Speech Enhancement Using Rasta and LPC

Gsharanya & M Ranjith

[*]PG Scholar, Department of ECE, Vaagdevi College of Engineering, Warangal, Telangana.

[**]Assosciate Professor, Department of ECE,Vaagdevi college of Engineering, Warangal, Telangana.

ABSTRACT:*The enhancement of speech is to recover and improve the speech quality and its intelligibility by using different techniques and algorithms. An iterative Kalman filter model is using the linear predation coefficients (LPC) parameter to estimate the noise present in degraded speech. But LPC parameters are sensitive with respect to different types of noise. To overcome this problem Kalman filter with overlapping frames are used. The cochlear implant is used for those who are suffering from human ear problem and it needs speech enhancement. The main challenge in enhancement of speech is analysis of speech signal features and designing of efficient filters. The complex speech spectrum is used to enhance the noisy speech. At low-frequency analysis of speech, substantially in order to bubble noise and car interior noise so that it contain a lot of energy at low frequency, more accurate result is given by iterative Wiener filter. A new method to change the magnitude and phase, a modified complex spectrum is to provide a better estimation. Another algorithm using feedback particle filter will reduce minimum mean square error (MMSE) using iterative method. A new algorithm is also there in literature using zero replacement signal of noisy speech. This method is using Fourier transform (FT) of cross-correlation function between noisy speech signal and zero replacement signals. The speech enhancement is using a pre-filter method to obtain a better auto regression (AR) co-efficient by using temporal and the simultaneous masking threshold. For the estimation of single channel speech enhancement algorithms, whose parameters are estimated by codebook method, the use of short time Fourier transforms (STFT) to reconstruct the noise observation and fundamental frequency of speech has been done. A non-zero mean condition that builds up by the Bayesian short time spectrum amplitude (STSA) algorithm is introduced in the stochastic deterministic (SD) speech. An analytical function is used for the magnitude of STFT of speech signal in the form of MMSE and phase in maximum-likelihood. Keywords Speech Enhancement; Amplitude and phase estimation; Gaussian*

process; Stochastic model; Kalman filter; Wiener filter; Particle filter; Cochlear implant.

1. INTRODUCTION

The speech signal is degraded by noise environment like car, train, aircraft, and background noise. The speech signal is also deteriorated due to additive noise. Thus noise signals create problems in real world like radio communications. There are numerous methods for the speech enhancement such as Kalman filter, Wiener filter, zero replacement noisy speech, STFT, MMSE estimation and spectrum substation methods. The Kalman filter is used to estimate the unknown state of a dynamic system using a linear combination of degraded speech and observing accurate speech signal is process by an iterative kalman filter of a noise degraded observation. The joint estimation of both magnitude and phase spectrum is observed by the nonstationaryKalman filter [1]. The use of analysis modification synthesis (ASM) on the filter bank technique, this algorithm using recursive Kalman filter in a Gammatone filter to enhance both the instantaneous amplitude and the phase, because it is deceptively improve speech quality and intelligibility. Thus technique is helpful in human cochlear filter bank where cochlea is a part of human ear and it is used to detect the strength of quality audio

signal [2, 3] .The pre-filtering is construction based on the temporal and simultaneously masking threshold and this method is applied to obtain a better AR coefficient in the speech production. The perceptual weighting filter technique is used for reducing the code noise from system and this technique is driven by AR speech algorithm [4]. The case of magnitude and phase spectrum compensation algorithm is Rajesh Kumar Dubey only to change the magnitude spectrum that is modified and phase spectrum remains unchanged. The combination of modified magnitude spectrum and unchanged phase spectrum are producing a modified complex spectrum. This estimation technique is a clean speech based magnitude spectrum method. In this method, magnitude noise is subtracted by the spectrum subtraction method [5]. In iterative Wiener filter (IWF) algorithm, it is provide a more accurate complex speech power spectrum and optimized distorted speech signal. This algorithm provides high performance at lower frequency for the case of bubble noise and car interior noise. In the complex linear prediction coding (LPC) analytic processor, analytic signal is observer the real part and Hilbert transform observer the imaginary part [6]. In this technique Kalman filter work based on speech enhancement frame work and the use of STP parameter for the

function of Kalman filter to estimate the approximation expectation-maximization (EM) algorithm. The main work in this paper is codebook based approach to evaluate the STP parameter [7]. Another algorithm is introduced in this paper is lower SNR and non-stationary noise environment. This algorithm uses cross-correlation function between the speech signal and zero replacement single of speech signal. The cross-correlation function is used due to the periodicity of the speech signal and this method is affected due to the non-stationary signal, because the past of signal is not utilized so this method avoids zero lag autocorrelation function . The problem of a non-linearity and nonGaussian is overcome by the use of feedback particle filter. In this feedback particle filter, it will provide a feedback which will be updated in every iterative step for enhancement of speech . The STFT estimate of the clean speech amplitude spectrum by spectrum substation is given in .The short time spectrum amplitude (STSA) estimation and long-spectrum amplitude (LSA) estimation are introduced by Ephraim and Malah. The non-zero mean hypothesis is the observation of the MMSE STSA process to characterize speech at the same instant to both the stochastic and the deterministic. The Stochastic and deterministic speech model as developed in Wiener filter

framework are given in This stochastic MMSE-STSA speech signal framework is present in this literature. The objective estimation of speech quality is important in speech processing procedures to observer and sustains of the quality of service (QoS) and system computerization and automation. The subjective speech quality is calculated by the absolute category rating (ACR) method and the objective speech quality is calculated by different computational techniques. The objective quality technique is computed in term of correlation between the subjective MOS and the objective MOS .

2 SPEECH ENHANCEMEN.

In this paper, different algorithms used for speech enhancement are reviewed and explained. A signal and noise model is written in form of state-space equation for the Kalman filtering. In this method, an unbiased and linear MMSE estimate xnn of the state vector x(n) at time n, are recursively calculated by the use of Kalman filter and a better LPC is obtain in initial iteration of speech by this process of Kalman iterative approach for overlapped frame. The Kalman parameter such as Kalman gain, error covariance, and state vector estimate xnn is continuous to update each iteration on sample by sample [1]. The second approach is

ASM on a Gammatone filter bank and it is used to improve the instantaneous amplitude and phase, for the modeling of three steps constants such as Analysis stage, modification stage, re-synthesis by inverse gamma tone filter bank. The Kalman filter used recursive solution for speech enhancement, which is widely used for statistical processing. The Kalman filter equation for state and observation is defined as [2].

$$S[m] = F.S[m-1] + W[m] \quad (1)$$

$$O[m] = H.S[m] + V[m] \quad (2)$$

Where, $S[m]$ is the state for discrete time m, $O[m]$ is observation for discrete m, $W[m]$ is drive noise and $V[m]$ is observation noise. To calculate the optimal estimate there are five steps for both representations, the instantaneous amplitude and the instantaneous phase [2]. There are two methods for estimating the LP coefficient, non-blind speech and blind speech. We can use blind speech method, whenever the noise speech signal is available and LP method is used for tanning phase. There are two methods used for estimating the LP coefficient, auto-correlation and covariance methodsand the signal to error (SER) is used for the restored speech signal [2]. The ordinary symmetric KullackLeibler distance (SKLD) sandwiched

between histogram and parameter distribution. In other algorithm, it is to show that how analog channel is transmitted from filter in cochlea implant of human ear. The fractional delay filter is used to overcome the overlapping problem in Conference Proceeding of 4th International Conference on Recent Development in Engineering, Science, Management and Humanities (ICRESMH-2017) at India different channel implantation device with suitable filtering techniques [3].
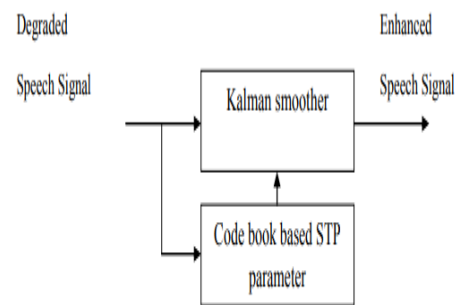

Fig 1: Basics of speech enhancement algorithm.

The pre-filter method is constructing the approach of masking properties of human auditory system together with time and frequency domain masking effect. The individual masking and simultaneous masking thresholds are to give the resultant of overall masking threshold processor [4]. The Kalman filtering technique has been used for signal channel speech enhancement and in figure 1, it is shown that degraded signal is fed as an input

to Kalman filter and codebook based STP parameter is estimated by the Kalman smoother. The Kalman filter speech enhancement perspective requires the AR signal model. The MMSE estimation parameter is also using codebook based approach. Degraded Enhanced Speech Signal Speech Signal Fig 1: Basics of speech enhancement algorithm. The parameter used in codebook based method for MMSE estimation and the use of LPC for speech and noise stored in trained code book are utilized to estimate STP parameter by the use of Kalman smoother. In IWF, the spectrum is estimated using speech power spectrum by LPC analysis for an input speech signal. Two power spectrums of speech and noise are estimated by the TV-CAR speech analysis in its place of LPC analysis [6]. In zero replacement process, half of the signal is replaced by zero by using seed number process. This method decreases distortion due to the noise by least numbers based on cross-correlation function between the noisy speech signal and zero replacement signals . The speech enhancement uses feedback particle filter. This will update at each iterative step for speech signal. Whereas adaptive filter has possibility to be improved nears the original approximation. The particle filter is flexible and power full tool to design and approximation of

the Gaussian and non-Gaussian, linear and also nonlinear. The is feedback provided at each iterative step to update the correction in the error. The gain and error calculation are used in the feedback particle equation governed by [9]. The observation of speech signal in the modification of magnitude spectrum which will be used for STFT domain is given. A best resultant of Ak is observed by y[n] and this resultant mentioned by the modification of Bayesian MMSE STSA speech .

$$Ak = EAkBk = \ , \beta k = \beta k \ (3)$$

In non-zero mean complex case STFT phase is normalized by the utilization of past and future frame and hold the current frame.

3.The Stochastic DieterMinistic Mass Effect Estimate

The MMSE and STSA estimation problem in the case of Xk and Dk variable are assumed to be an independent value for all k. The resultant of the problem to find out the expectation in case of magnitude of complex Gaussian with non-zero mean is given as.

$$\zeta k= \ \lambda x, \ \lambda x,+\lambda d,kBkei\beta k \ + \ \mu kei\theta k \ 1 \ − \ \lambda x,k\lambda x,k+\lambda d,k \ (4)$$

The value of parameter dependent on a value of μk which have been two cases considered $\mu k =$

0 and $\mu k \neq 0$. In the resultant SD MMSE STFT estimation, the case where $\mu k = 0$, qk is the observation yk with some attenuation

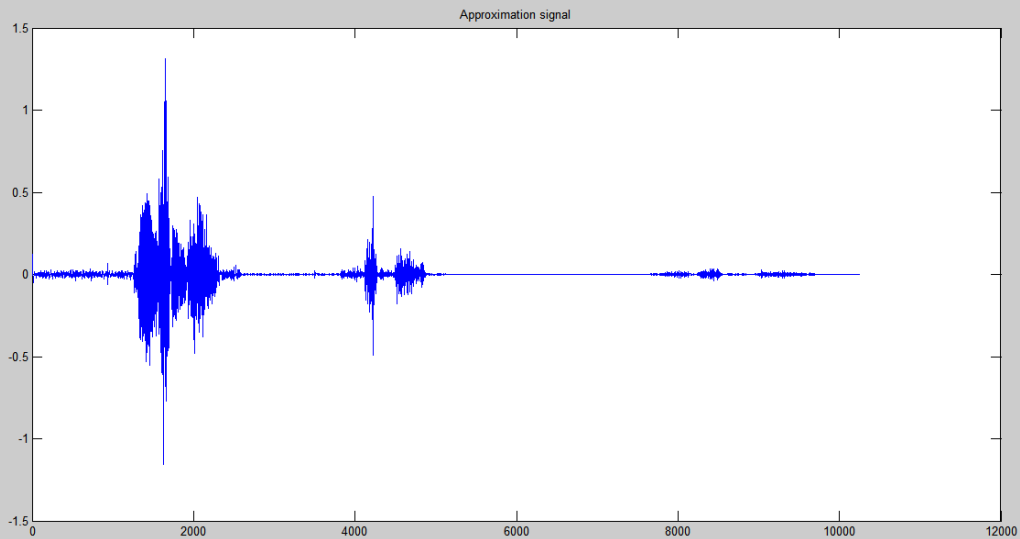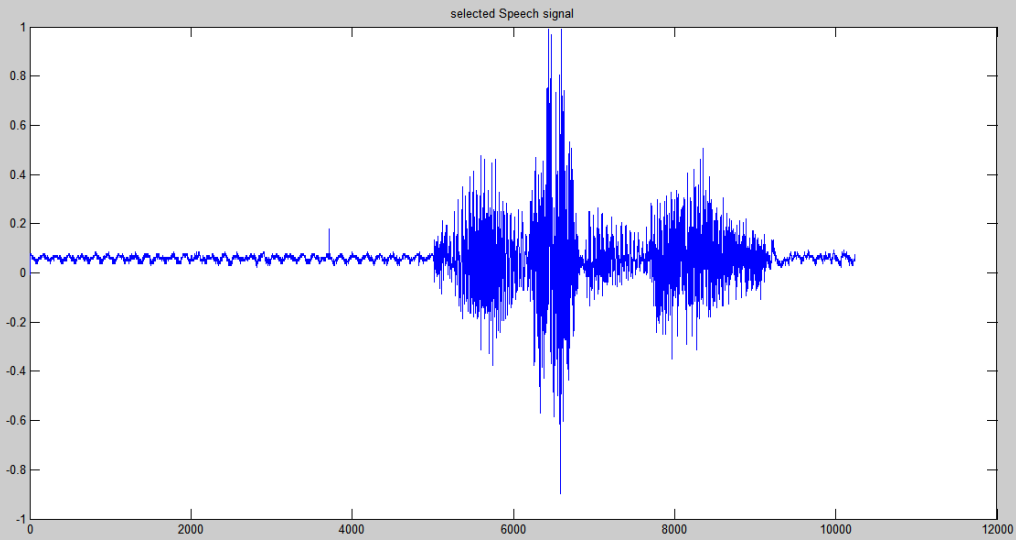$$\zeta k \mu k = 0 = \Omega k \, . \, \Omega k = \xi k \, 1 + \xi k \, (5)$$

Where, Ωk is the DFT frequency. The clean speech magnitude estimation Ak dependent on ζk and λk is dependent on the power ratio of two quantity.
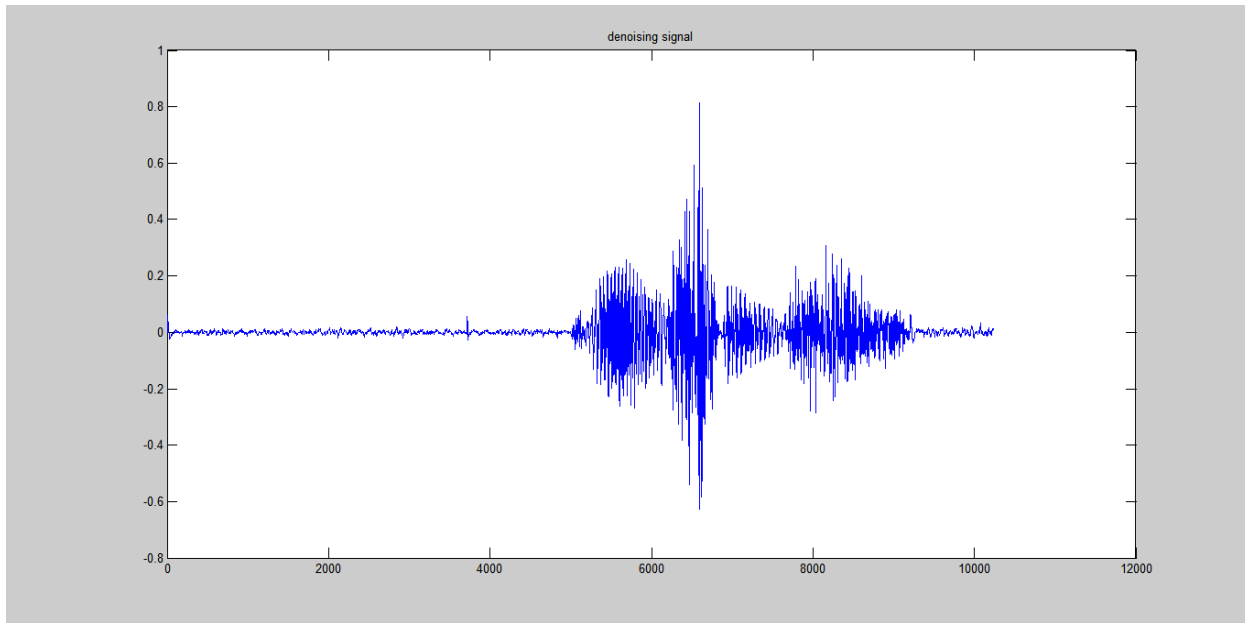
$$vk = \zeta k \, 2 \, \lambda k \, (6)$$

In 2nd case, when $\mu k \neq 0$ the value of Yk is dependent on the observation of ζk in (4) which may be large and small in magnitude thus is effecting the observation . If we set the exact value of ζk this is dependent on the value Wiener term Ω . For the same observation of case $\mu k = 0$ and the value of $\Omega k \approx 1$. Thus resultant Ωk is approximation equal or close to the observation of Yk in complex case. The value ζk is given in magnitude and phase difference and its value decreases when phase value increases. The use of BESSEL function is to estimate the magnitude and αk to give the resultant of maximum likelihood (ML). Where maximum likelihood condition gives the resultant of clean speech signal phase that is more accurate in noisy phase condition. The ML condition also used for the pitch track from a

clean speech process. The harmonic pulse noise speech model implement in non-zero mean concept with the help of STFT observation of speech signal, under this condition amplitude is estimated by analytical signal expression were phase is obtained by (ML) sense. This method recover sinusoidal component at low SNRs. The SD algorithm is used to minimize Gaussian distribution with correlated spectral components. The statistical analysis of listening test for speech signals at different categories noise suppression scheme to process the speech sentences of the databases NOIZEUS-2240 is given in [12]. In this, the correlation between the subjective MOS and the objective MOS is also compared

RESULTS:

International Journal of Research

Available at https://edupediapublications.org/journals

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue14
November 2017

CONCLUSION

In this survey, the different methods of speech enhancement algorithms to improve speech quality and intelligibility are studied such as speech enhancement method based on Kalman filtering, feedback particle filter, Wiener filtering, zero replacement method etc. The different performance parameters compared are MSE, MMSE, SS, PSE and SSUB under different noisy environmental condition.

REFERENCES

[1] Mellahi, N. and Hamdi, R., "LPCs enhancement in Iterative Kalman filter for speech enhancement using overlapped frames," 2014 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM), Tunis, 2014, pp. 1-5.

[2] Nower, N. and Unoki, M., " Restoration of instantaneous amplitude and phase using Kalman filter for speech enhancement," 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, 2014, pp. 4633-4637.

[3] Kashwan, K, R. and Sanketha, M., "Implementation of cochlear implant for speech enhancement and timbre detection using fractional delay filter," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, 2016, pp. 1-4.

[4] Wang, Y. and Ambikairajah, E., "Perceptually motivated pre-filter for speech enhancement using Kalman filtering," 2007 6th International Conference on Information, Communications & Signal Processing, Singapore, 2007, pp. 1-4.

[5] Ren, H. and Shimamara, T., " Speech enhancement using magnitude and phase spectrum compensation," 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), Okayama, 2016, pp. 1-4.

[6] Fanaki, K., "Speech enhancement based on iterative Wiener filter using complex speech analysis," 2008 16th European Signal Processing Conference, Lausanne, 2008, pp. 1-5.

[7] Kavalekalam, M., S. and Gran, F., " Kalman filter for speech enhancement in cocktail party scenarios using," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 2016, pp. 191-195.

[8] Iwai, Y. and Shimamara, T., " Speech enhancement with zero replacement signals at Low SNR," 2011 International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS), Chiang Mai, 2011, pp. 1-4.

[9] Singh, R. and Mishra, A.,"Speech enhancement using feedback particle filter," 2015 International Conference on Computational Intelligence and

Communication Networks (CICN), Jabalpur, 2015, pp. 682-685. Conference Proceeding of 4th International Conference on Recent Development in Engineering, Science, Management and Humanities (ICRESMH-2017) at India International Center,Max Muller Marg, Delhi, India on 8th January 2017 ISBN: 978-81-932712-3-0 45 BrijendraYadav, Rajesh Kumar Dubey

[10] Krawczyk, M. and Gerkmann, T., "STFT phase reconstruction in voiced speech for an improved single-channel speech enhancement," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 22, no. 12, pp. 1931-1940, Dec. 2014.

[11] Mccallum, M. and Guillemin, B., "Stochastic–Deterministic MMSE STFT speech enhancement with general aprior information," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 7, pp. 1445-1457, July 2013.

[12] Dubey. R., K. and Kumar.A., "Comparison of subjective and objective speech quality assessment for different degradation / noise conditions," 2015 International Conference on Signal Processing and Communication (ICSC), Noida, 2015, pp. 261-266