ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Dual Adaption based Speech Enhancement & Echo cancellation in Mobile Communication

Vijay Kiran Battula[1], Nagaraju Tatipudi[2]

M. Tech Scholar, Department of ECE, JNTUK-UCEV, Vizianagaram, Andhra Pradesh, India[1]

M. Tech Scholar, Department of ECE, GMRIT, Srikakulam, Andhra Pradesh, India[2]

ABSTRACT: A scenario of requirement for a robust algorithm to deal with both Speech enhancement and Echo cancellation is seen in mobile communication**.**Speech enhancement and Echo cancellation are adaptive signal processing applications and they use adaptive algorithms. In this paper, a study on situations involving both the applications in mobile phone communication is done. General Kalman Filter is used along with estimate maximization framework to perform dual adaption and effectively handle both the applications simultaneously. Simulation results show it is dominant over popular adaptive algorithms

KEYWORDS: Speech Enhancement; Echo cancellation; General Kalman Filter; EM frame work

I. INTRODUCTION

Mobile Communication is mostly based on speech as a base for information exchange. A person using a mobile phone is evident to move from place to place. Speech is almost a non-stationary signal, and thus it is always a challenging problem to model it perfectly with all possible considerations. On the other hand speech is every time degraded by various noise characteristics present along with it. Hence, we require methods to deal with the issue. Speech enhancement is performed to increase intelligibility and enhance quality of speech under the specific conditions of degradation. Echo cancellation is seen when the reverberated sound produced by the system in an acoustic atmosphere of far end speaker corrupts the near end speaker speech.

In Echo cancellation, eliminating the background noise caused due to reverberation phenomenon is required. In fact the background noise is nothing but estimating the echo path and decreasing its effect on speech of near end speaker. But this in turn refers to a speech enhancement application indirectly. Most of the literature concentrates more on decreasing degradation of near end speaker caused by reverberating signal of far end speaker and least considers the enhancement of near end speech addicted by other noise phenomena in the acoustic system.

Speech enhancement and Echo cancellation are performed through dual adaption procedure using General Kalman Filter[4].There are popular algorithms available for speech enhancement and echo cancellation respectively. It is aimed to use GKF for both the applications and show it is optimum for dual adaption than some well-known adaptive algorithms. General Kalman filter is first used for Echo cancellation[4]. It is time domain derived form of kalman filter. In present work it is extended to speech enhancement. Therefore the strong resemblance between the algorithms used in the two applications gave the idea to develop an integrated method involving both the applications which can utilized in the area of mobile communication.

II. RELATED WORK

There is a need to study the existing methods and their advantages on demand basis of the application. The classification of different algorithms developed in past are mostly of either frequency domain or time domain. Using kalman filter for statistical analysis is quite efficient, but implementing it in time domain had not shown better results. The kalman filter implementation in frequency domain of the signal is seen in the work of G.Enzner[2] with better performance. Spectral subtraction is frequency domain method proposed by Boll[8] stands effective even today but it needs a good voice activity detector. The work of Widrow[7] is still in use for many speech applications but it has very lower convergence. Affine projection algorithm is variable step size filter type and its ability of varying the step size led to early convergence than LMS(Least mean square) algorithm and it is also prominent in certain specific

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

applications. Recursive least squares is more compared to General Kalman filter in every significant stage of their algorithm procedures respectively[4]. RLS(Recursive least squares) has computational complexity due to matrix inversion. The periodicity based type has techniques like using comb filter with taps for different pitch periods. Speech production based models are also present like linear predictive coding (LPC) modelling of the signal. Whenever we speak about optimum filtering wiener filter satisfy the most of the conditions provided when both the noise and speech power spectral densities are measurable and are non-stationary.

III. Problem Statement and Proposed Method

A. *Problem statement:*

$z(n)$ is a noisy near end speech signal, $s(n)$ and $v(n)$ are its two components. $s(n)$ represents the clean speech signal and it has to be separated from additive noise $v(n)$.

$$z(n) = s(n) + v(n) \tag{1}$$

The above signal is added with reverberating signal $y(n)$ of far end speech signal $x(n)$ at near end which is reproduced and $d(n)$ is desired output of microphone with impulse response of present acoustic system h .Our objective is to estimate with an adaptive filter $\hat{h}$

$$y(n) = x(n)h \tag{2}$$

$$d(n) = z(n) + y(n) \tag{3}$$

B. *Description of the Proposed Algorithm:*

The far end signal received by the mobile phone is a known signal but the reproduction environment and response signal of this signal in that particular acoustic environment is not known. We use an adaptive filter(GKF), the same far end signal which reproducing is given as input to it. Adaption is performed such that the filter response is almost similar to acoustic response of the environment it is reproduced. The near end side microphone picked up signal contains the reverberating signal and is cancelled by adaptive filter output. All the above stated falls under Echo cancellation stage of proposed algorithm. A simultaneous stage of enhancement performing adaptive filter is also fed with input of speech absent periods as initial condition, which is nothing but noise of the acoustic environment through the same micro phone. The second filter is proposed with General Kalman filter and Estimate maximization frame work. Since, noise and speech are non-stationary no algorithm works without knowing the characteristics of at least one among them to separate one from another. Noise obtained in speech absent periods is modelled as Auto Regressive process and used in adaption procedure. Parameters of AR model is also updated accordingly for every fixed number of iterations of adaption. Finally, the outputs of two adaptive filters are subtracted from recorded signal of microphone. Thus we obtain a clean speech signal free from echo and noise.

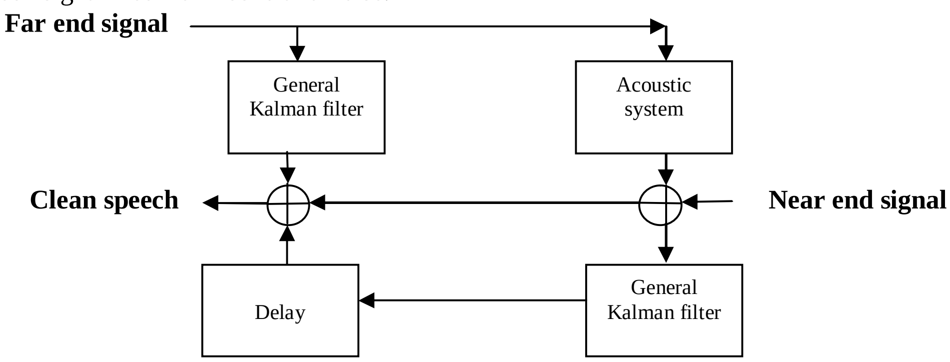

Fig.1.Block diagram of the proposed method

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

IV. PSEUDO CODE

Step 1: Consider speech absent periods collect noise characteristics.
Step 2: Initial noise characteristics are modeled as Auto regressive model delayed until the speech presence is seen.
Step 3: A situation of reproduction of far end signal at near end signal is considered.
Step 4: A combined signal of reverberated, near end speech and background noise is collected using microphone.
Step 5: In the initial step the echo cancellation is performed using General Kalman filter.
Step 6: Hence we obtain reverberation effects free signal now enhancement of this signal is performed.
Step 7: GKF and EM frame work considers the delayed noise signal and performs adaption based on it.
Step 8: This recursive process is performed for fixed iterations to obtain a clean speech.

V. FORMULATION

A. *Echo cancellation:*

The following are equations related to echo cancellation algorithm.

$$d(n) = z(n) + h_x^T(n)x_k(n) \tag{4}$$

$$\hat{y}(n) = h_x^T(n)x_k(n) \tag{5}$$

$$\hat{h}_x(n) = \hat{h}_x(n-1) + w_x(n) \tag{6}$$

Initialize with $\hat{h}_x(n) = 0$ and $R_{\mu x}(0) = \varepsilon_x I_{qx}$ where ε_x is small positive constant

$$\hat{h}_x(n) = \hat{h}_x(n-1) + w_x(n) \tag{7}$$

$$d(n) = z(n) + h_x^T(n)x_k(n) \tag{8}$$

$$e_x(n) = d(n) - \hat{y}(n) \tag{9}$$

$$R_{mx}(n) = R_{\mu x}(n-1) + \sigma_{wx}^2(n)I_{qx} \tag{10}$$

$$R_{ex}(n) = x_k^T(n)R_{mx}(n)x_k(n) + \sigma_x^2(n)I_{px} \tag{11}$$

$$K_x(n) = R_{mx}(n)x_k(n)R^{-1}{}_{ex}(n) \tag{12}$$

$$e(n) = z(n) - x_k^T(n)\hat{h}_x(n-1) \tag{13}$$

$$\hat{h}_x(n) = \hat{h}_x(n-1) + K_x(n)e(n) \tag{14}$$

$$R_{\mu x}(n) = [I_{qx} - K_x(n)x_k{}^T(n)]R_{mx}(n) \tag{15}$$

Where $R_{mx}(n)$ is priori misalignment, $R_{\mu x}(n)$ is posteriori misalignment correlation matrix, $K_x(n)$ is Kalman gain, I_{qx} is identity matrix, $\sigma_{wx}^2(n)$ is variance of process noise $w_x(n)$, $R_{ex}(n)$ is priori error vector correlation matrix and $e_x(n)$ is error between signal and estimated.

B. *Speech enhancement*

Let us consider the problem in a discrete time index of n,

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

$$d(n) = z(n) + y(n) \tag{16}$$

Noise and speech are non-stationary. Hence, any one of them has to be known in order to go for formulation. Consider the speech absent periods to study noise characteristics,

$$z(n) = v_{ab}(n) \tag{17}$$

$$d(n) = v_{ab}(n) \tag{18}$$

Where $v_{ab}(n)$ is signal collected by recording device when speech is absent and use as initial signal vector of noise. Noise is modelled as stochastic AR process:

$$v(n) = -\sum_{q=1}^{k} \alpha_q v(n-k) + g_v u_v(n) \tag{19}$$

$\alpha_1, \alpha_2 \ldots\ldots \alpha_k$ represents AR parameters of noise process and g_v represents power level. $u_v(n)$ is normalized (zero-mean unit variance),white Gaussian noise.

AR model of noise process is converted into state space formulation $v_k^{\,T}(n) = [v(n-k+1), v(n-k+2), \ldots\ldots, v(n)]$

$$\Phi_v = \begin{bmatrix} 0 & 1 & 0 & 0 & \ldots\ldots\ldots\ldots & 0 & 0 \\ 0. & 0 & 1 & 0 & \ldots\ldots\ldots\ldots & 0 & 0 \\ . & . & . & . & & . & . \\ . & . & . & . & \ldots\ldots\ldots\ldots & . & . \\ . & . & . & . & & . & . \\ . & . & . & . & \ldots\ldots\ldots\ldots & . & . \\ -\alpha_k & -\alpha_{k-1} & -\alpha_{k-2} & -\alpha_{k-3} & \ldots\ldots\ldots\ldots & -\alpha_2 & -\alpha_1 \end{bmatrix}$$ is noise transition matrix and k-dimensional vectors $g_v^T = [0 \quad 0 \quad \ldots\ldots \quad g_v]$.The below equation provides estimate of noise

$$v_k(n) = \Phi_v v_k(n-1) + g_v(n) u_v(n) \tag{20}$$

$$z(n) = s(n) + h_v^T(n) v_k(n) \tag{21}$$

$$d(n) = s(n) + h_v^T(n) v_k(n) + h_x^T(n) x_k(n) \tag{22}$$

Parameters for estimate and AR model, GKF algorithm procedural steps results in finding the minimal error and optimal estimate. The following are the related equations

Initialize with $\hat{h}_v(n) = 0$ and $R_\mu(0) = \varepsilon I_q$ where ε is small positive constant

$$\hat{h}_v(n) = \hat{h}_v(n-1) + w_v(n) \tag{23}$$

$$z(n) = s(n) + h_v^T(n) v_k(n) \tag{24}$$

$$e(n) = z(n) - \hat{v}(n) \tag{25}$$

$$R_m(n) = R_\mu(n-1) + \sigma_w^2(n) I_q \tag{26}$$

$$R_e(n) = v_k^{\,T}(n) R_m(n) v_k(n) + \sigma_v^2(n) I_p \tag{27}$$

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

$$K(n) = R_m(n) v_k(n) R_e^{-1}(n) \tag{28}$$

$$e(n) = z(n) - v_k^{\ T}(n)\hat{h}_v(n-1) \tag{29}$$

$$\hat{h}_v(n) = \hat{h}_v(n-1) + K(n)e(n) \tag{30}$$

$$R_\mu(n) = [I_q - K(n)v_k^{\ T}(n)]R_m(n) \tag{31}$$

Where $R_m(n)$ is priori misalignment, $R_\mu(n)$ is posteriori misalignment correlation matrix, $K(n)$ is Kalman gain, I_q is identity matrix, $\sigma_w^2(n)$ is variance of process noise $w_v(n)$, $R_e(n)$ is priori error vector correlation matrix and $e(n)$ is error between signal and estimated.

Let θ be the vector of unknown parameters and given as vector $\theta^T = \left[\alpha^T \quad g_v\right]$

$$X = v_q(n-1)v_q^T(n-1) \tag{32}$$

$$Y = v_q(n-1)v(n) \tag{33}$$

Where updated parameters are defined as

$$\hat{\alpha}^{(q+1)} = -\sum E[X]\sum Y \tag{34}$$

$$\hat{g}_v^{(q+1)} = \sum\left[Z + \left(\hat{\alpha}^{(q+1)}\right)^T E[Y]\right] \tag{35}$$

VI. Results

The NOIZEUS speech corpus is used to test the proposed method. Simulations are performed using MATLAB software. Fig.2 is Noisy speech before dual adaption procedure ,Fig.3 is output of the Echo cancellation stage and Fig.4 is output of the speech enhancement stage an enhanced noisy speech. A known signal is taken as the far end signal and is used for echo cancellation and a random noise is considered for speech enhancement.The objective measures of speech quality assessment namely signal to noise ratio(SNR),mean square error(MSE), perceptual evaluation of speech quality (PESQ),Segmental SNR(SegSNR),log likelihood ratio(LLR),Weight Spectral Slope(WSS) are calculated for proposed method and compared with existing methods.

Algorithm	SNR	MSE	PESQ	WSS	LLR	SegSNR
Affine projection algorithm	18.6964	0.0190	0.82382	84.548	1.1378	-5.2797
Least mean square	5.3090	0.5001	0.873704	179.730	1.3874	-10.00
Normalized least mean square	11.3423	0.0839	1.7459	163.742	1.9886	-9.9957
Kalman filter	9.4235	0.0683	1.1657	203.258	1.8601	-9.9028
General Kalman filter	19.7544	0.0164	1.3746	228.4815	1.5989	-9.9799

Table.1. Objective measures comparison of methods for dual adaption on SNR 5dB babble noise type

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

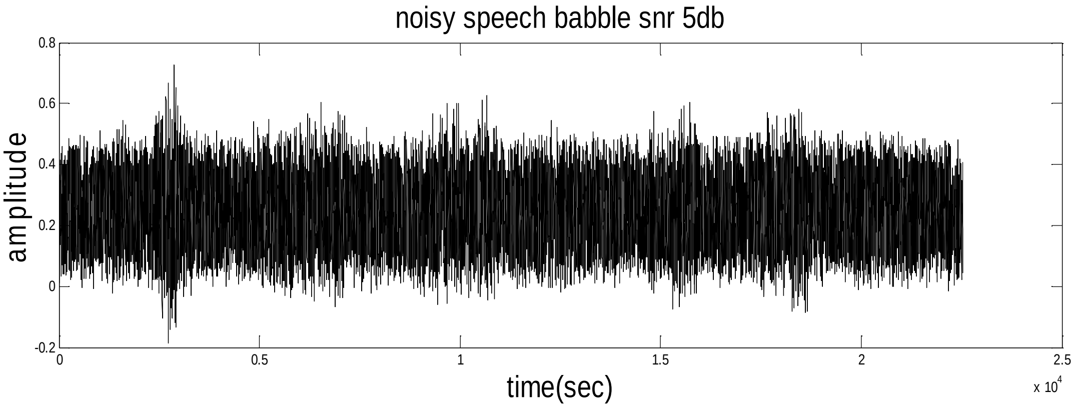

Fig.2. Noisy Speech signal before dual adaption

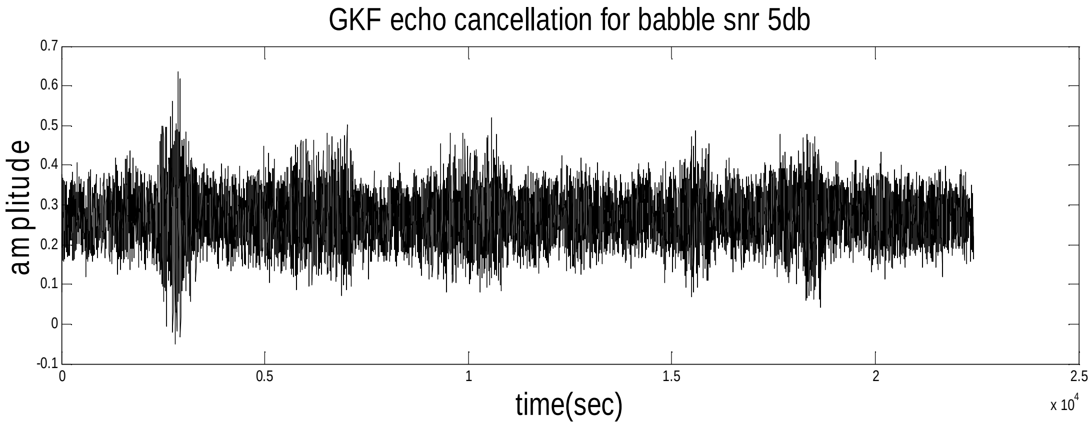

Fig.3. Signal after echo cancellation stage

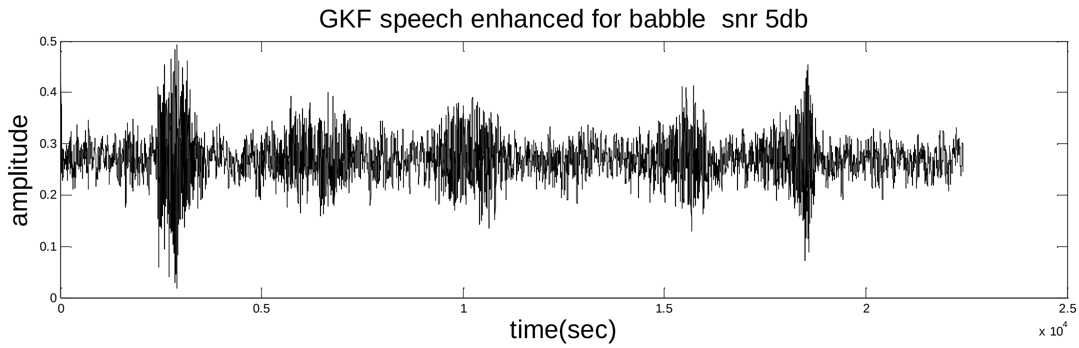

Fig. 4.Signal after speech enhancement and echo cancellation

VII. CONCLUSION AND FUTURE WORK

Results of simulations performed in MATLAB are better for the proposed algorithm in Table.1. The Enhancement stage uses modeling of only noise as AR process because it is non-stationary considering speech absent period. The performance of the proposed algorithm is analyzed for off-line procedure. Future work is to extend the algorithm for on-line procedure. Similarly better performance can be expected if speech is also modeled.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

REFERENCES

1. R. E. Kalman, “A new approach to linear filtering and prediction problems”, J. Basic Eng., vol. 82, pp. 35–45, Mar. 1960.
2. G. Enzner and P. Vary, “Frequency-domain adaptive Kalman filter for acoustic echo control in hands-free telephones ”, Signal Process., vol.86, pp. 1140–1156, 2006.
3. Jae Lim, A. Oppenheim, “All-pole modeling of degraded speech”, IEEE Transactions on Acoustics, Speech, and Signal Processing, Volume: 26, Issue: 3, Year: 1978
4. C. Paleologu, J. Benesty, and S. Ciochina, “Study of the General Kalman filter for echo cancellation”, IEEE Trans. Audio, Speech, Language Processing, vol. 21, pp. 1539–1549, Aug. 2013.
5. S. E. Bou-Ghazale and K. Assaleh, “A robust endpoint detection of speech for noisy environments with application to automatic speech recognition”, in Proc. ICASSP2002, vol. 4, pp. 3808–3811, 2002.
6. S. Gannot, “Speech processing utilizing the Kalman filter” , IEEE Instrum. Meas. Mag., vol. 15, no. 3, pp. 10–14, Jun. 2012.
7. B. Widrow, J.R. Glover Jr., J.M. McCool, J. Kaunitz, C.S. Williams, R.H.Hearn, J.R. Zeider, E. Dong Jr., and R.C. Goodlin, ”Adaptive Noise Cancelling: Principals and Applications”, Proceeding of the IEEE, 63(12):1692–1716, Dec 1975.
8. S. F. Boll. “Suppression of acoustic noise in speech using spectral subtraction”, Jae S. Lim, editor, Speech Enhancement, Alan V. Oppenheim series, pages 61–68. Prentice-hall, 1983.
9. Yi Hu and Philipos C. Loizou, “Evaluation of Objective Quality Measures for Speech Enhancement”, IEEE Transactions on audio, Speech, and language processing, Vol. 16, No. 1, January 2008.

BIOGRAPHY

Battula Vijay Kiran had received B.Tech degree from JNTUK university and currently pursuing M.Tech degree(Systems and Signal Processing) in JNTUK University College of Engineering Viziangaram, Andhra Pradesh, India.

Tatipudi Nagaraju had received B.Tech degree from JNTUK university and currently pursuing M.Tech(VLSI and Embedded systems) degree in GMRIT, Rajam, Srikakulam, Andhra Pradesh, India.