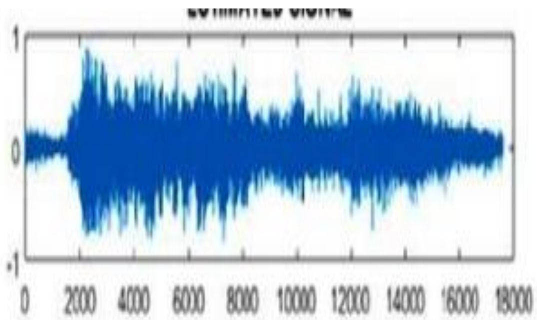

Fig.3.3. Speech enhancement of entire signal by Kalman filter.

IV. CONCLUSION

The enhancement of speech is important in various fields of communication and we have proposed kalman model which estimates the output, Kalman filter is implemented using

NOZIEUS database. The windowing process model has some big matrices (P and A for instance), whose sizes are determined by choosing appropriate autoregressive filter order. The process is slow and it is not surprising given the number of matrix multiplications it has to do for every sample. The filter have some advantages compared to LMS, RLS and wiener, it works on real-time execution without storing observations or previous estimates, provides variance of the estimation error. The filter doesn't need any memory as it works in real time and is also good for stationary and non-stationary signals. Since the proposed algorithm has been enhanced the speech in efficient way. Hence, in future, it is expected to work better on music enhancement by using this algorithm. In case of low order of the AR (Autoregressive) model, the harmonic structure of music is often lost. Further advancement in this work is to test the algorithm with automatic order determination on music signals.

V. REFERENCES

[1]. Hu.Y and Loizou.P ," Evaluation of objective quality measures for speech enhancement," IEEE Trans. On speech and audio processing, 2008.

[2]. B. Chabane and B. Daoued " On the use of Kalman Filter for enhancing speech corrupted by Colored Noise," WSEAS Trans. On signal processing 2008 Dec.

[3]. "Speech Enhancement" by J. Benesty , J. Chen, Eds., and S. Makino, Springer, Berlin, 2005.

[4]. J. Jensen and J.H.L Hansen , " Speech enhancement using a constrained iterative sinusoidal model" IEEE Trans. Speech Audio Process., vol.9 ,2001.

[5]. J.Jensen,"Speech enhancement using a constrained iterative sinusoidal model,"vol.9,2001.

[6]. Paliwal KK, Basu A. A speech enhancement method based on Kalman filtering. Proceedings of ICASSP'; Dallas, TX, USA. 1987; 7(5):510–24.

[7]. Ganesan V, Manoharan S. Surround noise cancellation and speech enhancement using sub band filtering and spectral subtraction. Indian Journal of Science and Technology. 2015 Dec; 8(33). DOI: 10.17485/ijst/2015/v8i33/77774.

[8]. Kallakunta RK, Bhavani L, Spandana P, Rishitha K, Ashitha G. Removal of real world noise in speech: Comparision of various parameters using kalman and h-infinity filter algorithms. Indian Journal of Science and Technology. 2016 Aug; 9(30).