World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:8, No:3, 2014

Speech Enhancement Using Kalman Filter in Communication

Eng. Alaa K. Satti Salih

***Abstract*—**Revolutions Applications such as telecommunications, hands-free communications, recording, etc. which need at least one microphone, the signal is usually infected by noise and echo. The important application is the *speech enhancement*, which is done to remove suppressed noises and echoes taken by a microphone, beside preferred speech. Accordingly, the microphone signal has to be cleaned using digital signal processing DSP tools before it is played out, transmitted, or stored. Engineers have so far tried different approaches to improving the speech by get back the desired speech signal from the noisy observations. Especially Mobile communication, so in this paper will do reconstruction of the speech signal, observed in additive background noise, using the Kalman filter technique to estimate the parameters of the Autoregressive Process (AR) in the state space model and the output speech signal obtained by the MATLAB. The accurate estimation by Kalman filter on speech would enhance and reduce the noise then compare and discuss the results between actual values and estimated values which produce the reconstructed signals.

***Keywords*—**Autoregressive Process, Kalman filter, Matlab and Noise speech.

I. Introduction

SPEECH is a form of communication between human civilizations which has existed since long time ago till now and it still remains as the most desirable medium of communication between humans.

Human beings live in the digital information age, so speech applied to different technologies, such as the network telephone, cellular and satellite technology with digital speech, have become widely used today because information stored in a compressed digital form is more efficiently transmitted, stored and is easier to process. However, human beings always look for high quality, so telecommunication researchers/ engineers are trying various approaches of improving and enhancing the speech signals. [1]

In order to achieve this enhancement, Sridharan, Leis and Paliwal stated that the main attributes of a speech coder such as bit rate, subjective quality, complexity, delay, error. [2]

Sensitivity and bandwidth must be taken care of. Therefore in this paper, the research would probe into the powerful technology known as the Kalman filter [2], which is also known as a mathematical power tool that can play an increasingly important role in speech processing.

A time series is a gathering of observations made sequentially through time; these measurements may be made continuously through time or be taken at a discrete set of time point. Therefore, the main objectives of time series analysis include the description, modeling, controlling and forecasting in this thesis primarily concerned with forecasting/ prediction and other often prerequisites.

Many methods associated with analysis of stationary time series were invented by statisticians, merchants, financiers, speculators and engineers. These methods include the harmonic analysis, first differences and logarithms of data, autocorrelation, auto regression, regression on functions of time, moving averages, and decomposition of time series. Models associated with these methods are the conventional ARIMA model, GARCH model, Spectral Analysis model, Hidden Markov Model, Wavelet model, State space model, Change point detection and independent component analysis.

MSc. Alaa K. Satti was with Telecommunication Department at University of Malaya - Malaysia and now with the Future University in Telecommunication and satellite engineering & Research and development Project – Sudan, 10553 (phone: 00249912675095 ; fax: 00249-183 ; e-mail: alakamal@gmail.com).

II. Problem Statement and Aims

A. Problem Statement

The speech signal is experiential in noisy environment using microphone with discrete time observation signal is represented as:

$$\mathbf{Y\,(n) = s\,(n) + v\,(n)} \tag{1}$$

where s (n) is clean speech sequence and v (n) is additive noise.

B. Objectives

The primary objective of this thesis was to come up with the reconstruction and enhancement of the speech signal, corrupted by additive background noise, with the aid of the Kalman filtering technique, using the MATLAB simulator. This technique would be used to model an autoregressive process (AR) represented in the state space domain by the Kalman filter.

C. Speech Enhancement

Enhancement means improvement and increase in quality of something, when we applied speech that means improvement of quality of degraded speech signal by using processing tool. The two main perceptual criteria for measuring the performance of a speech enhancement are *Quality* of the enhanced signal measures its clarity, distorted nature, and the level of residual noise in that signal. The quality is a subjective measure that is indicative of the extent to which the listener is comfortable with the enhanced signal.

The second criterion measures the *intelligibility* of the enhanced signal.

This is an objective measure which provides the percentage of words that could be correctly identified by listeners. [3]

Many filters used to remove noise **Wiener filter** is one of optimal filter proposed by Norbert wiener during 1940 which reduce the noise comes in signal through estimating of the noiseless signal based on statistical approach.

The input to the Wiener filter is assumed to be a signal *s* (*t*), corrupted by noise, *n* (*t*). The output *s^(t),* is calculated by means of a wiener filter g (*t*), using the following convolution:

$$s\hat{}(t)=g(t)*(s(t)+n(t)) \quad (2)$$

Working by different signals based on their frequency spectra mean it seems logical that the "mostly signal" frequencies should be passed through the filter, while the "mostly noise" frequencies should be blocked. The Wiener filter takes this idea a step further; the gain of the filter *at each frequency* is determined by the relative amount of signal and noise *at that frequency.*

D.Kalman Filter

Kalman filter is one of the most important optimal filters to estimate the state of dynamic system from series of noise measurement and resolve an error which is contained in speech, after passing it through a distorted channel. For this, the researcher found that the Kalman filter could be used to solve and suppress the very slow computations of the Wiener filter, which is also one of the common adaptive filtering techniques applied to speech.

Based on this finding, it was therefore concluded that the Kalman filtering possesses an accurate estimation and good constructing method for speech. It includes mainly two purposes: *estimation* and *performance analysis* of estimator.

Kalman filtering includes mainly two purposes: estimation and performance analysis of estimator. As it uses a complete description of the probability of its estimation errors the "design parameters" of the following estimation systems are

- Types of sensors to be used;
- Locations and orientations of the various sensor types with respect to the system to be estimated;
- Allowable noise characteristics of the sensors;
- Pre-filtering methods for smoothing sensor noise;
- Data sampling rates for the various sensor types and
- The level of model simplification for reducing implementation requirements.

III. Math

A. Computational Origins of the Filter

Consider $\overline{X}\kappa \in \Re$ with "super mines" to be our a priori state estimate at step *k* given knowledge of the process priori to step *k* and $\overline{X}\kappa \in \Re$ without super minus to be our a posterior state estimate at step *k* given measurement *Zκ* . Priori and posteriori estimate errors define as

$$\overline{\ell} = \chi_\kappa - {}^\wedge\chi_\kappa^- \text{ and } \ell = \chi_\kappa - {}^\wedge\chi_\kappa \quad (3)$$

A priori estimate error **covariance** is then

$$\mathrm{P}_\kappa^- = \mathrm{E}[\ell_\kappa^- \ell_\kappa^{-\mathrm{T}}] \quad (4)$$

and the a posteriori estimate error **covariance** is

$$\mathrm{P}_\kappa = \mathrm{E}[\ell_\kappa \ell_\kappa^{\mathrm{T}}] \quad (5)$$

After obtain equations for Kalman filter we have to finding an equation to compute an a posteriori state estimate $\hat{\chi}_\kappa$ as linear combination of an a priori $\overset{\barwedge}{\chi_K}$ and a weighted difference between an actual measurement Z_κ and a measurement prediction $\mathrm{H}\overset{\barwedge}{\chi_K}$ as shown in (6) below (Justification is rooted in the probability of a priori estimate $\overset{\barwedge}{\chi_K}$ from (Bayes' rule))

$$\hat{\chi}_\kappa = \overset{\barwedge}{\chi_K} + \mathrm{K}(Z_\kappa - \mathrm{H}\overset{\barwedge}{\chi_K}) \quad (6)$$

The difference $(Z_\kappa - \mathrm{H}\overset{\barwedge}{\chi_K})$ in (6) is called the *measurement innovation or residual* which reflect the difference between the measured value and the estimated value as well as the samples *cannot* be predicted exactly, when residual equal zero means that the two are in complete agreement.

In (6) ***K*** is gain factor with *nxm* matrix to minimize the a *posteriori error* covariance in (5) this can be accomplished by first substituting (6) into above definition for ℓ_κ, substituting that into (5).

IV. Comparative Advantages of Kalman Filter and Wiener Filter

This some advantages of Kalman filter comparing with another filter famous as wiener filter was appeared before Kalman filter in order is understood from [4]:

- The Kalman filter compatible with state space model formulation which was able to estimate and control system.
- Kalman requires less mathematical computation than wiener filter.
- Kalman filter provides necessary information for mathematically sound, statically based decision methods for detecting and rejecting irregular measurements.
- Kalman filter algorithm implement able on a digital computer which is slower compared with wiener; but it is capable of greater accuracy.
- Due to stationary properties of Kalman filter are not require deterministic dynamics or random processes.

V. Discrete Kalman Filter

A. Process to Estimate

Subsequent to the introduction of using Kalman filter, we have to know the formulation analysis and process to estimate the state at discrete points in time.

The minimize form for Kalman gain **K** is given by

$$\mathrm{K}_\kappa = (\mathrm{P}_\kappa^- \mathrm{H}^{\mathrm{T}} + \mathbb{R})^{-1} = \frac{\mathrm{P}_\kappa^- \mathrm{H}^{\mathrm{T}}}{\mathrm{H}\mathrm{P}_\kappa^- \mathrm{H}^{\mathrm{T}} + \mathbb{R}} \tag{7}$$

Looking to (7) we see that as the measurement error covariance R approaches zero, the gain **K** weights the residual more heavily. Specially

$$\lim_{\delta x \to 0} \mathrm{K}_\kappa = \mathrm{H}^{-1}$$

Similarly a *priori* estimate error covariance P_κ^- approaches zero, the gain ***K*** weights the residual less heavily. Specially,

$$\lim_{\delta x \to 0} \mathrm{K}_\kappa = 0$$

The process of Kalman filter begins with trying to estimate the state $\overline{X}\kappa \in \Re$ of discrete time managed by linear stochastic difference equation.

True State:

$$\chi_\kappa = \mathrm{A}\kappa_{\kappa-1} + \mathrm{B}\upsilon_\kappa + w_{\kappa-1} \tag{8}$$

With measurement $Z \in \Re^{\mathfrak{M}}$ that is

Observation Measurement:

$$\mathrm{Z}_\kappa = \mathrm{H}\chi_\kappa + \nu_\kappa \tag{9}$$

Wκ and Vκ represent the process and measurement noise respectively which independent of each other, white and with normal probability distributions.

$$\begin{aligned} \rho(\mathrm{w}) - \mathrm{N}(0,\mathrm{Q}) \\ \rho(\nu) - \mathrm{N}(0,R) \end{aligned} \tag{10}$$

Q is the process noise covariance and R is measurement noise covariance matrices change with each measurement here assume they are constant.

B. Discrete Kalman Filter Algorithm

The Kalman filter works by estimates the process using feedback control mean estimates the process state at some time then obtains feedback in the form of noisy measurements. Equations of Kalman filter divided into two groups*: time update* and *measurement update* equations. Responsibilities of time update equations are for projecting forward (in time) the current state and error covariance estimates to obtain the priori estimates for the next time step as (predictor equations) and measurement update equations are responsible for the feedback i.e. integrating a new measurement into the *a priori* estimate to get an enhanced *a posteriori* estimate as (corrector equations).

Certainly the final estimation algorithm to solving numerical problems as shown in Fig. 1 below is known as *predictor-corrector* algorithm.

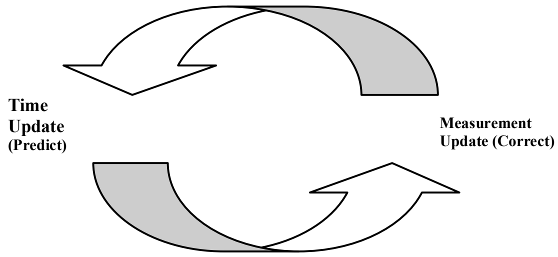

Fig. 1 The continuing discrete Kalman filter cycle

The *time update* projects the current state estimate ahead in time. The *measurement update* adjusts the projected estimate from *time update* by an actual Measurement at certain time. The particular equations for the time and update are obtainable in tables below

TABLE I
DISCRETE KALMAN FILTER TIME UPDATE EQUATIONS

Symbol	Quantity	Time update Equations
$\hat{\chi}_\kappa^-$	Priori state estimate at step k	$\hat{\chi}_\kappa^- = \mathrm{A}\hat{\chi}_{\kappa-1} + \mathrm{B}\upsilon_\kappa$ (11)
P_κ^-	Priori estimate error covariance	$\mathrm{P}_\kappa^- = \mathrm{A}\mathrm{P}_{\kappa-1}\mathrm{A}^{\mathrm{T}} + Q$ (12)

A, B are from (1) while *Q* is from (10), initial condition already discussed.

TABLE II
DISCRETE KALMAN FILTER MEASUREMENT UPDATE EQUATIONS

Symbol	Quantity	Measurement update Equations
K_κ	Kalman Gain	$\mathrm{K}_\kappa = \mathrm{P}_\kappa^{-\mathrm{T}}(\mathrm{H}\mathrm{P}_\kappa^- \mathrm{H}^{\mathrm{T}} + \mathrm{R})^{-1}$ (13)
$\hat{\chi}_\kappa$	Posteriori state estimate	$\hat{\chi}_\kappa = \hat{\chi}_\kappa^- + \mathrm{K}_\kappa(\mathrm{Z}_\kappa - \mathrm{H}_\kappa \hat{\chi}_\kappa^-)$ (14)
P_κ	Posteriori error covariance	$\mathrm{P}_\kappa = (\mathrm{I} - \mathrm{K}_\kappa \mathrm{H}_\kappa)\mathrm{P}_\kappa^-$ (15)

Observable that the first step in measurement updates equations is to compute Kalman gain K_κ, next step is to measure the process to obtain Z_κ and then to generate an a posteriori state estimate in (14). The final step is to obtain an a posteriori error covariance estimate via (15).

The process repeated with the previous a *posteriori* estimates used to project or predicts the new a *priori* estimates as recursive solution one of the attractive nature of Kalman filter as in Fig. 2 present below a complete operation of the filter combining Tables I & II with Fig. 1.

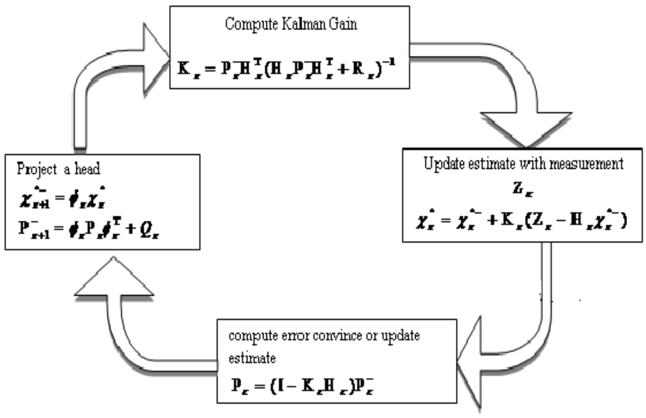

Fig. 2 Complete loop operation of Kalman filter

C. Parameters Estimation or Changing

In real execution of the filter, *R* the measurement noise covariance is usually measured prior to operation of the filter. It is possible to measuring R practical because we need to measure the process of filter however we should be able to take some off-line sample from measurements.

Find out of *Q* process noise covariance is generally more difficult as we don't have ability to directly observe the process we are estimating.

Moreover if we don't have rational base for choosing the parameters so the times greater filter performance can be found by tuning the filter parameter R, *Q* off-line time.

When *Q*, R are in fact constant then both the estimation error covariance P_κ and the Kalman gain K_κ will remain constant so these parameters can be recomputed by running filter offline or in steady state value. [5]

D. Purpose of Kalman Filter

Proposed Kalman filtering method which allows for stationary and non stationary of speech and at same time it is capable to estimating errors accurate more than another filters.

Kalman Filter is an adaptive least square error filter that provides an efficient computational recursive solution for estimating a signal in presence of Gaussian noises to continuously update the best estimate of the system's current state. Extra than 20 years have elapsed kalman's original paper is still dealing with many newspapers, applications and variation based on the original of Kalman filter. It has withstood the test of time!

VI. Implementation of Kalman Filter

In the stage of estimation, the values of some unknown parameters are estimated previous to the operation of the filter. For instance, the variance of measurement noise can be estimated from some off-line sample measurements. The variance of the process noise is more heuristic and tuning can be done to change its initial value to get better accuracy of model. The values of other parameters are assumed to be known either from the laws of physics or from other logical assumption.

Once the values of the parameters are identified, the accuracy of the estimated model will be tested in the model justification stage. A series of state predictions are estimated through the time update and measurement update equations (loop operation of Kalman filter). The precision of these predictions is compared next to the actual observations and residuals are calculated. If the values of the residuals are too big (unacceptable), fine tuning is done by changing the values of the parameters R, Q (measurement and process noise respectively) of the model until an acceptable residual level is achieved. Assuming *A*, *H*, R and Q are static, a Mat lab and R software are written to find the best values for these parameters that would produce the estimated $X\kappa$ (produced by those parameters) are compared to the actual time series.

The range for A and H is restricted between 0 and 2 since these two scalars that should carry positive correlation. The increment for the values of A and H is 0.1 for each step. The variance of the time series data can be calculated off-line and the range for R and Q is restricted between 0 and the variance of the data. The values of R and Q are incremented by 1 in each step. The initial values chosen for $\hat{x}^-_k$ and P^-_k are usual 0 and 1 respectively.

The tasks at hand is to find the values of these parameters and determine the best estimate of state of speech mathematically, plus find the estimator with the smallest possible error variance.

By and large of that all pole speech models can be considered as filtered excitation noise sequence and formally represented as

$$y_\kappa = \left[\frac{1}{1 - \sum_{i=1}^{N} a_i z^{-i}}\right] w_\kappa$$

where $y\kappa$ represents the input speech samples, $w\kappa$ represents the excitation white noise sequence, a_i are the filter coefficients, K number of iterations and N is the order of the model.[6]

VII. Round-Off Error

A **round-off error**, also called **rounding error**, is the dissimilarity between the calculated approximation of a number and its exact mathematical value.

Computer round off limits the accuracy of numerical computation of Kalman filter implementation which cause degradation in the performance of filter, alternative method of execution to Kalman filter equations like Riccati equations, mostly used to robustness against round-off errors.[7]

There are a number of ways in which the sensitivity due to round off errors can be reduced; care must be used to choosing of state variables of a Kalman filter more details about methods in [8].

VIII. Simulated Result

In striving to achieve the aim of this paper the simulation source code written in MATLAB 6 to employ Kalman filter in speech recorded and S software to get the parameters need to complete the simulation. Using MATLAB for the reason that is very influential programming language for matrices while the signal noise is simply an array of numbers; this is very practical approach.

The simulated results are obtained by estimating coefficients in 4th order of Kalman filtering. The reconstructed speech signal Y k followed by an input signal with random generated noise show as

$$Y_\kappa = a_1 Y_{\kappa-1} + a_2 Y_{\kappa-2} + \ldots + a_N Y_{\kappa-N} + w_\kappa$$

Since the value of Y_κ is the input of Kalman filter process which has accurate estimation capability. In order to reconstruct the signal the parameters w_κ and coefficients a_i are assumed. Then substitute a_i by setting Kalman filter in 4 coefficients [9] as

$$Y_\kappa = (0.9)^* Y_{\kappa-1} + (-0.1)^* Y_{\kappa-2} + (0.4)^* Y_{\kappa-3} + (-0.5)^* Y_{\kappa-4} + W\kappa$$

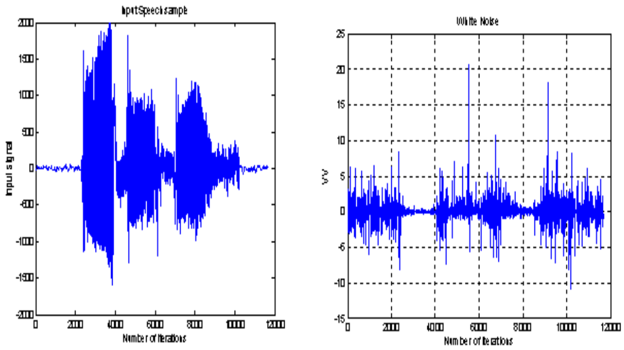

Fig. 3 Input speech sample with estimated white noise respectively

A. Different Coefficients at Different Iterations

In case of 4th order Kalman filter will have 4 groups of coefficients, however each set of coefficients have different values at different iterations. As the order of Kalman filter increases to create accurate coefficients estimates will guide to more duration.

Different iterations found that different coefficients producing show in Fig. 4 include a, b, c and d (green, red, blue, and pink respectively) which suggest that results of Kalman filter will employ accurate when applied to speech samples.

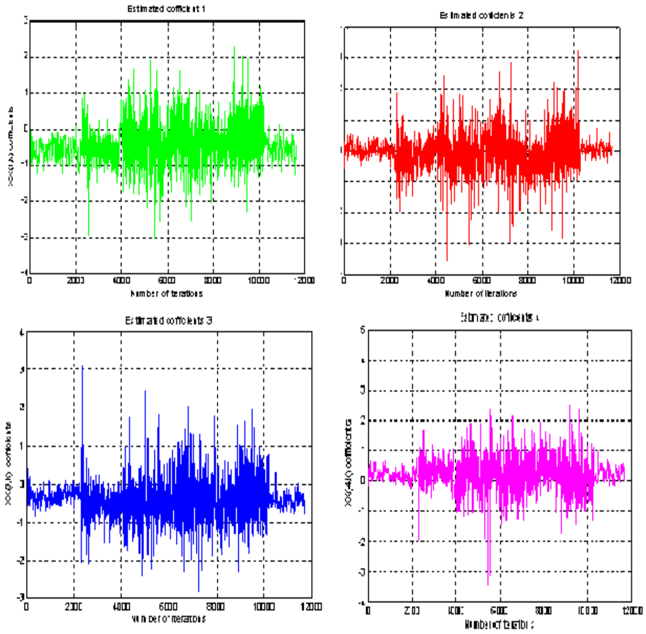

Fig 4 Kalman Coefficient (1st, 2nd, 3rd and 4th) values in different iteration

The 4-cofficients set estimated from figures above are shown in Table III below in different iterations:

TABLE III
NUMBER OF COEFFICIENT WITH DIFFERENT ITERATION

Coefficients Iterations	1st Set	2nd Set	3rd Set	4th Set
1 to 2000	-0.5	0.4	-0.7	0.7
2001 to 4000	-0.7	0.2	-0.3	0.8
4001 to 6000	-0.3	0.4	-0.8	0.5
6000 to 8000	-0.8	0.7	-0.7	0.8
8000 to 10000	-0.6	0.8	-0.6	0.3
10001 to 12000	-0.3	0.6	-0.5	0.4

B. Ruling of Comparison

A phenomenon observed in Kalman filter when employing in estimating the result in figures before, is that Kalman needs some time in order to become stable and estimate the reconstructed output of speech. And Q and R have to be tuned to meet the identical of input signal.

Two tools to realize the similarity and the minim errors between the simulated signals and true signal:

1. **Least square error:** The *least square is* one that has a smaller error or residuals than any other straight-line model. *Residuals* are the difference between the actual value of the dependent variable and the value predicted by the regression model. And *Regression model* is for predict variables form others variables.
2. **Cross correlation:** Is professional tool to match and measure the similarity of two waveforms based upon the amount of common parts for prediction and compare the differences between the two signals.[10]

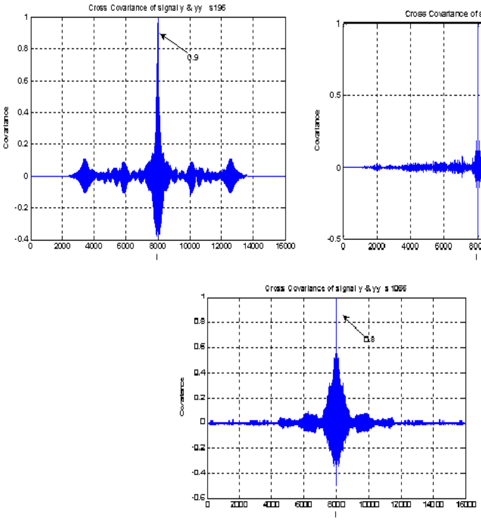

Fig. 5 Cross correlation of sample s196, s306 and s1066

IX. Conclusion

Till the day date, the reduction of noise mainly in telecommunication applications has attracted amount of research attention however Kalman filter has been found to be the title of many researchers in different applications, mainly in the area of navigation and GPS, because of its high accuracy in estimating the position of objects. As mentioned previously Kalman filter is used to implement to reconstruct and enhancement the speech signal developed by Mat lab software.

The results shown had been obtained by tuned or testing signal in different orders of Kalman and iterations to provide optimal performance. By means of implement five Kalman filter's equations which called complete loop operation of Kalman *(predict* ↔ *correct)* and tuning parameters R and Q to meet the aim of paper. However R parameter is of superfluous to be changed whereas Q has to be tuned.

Furthermore, tests by least square and cross correlation had been performed during simulation for measuring the similarity and not identically between input and output speech signals.

On the whole, this thesis has been doing well to achieve and solve the problem statement. But the difficulties take place when trying and iterating to estimate the coefficients. For the most significant part is the time management builds up during practical the research.

Acknowledgment

Alaa K. Satti thanks Future University- Sudan for the moral and financial support with full network connection and a well equipped library provided which made work much easier. Also Dr. *Hamza Aroof* the supervisor in University of Malaya Kuala Lumpur for his wide knowledge, understanding, encouraging and personal guidance.

References

[1] R.V. Cox, "Current Methods of Speech Coding", *International Journal of High Speed Electronics and Systems*, Vol. 8, No. 1, World Scientific Publishing Company, pp-13-68.

[2] S. Sridharan, J. Leis and K. K. Paliwal "*SPEECH CODING*" Handbook of Neural Networks for Speech Processing (First edition) pp. 3-5 Ed. S. Katagiri, 1999 Artech House.

[3] Philipo C. loizou Speech Enhancement Series: *Signals and Communication Technology* Benesty, Jacob; Makino, Shoji; Chen, Jingdong (Eds.) Springer-Verlag Berlin Heidelberg 2005, Printed in The Netherlands.

[4] M.S. Grewal and A.P. Andrews, *Kalman Filtering Theory and Practice Using MATLAB 2nd edition*, John Wiley & Sons, Canada, 2001

[5] M. S. Grewal and A.P. Andrews, *Kalman Filtering Theory and Practice*, Upper Saddle River, NJ USA, Prentice Hall, 1993

[6] S. Crisafulli, J.D. Mills, and R.R Bitmead, "Kalman Filtering Techniques in Speech Coding". *In Proc. IEEE International Conference on Acoustics, Speech, and Signal* Processing, San Francisco, March 1992

[7] Global Positioning Systems, Inertial Navigation, and Integration, By Mohinder S. Grewal, Lawrence R. Weill, Angus P. Andrews, Published by Wiley-IEEE, 2001

[8] Filtering for Stochastic Processes with Applications to Guidance, By Richard S. Bucy, Peter D. Joseph, Published by AMS Bookstore, 1987]

[9] Shumway, Robert H., Stoffer**,** David S., *Time Series Analysis and Its Applications with R Examples*, second edtition, Springer Texts in Statistics Series, 2006

[10] B. P. Lathi, "Signal Comparison: Correlation", *Signal, Systems and Communication*, John Wiley & Sons, USA, 1965.

Alaa K. Satti born in Khartoum – Sudan. got her Bachelor degree in Computer and communication Engineering with first class degree from Future University (Formally Computer Man) – Sudan 2005, and Master Degree in Telecommunication Engineering from University of Malaya with first class degree Malaysia 2009, Senior Lecturer and Project Unit Manager in Future University – Khartoum Sudan; previous positions Lecturer and demonstrator in university of Malaya , Malaysia Kulala lumpur . Became a Member (M) of **IEEE** in 2012.

Finally, Ms. Satti won awards of best graduate project in telecommunication engineering and participates in many committees and publications.