Comparison Of LDM and LMS for an Application of a Speech

V.A.Mane vikram_mane34@yahoo.com
Department of E&TC
Annasaheb Dange COE
Ashta, 416302, India

K.P.Paradeshi kiranparadeshi@redifmail.com
Department of E&TC
Annasaheb Dange COE
Ashta, 416302, India

S.A.Harage shrikant.harage@rediffmail.com
Department of E&TC
Annasaheb Dange COE
Ashta, 416302, India

M.S.Ingavale mangesh_ingavale1982@rediffmail.com
Department of E&TC
Annasaheb Dange COE
Ashta, 416302, India

Abstract

Automatic speech recognition (ASR) has moved from science-fiction fantasy to daily reality for citizens of technological societies. Some people seek it out, preferring dictating to typing, or benefiting from voice control of aids such as wheel-chairs. Others find it embedded in their Hitech gadgetry – in mobile phones and car navigation systems, or cropping up in what would have until recently been human roles such as telephone booking of cinema tickets. Wherever you may meet it, computer speech recognition is here, and it's here to stay.

Most of the automatic speech recognition (ASR) systems are based on Gaussian Mixtures model. The output of these models depends on subphone states. We often measure and transform the speech signal in another form to enhance our ability to communicate. Speech recognition is the conversion from acoustic waveform into written equivalent message information. The nature of speech recognition problem is heavily dependent upon the constraints placed on the speaker, speaking situation and message context. Various speech recognition systems are available. The system which detects the hidden conditions of speech is the best model. LMS is one of the simple algorithm used to reconstruct the speech and linear dynamic model is also used to recognize the speech in noisy atmosphere..This paper is analysis and comparison between the LDM and a simple LMS algorithm which can be used for speech recognition purpose.

Keywords : White Noise, Error Covariance Matrix, kalman Gain, LMS Cross Correlation

1. INTRODUCTION

Speech is a form of communication in everyday life. It existed since human civilizations began and even till now, speech is applied to high technological telecommunication systems. A particular field, which I personally feel, will excel be speech signal processing in the world of telecommunications. As applications like Cellular and satellite technology are getting popular among mankind, human beings tend to demand more advance technology and are in search of improved applications. For this reason, researchers are looking closely into the four generic attributes of speech coding. They are complexity, quality, bit rate and delay. Other issues like robustness to transmission errors, multistage encoding/decoding, and accommodation of non-voice signals such as in-band signaling and voice band modem data play an important role in coding of speech as well.

Presently Speech processing has been a growing and dynamic field for more than two decades and there is every indication that this growth will continue and even accelerate. During this growth there has been a close relationship between the developments of new algorithms and theoretical results, new filtering techniques are also of consideration to the success of speech processing.
A least mean square (LMS) adaptive filtering approach has been formulated for removing the deleterious effects of additive noise on the speech signal; unlike the classical LMS adaptive filtering scheme, the proposed method is designed to cancel out the clean true speech signal. This method takes advantage of the quasi-periodic nature of the speech signal to form an estimate of the clean speech signal at time t from the value of the signal at time t minus the estimated pitch period. For additive white noise distortion, preliminary tests indicate that the method improves the perceived speech

One of the common adaptive filtering techniques that are applied to speech is the Wiener filter. This filter is capable of estimating errors however at only very slow computations. On the other hand, the Kalman filter suppresses this disadvantage. As widely known to the world, Kalman filtering techniques are used on GPS (Global Positioning System) and INS (Inertial Navigation System). Nonetheless, they are not widely used for speech signal coding applications. According to, the reason why Kalman filter is so popular in the field of radar tracking and navigating system is that it is an optimal estimator, which provides very accurate estimation of the position of either airborne objects or shipping vessels. Due to its accurate estimation characteristic, electrical engineers are picturing the Kalman filter as a design tool for speech, whereby it can estimate and resolve errors that are contained in speech after passing through a distorted channel. Due to this motivating fact, there are many ways a Kalman filter can be tuned to suit engineering applications such as network telephony and even satellite phone conferencing. Knowing the fact that preserving information, which is contained in speech, is of extreme importance, the availability of signal filters such as the Kalman filter is of great importance.

2. EARLY APPROACHES TO SPEECH RECOGNITION

Automatic speech recognition might appear to be an almost unattainable goal. However, by concentrating on a reduced specification and by tracking the problems in a scientific and staged manner, it has been possible to make considerable progress in understanding the precise nature of the problems and in development of relevant and practical solutions [12]. However, this has not always been the case. Some of the early work, which interesting in the context of a review of different approaches to automatic speech recognition, tended to be either overambitious about the achievements that could realistically be expected to be realized or somewhat naive with regard to the real difficulties that were being tackled.

Early attempts can thus be categorized into one of two main approaches. In the fifties and sixties the main approach was based on simple principles of ‘pattern matching’ that in the seventies gave way to a ‘knowledge engineering ‘or rule based approach. Only towards the end of the seventies there was a growing awareness of the need to integrate these two approaches and move towards a clear and scientific recognition- a move that ultimately led to a maturation of ides and algorithms, which are now beginning to provide powerful exploitable solutions [10].
The following section reviews some of these early approaches to automatic speech recognition.

2.1 Pattern matching

Such systems employ two modes of operation: a ‘training mode’ in which example speech patterns (usually words) are stored as reference ‘templates’ and a recognition mode in which incoming speech patterns are compared with each reference pattern that is most similar to the input pattern determines the result. In this scheme the acoustic pattern of a speech signal typically consisted of a sequence of vectors, which had been derived from the speech waveform using some form of ‘preprocessing’. For example it was common to perform a frequency analysis by means of an FFT or a filter bank in order to produce vectors that correspond to the short-time power spectrum of the signal into discrete pattern segments.

The key to success of this approach is the comparison process, and a technique called 'linear time normalization' was commonly used in order to overcome variability in the duration of spoken words. In this situation, the lengths of the patterns were 'time normalized' to a standard duration by lengthening (or shortening) the patterns the appropriate amount by using a fixed expansion (or compression) of the time scale uniformly over the entire pattern

2.2 Knowledge engineering

The knowledge-based approach popular in the early seventies was based on techniques from the field of artificial intelligence, which was the newly emerging. These techniques were applied to traditional concepts from the disciplines of phonetic and linguistics about how speech signals was organizing. The key principle was to exploit the speech knowledge through its exploit use within a rule-based framework aimed at deriving and interpretation that would be suitable for the purpose of understanding the semantic content of the signal.

2.3 Integrated approach:

This new process (popular in the late seventies) became known as 'dynamic time warping'(DTW) and it has been a highly successful technique in terms of raising performance to a level at which serious commercialization of automatic speech recognition systems could begin.

3. LMS ALGORITHM

A linear mean square (LMS) adaptive filtering approach has been formulated for removing the deleterious effects of additive noise on the speech signal; unlike the classical LMS adaptive filtering scheme, the proposed method is designed to cancel out the clean true speech signal [11]. . An adaptive LMS filter was employed to process speech in signal-to-noise ratios (S/N) varying from -8 to +12 dB. The filter configuration is commonly called noise cancellation [12][14][15].

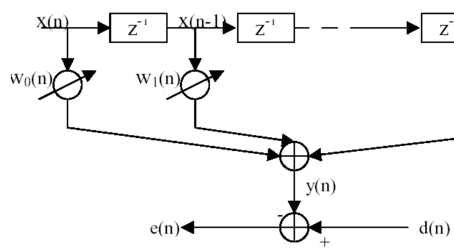

FIGURE 1: LMS model

$$Y(n) = \sum_{i=0}^{N-1} W_i(n)x(n-i)$$

e(n)=d(n)-y(n)

We assume that the signals involved are real-valued.

The LMS algorithm changes (adapts) the filter tap weights so that e(n) is minimized in the mean-square sense. When the processes x(n) & d(n) are jointly stationary, this algorithm converges to a set of tap-weights which, on average, are equal to the Wiener-Hopf solution.

The LMS algorithm is a practical scheme for realizing Wiener filters, without explicitly solving the Wiener-Hopf equation.

The conventional LMS algorithm is a stochastic implementation of the steepest descent algorithm. It simply replaces the cost function

ζ =E [e2(n)] Substituting ζ =e2(n)

For in the steepest descent recursion, we obtain

$$\overline{W}.(n+1) = \overline{W}(n) - \mu\nabla e^2(n)$$

.....W N-1 (n)

$$\left[\overline{\nabla} = \frac{\partial}{\partial w0} \quad \frac{\partial}{\partial w1} \cdots\cdots \frac{\partial}{\partial w\ \ n-1}\right]$$

Note that the i-th element of the gradient vector is

$$\frac{\partial e^2(n)}{\partial wi} = 2e(n)\frac{\partial e^2(n)}{\partial wi}$$
$$= -2e(n)\frac{\partial y(n)}{\partial wi}$$

$$= -2e(n)x(n-i)$$
Then

Where (n) =[
Finally we obtain-
+2μe
Equation is referred to as the LMS recursion.
Summary of the LMS algorithm,
Input:
Tap -weight vector:
Input vector:
Desired output: d
Output:
Filter output: y
Tap -weight vector update:
1. Filtering: y
2. Error estimation:
3. Tap-weight vector adaption: +2μe

4. LINEAR DYNAMIC MODEL (KALMAN FILTER)

In 1960, R.E. Kalman published his famous paper describing a recursive solution to the discrete-data linear filtering problems [10]. Since that time, the Kalman filter has been the subject of extensive research and application, particularly in the area of autonomous or assisted navigation. It is only a tool – It aids mankind in solving problems, however, it does not solve any problem all by itself. This is however not a physical tool, but a mathematical one, which is made from mathematical models. In short, essentially tools for the mind. They help mental work become more efficient, just like mechanical tools, which make physical work less tedious. Additionally, it is important to understand its use and function before one can apply it effectively .It is a computer program - It uses a finite representation of the estimation problem, which is a finite number of variables; therefore this is the reason why it

is said to be "ideally suited to digital computer implementation". However, assuming that these variables are real numbers with infinite precision, some problems do happen. This is due from the distinction between finite dimension and finite information, and the distinction between "finite" and "manageable" problem sizes. On the practical side when using Kalman filtering, the above issues must be considered according to references [1; 2; 3; 4].
Mathematical analysis of Kalman filter

Following discussions from references [4; 5; 6; 7; 10]
After going through some of the introduction and advantages of using Kalman filter, we will now take a look at the process of this magnificent filter. The process commences with the addresses of a general problem of trying to estimate the state $x \in \Re^n$ of a discrete-time controlled process that is governed by a linear stochastic difference equation:
xk=Axk-1+Buk+wk-1 (1.1)
With measurement
zk=Hxk+vk (1.2)
The random variables wk and vk represent the process and measurement noise respectively. We assume that they are independent of each other, and with normal probability distributions
p(w)~N(0,Q) (1.3)
p(v)~N(0,R) (1.4)

.
Ideally, the process noise covariance Q and measurement noise covariance R matrices are assumed to be constant, however in practice, they might change with each time step or measurement.

In the absence of either a driving function or process noise, the matrix nxn A in the difference equation (1.1) relates the state at the previous time step to the state at k-1 to the current step k. In practice, A might change with each time step, however here it is assumed constant. The nxl matrix B relates the optional control input to the state x., which is a matrix in the measurement equation (1.2), which relates the state to the measurement**.** In practice x might change with each time step or measurement, however we assume it is constant.

4.1 Discrete Kalman Filter
This section will begin with a broad overview, covering the "high-level" operation of one form of the discrete Kalman filter. After presenting this high-level view, I will narrow the focus to the specific equations and their use in this discrete version of the filter. How does the Kalman filter works? Firstly, it estimates a process by using a form of feedback control loop whereby the filter estimates the process state at some time and then obtains feedback in the form of (noisy) measurements. As such, these equations for the Kalman filter fall into two groups: “Time Update equations” and “Measurement Update equations”. The responsibilities of the 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. The measurement update equations are responsible for the feedback i.e. for incorporating a new measurement into the priori estimate to obtain an improved posteriori estimate. The time update equations can also be thought of as “predictor” equations, while the measurement update equations can be thought of as “corrector” equations. By and large, this loop process of the final estimation algorithm resembles that of a predictor-corrector algorithm for solving numerical problems just like the one shown in fig below

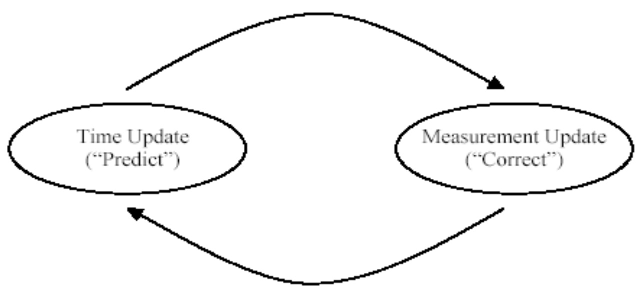

FIGURE 2: Discrete Kalman Filter Cycle

As the time update projects the current state estimate ahead in time, the measurement update adjusts the projected estimate from the time update by an actual measurement at that particular time. The specific equations for the “time” and “measurement” updates are presented below in Table1

$Xk=Axk-1+Buk$ 1.5
$Pk=Apk-1AT+Q$ 1.6

Table 1: Time update equations

$Kk=PkHT(HPkHT+R)-1$ 1.7
$Xk=xk+Kk(zk-Hxk)$ 1.8
$Pk=(I-KkH)Pk$ 1.9

TABLE 2: Measurement equations

Once again, notice how the time update equations in Table.1 project its state, x the filter are discussed in the earlier section. By referring to Table 1, it is obvious that the first task during the measurement update is to compute the Kalman gain, By (1.7) in the table above is to actually measure the process in order to obtain, and then to generate a posteriori state estimate, by incorporating the measurement as in (1.8). Once again, notice the repeated equation of (1.7) here and (1.8) for completeness. Finally, the last step is to obtain a posteriori error covariance estimate via (1.9). Thus, after each time and measurement update pair, this loop process is repeated to project or predict the new time step priori estimates using the previous time step posteriori estimates. This recursive nature is one of the very appealing features of the Kalman filter it makes practical implementations much more feasible than (for example) an implementation of a Wiener filter which is designed to operate on all of the data directly for each estimate. Instead, the Kalman filter recursively conditions the current estimate on all of the past measurements. The high-level diagram of Fig 2 is combined with the equations from Table .1 and Table 2, in Fig.2 as shown below, which offers a much more complete and clear picture of the operation of the Kalman Filter.

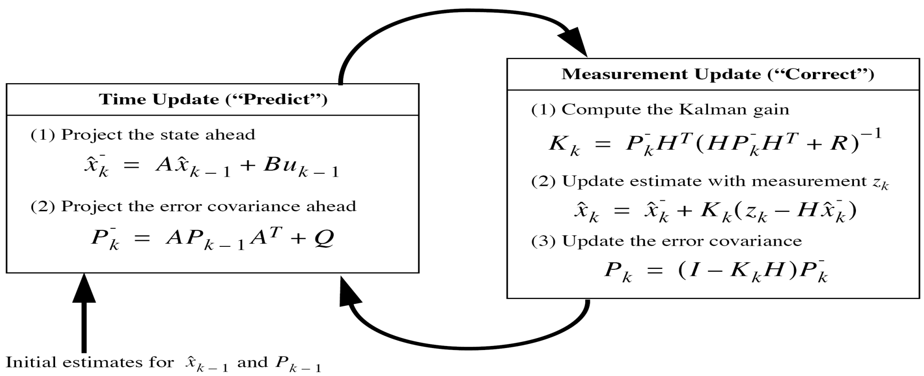

FIGURE 1: Complete picture of Kalman filter

The measurement noise covariance is usually measured before the operation of the filter when it comes to the actual implementation of Kalman filter. Generally, measuring the measurement noise covariance is practically possible due to the fact that the necessary requirement to measure the process noise covariance (while operating the filter), therefore it should be possible to take some off-line sample measurements in order to determine the variance of the measurement noise. As for determining of the process noise covariance, it will be generally more difficult. This is due to the reason that the process to be estimated is unable to be directly observed. Sometimes a relatively simple (poor) process model can produce acceptable results if one "injects" enough uncertainty into the process via the selection of. (Certainly, one would hope that the process measurements are reliable**).** In either case, whether or not a rational basis is chosen for the parameters, superior filter performance (statistically speaking) can be obtained by tuning the filter parameters R and Q. closing under conditions where R and Q are in fact constant, both the estimation error covariance and the Kalman gain will stabilize quickly and then remain constant (see the filter update equations in Fig 2). If this is the case, these parameters can be pre-computed by either running the filter off-line, or for example by determining the steady-state value.

5. COMPARISON OF LDM AND LMS

Optimal adjustment parameters of the adaptive filter with LMS algorithm in the practical application of suppression of additive noise in a speech signal for voice communication with the control system. By the proposed method, the optimal values of parameters of adaptive filter are calculated with guarantees the stability and convergence of the LMS algorithm [9] same as that of the LDM[16].
The proposed methods of recognition of speech give the following results on three different speeches
S1 is a noiseless speech sample and S1white is the S1 speech captured by white noise. The original speech sample, S1white and reconstructed speech samples with LDM and LMS are shown below. The cross covariance between the reconstructed and noiseless speech samples is also shown below.

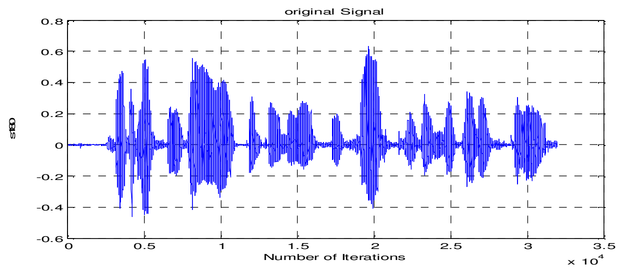

FIGURE 1a. S1 original speech sample

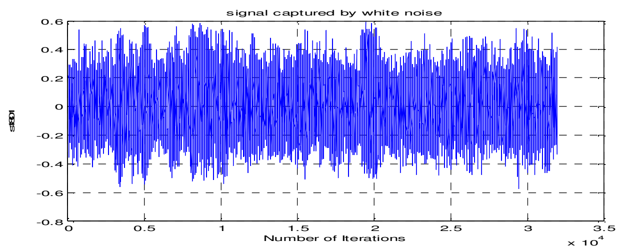

Fig 1b. S1 white speech sample captured by white noise

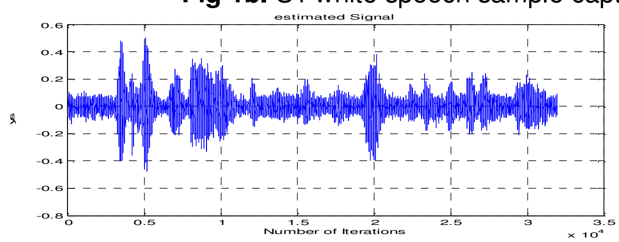

Fig 1c. Reconstructed speech by LDM

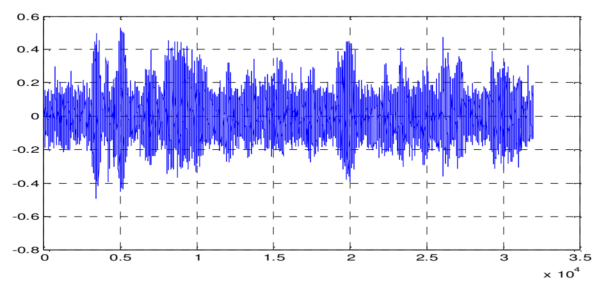

Fig 1d. Reconstructed speech by LMS

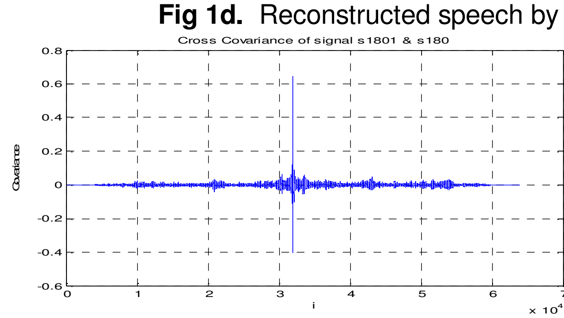

Fig 1e. Cross correlation between original and reconstructed signal by LDM

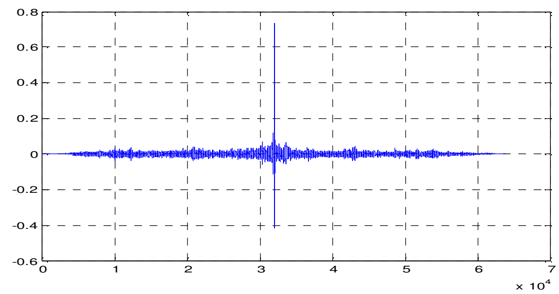

Fig 1f. Cross correlation between original and reconstructed signal by LMS

S2 is a noiseless speech sample and S2street is the S2 speech captured by street noise. The original speech sample S2, S2street and reconstructed speech samples with LDM and LMS are shown below. The cross covariance between the reconstructed and noiseless speech samples is also shown below.

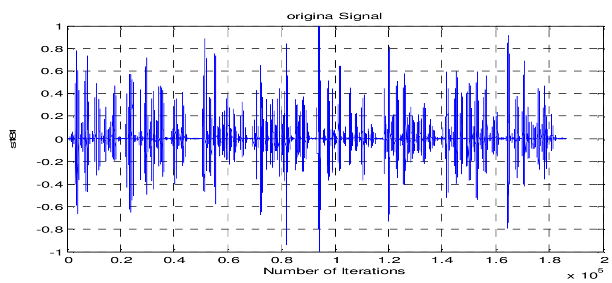

Fig 2a. S2 original speech sample

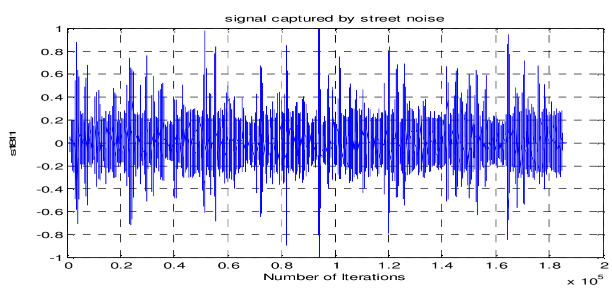

Fig 2b. S2street speech sample captured by street noise

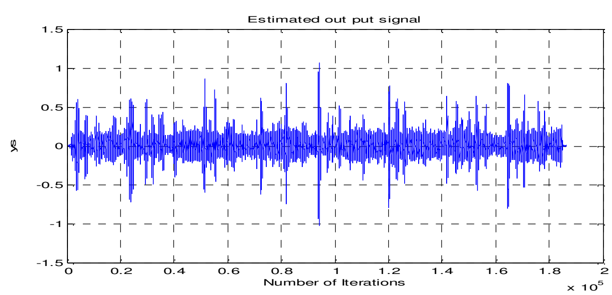

Fig 2c. Reconstructed speech by LDM

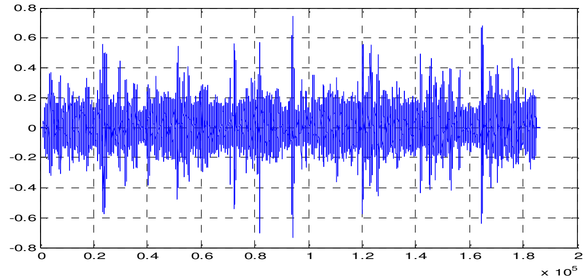

Fig 2d. Reconstructed speech by LMS

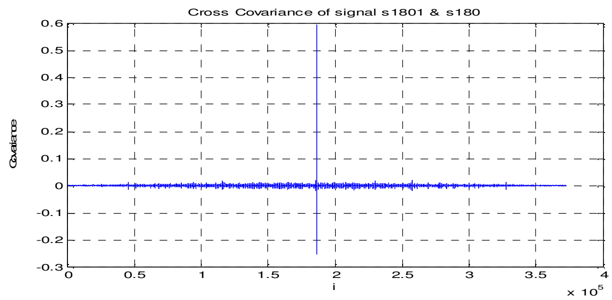

Fig 2e. Cross correlation between original and reconstructed signal by LDM

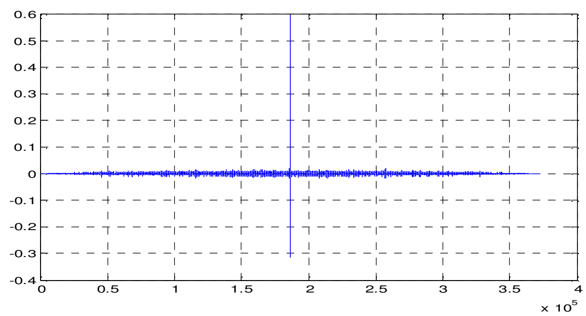

Fig 2f. Cross correlation between original and reconstructed signal by LMS

S3 is a noiseless speech sample and S3 and is the S3 speech captured by random noise (artificially generated). The original speech sample S3 and reconstructed speech samples with LDM and LMS are shown below. The cross covariance between the reconstructed and noiseless speech samples is also shown below.

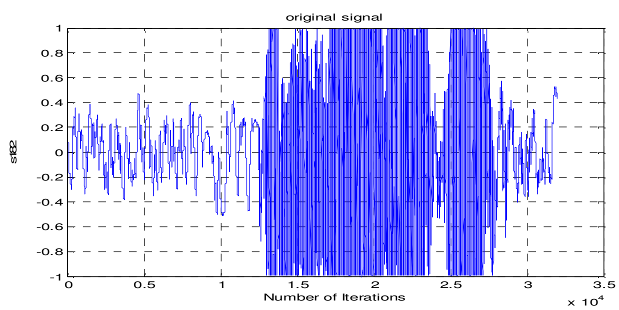

Fig 3a. S3 original speech sample

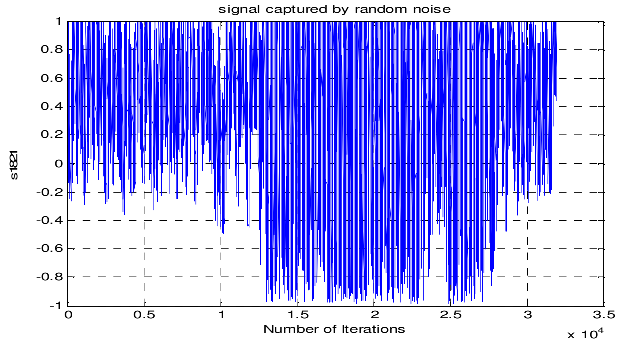

Fig 3b. S3street speech sample captured by random noise

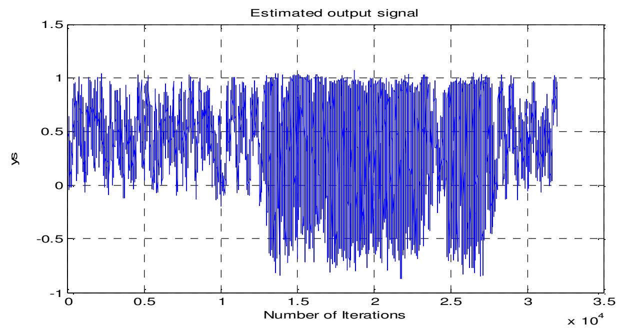

Fig 3c. Reconstructed speech by LDM

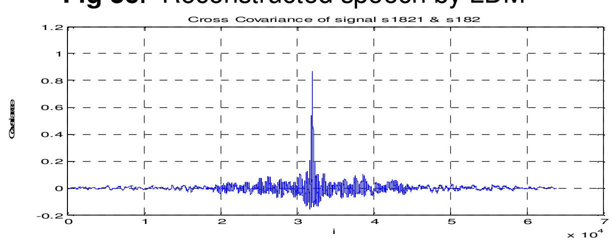

Fig 3 Cross correlation between original and reconstructed signal by LDM

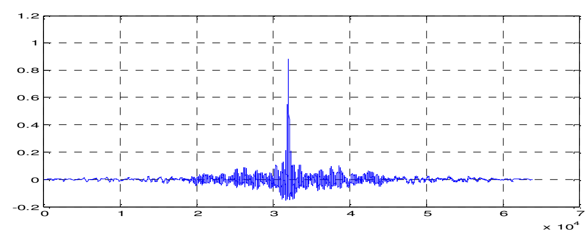

Fig 3e. Cross correlation between original and reconstructed signal by LMS

Speech	LDM	LMS
S1	0.65	0.736
S2	0.595	0.5996
S3	0.870	0.8836

TABLE 3: comparison of LDM and LMS in terms of cross correlation

6. CONCLUSION

After the working over the two models it is concluded that we can use the LDM for speech but LMS algorithm is also one of the methods which is simple and efficient algorithm. After comparing the results we found that for Speech sample S1 captured by White noise LMS algorithm is having better results than LDM. For Speech sample S2 captured by Street noise both algorithms are having approximately same results. For Speech sample S3 captured by Random noise again LMS algorithm is having better results than LDM. Overall we found LMS Algorithm is giving promising results for the above speech samples considered for experimentation.

7. REFERENCES

Book Chapters

[1] M.S. Grewal and A.P. Andrews, Kalman Filtering Theory and Practice Using MATLAB 2nd edition, John Wiley & Sons, Canada, 2001, pp 1-12

Book

[2] An Introduction to the Kalman FilterGreg Welch and Gary Bishop UNC-Chapel Hill, TR 95-041, July 24, 2006

Dissertations and Theses

[3] J. Frankel,"Linear Dynamic Models for automatic speech recognition", Ph.D. dissertation, The center for Speech Technology Research, University of Edinburgh, UK, 2003.

Article in a Journal

[4] Joe Frankel and Simon King, speech Recognition using Linear Dynamic Models", Member, IEEE Member,IEE Manuscript received September 2004. This work is supported by EPSRC grant GR/S21281/01 Joe Frankel and Simon King are both with the Center for speech Technology Research, University of Edinburgh.

[5] Discriminative training for large vocabulary speech recognition using minimum classification errors" by Eric McDermott member IEEE,TimonthyJ Hazen,Member IEEE, Jonathan Le Roux,Atsushi Nakamura,Member,IEEE,and Shigeru Katagiri,Fellow,IEEE IEEE transaction on Audio speech and language processing vol 15 No1 2007

[6] Kalman, R. E. 1960. "A New Approach to Linear Filtering and Prediction Problems," Transaction of the ASME--Journal of Basic Engineering, March 1960.

[7] Kalman-Filtering Speech Enhancement Method Based on a Voiced-Unvoiced Speech Model, Zenton Goh, Kah-Chye Tan, Senior Member, IEEE, and B. T. G. Tan IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, VOL. 7, NO. 5, SEPTEMBER 1999

[8] The stability of variable step-size LMS algorithms Gelfand,S.B.; Yongbin Wei; Krogmeier, J.V.; Sch. of Electr. & Comput. Eng., Purdue Univ., West Lafayette, IN Signal Processing, IEEE Transactions on Issue Date Dec 1999.

[9] Application of optimal settings of the LMS adaptive filter for speech signal processing Computer Science and Information Technology (IMCSIT), Proceedings of the 2010.

[10] C. R. Watkins, "Practical Kalman Filtering in Signal Coding", New Techniques in Signal Coding, ANU, Dec 1994.

Articles from Conference Proceedings (published)

[11] LMS Adaptive filtering for enhancing the quality of noisy speech Sambur, M. ITT Defense Communications Division, Nutley, N. J. Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '78.

[12] Application of the LMS adaptive filter to improve speech communication in the presence of noise Chabries, D. Christiansen, R. Brey, R. Robinette, M. Brigham Young University, Provo, UT, USA This paper appears in: Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '82.

[13] Speech enhancement using a Kalman-based normalized LMS algorithm Mahmoodzadeh, A. Abutalebi, H.R. Agahi, H. Electr. Eng. Dept., Yazd Univ., and Yazd This paper appears in: Telecommunications, 2008. IST 2008. International Symposium on Issue Date 27-28 Aug. 2008

[14] Reduction of nonstationary acoustic noise in speech using LMS adaptive noise cancelling Pulsipher, D. Boll, S. Rushforth, C.Timothy, L. Sandia Laboratories, Livermore, CA This paper appears in: Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '79.

[15] Adaptive noise canceling for speech signals Sambur, M. ITT Defense Communication Division, Nutley, NJ This paper appears in: Acoustics, Speech and Signal Processing, IEEE Transactions on Issue Date : Oct 1978

[16] Paper in IEEE explore entitled "Comparison of LDM and HMM for an Application of a Speech" by Mane, V.A., Patil, A.B., Paradeshi, K.P., Dept. of E&TC, Annasaheb Dange COE, Ashta, India in International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom), 2010 Issue Date: 16-17 Oct. 2010 On page(s): 431-436 Location: Kottayam Print ISBN: 978-1-4244-8093-7 References Cited: 13 INSPEC Accession Number: 11679354 Digital Object Identifier: 10.1109/ARTCom.2010.65 Date of Current Version: 03 December 2010