A review on speech filtering and its different techniques

Minajul Haque[1] and Kaustubh Bhattacharyya[2]
[1]Don Bosco College of Engineering and Technology, Assam Don Bosco University
Airport Road, Azara, Guwahati - 781017, Assam. INDIA.
[1]minajul.h.electronics@gmail.com
[2]kaustubh.bhattacharyya@dbuniversity.ac.in

Abstract: ***Speech is a form of communication that most people came across in their day to day life. Speech can be used for many purposes like speech communication, speech recognition, speaker identification etc. In all of these applications a noise free speech signal is highly desired. Therefore to get a noise free speech, different speech processing and speech filtering techniques are used. In this paper we have presented a review on speech processing and different speech filtering techniques. The techniques that we have investigated in this paper are Adaptive filtering, Kalman Filtering, Neural Networks and adaptive neuro-fuzzy filtering and it is found that LMS gives good result in linear noise depending upon the order of the filter and hybrid filtering techniques like ANFIS because it gives the better result in case of non-stationary noise.***

Keywords: Adaptive Filter, noise cancellation, least mean square adaptive filter, Autoregressive, Kalman filter, neural network, Adaptive neuro-fuzzy filter

1.INTRODUCTION

Speech processing has many applications like wireless communication, teleconferencing system, long distance communication and radar communication [1]. In all these application a noise free speech is highly desired. So to have a noise free speech we need to perform certain task to remove the noise before processing it. Basically noise is mixed with the speech during the generation of the speech or may be in transmitting the speech over a noisy channel or at the receiver end. A noise can be anything like it may be environmental noise or system noise or channel noise but it is essential to remove this noise to enhance the quality of the speech. Removal of noise from speech is also called filtering of speech, for filtering of a speech signal numerous methods and techniques are available among those adaptive filtering is one of the efficient and mostly used technique [2] other methods such as kalman filtering, ANN (Artificial neural network), Neuro fuzzy filters like ANFF (Adaptive neuro fuzzy filter) and ARNFF (Adaptive recurrent neuro-fuzzy filter are also used. Adaptive filters have great impact in noise cancellation in speech processing [2]. All adaptive filters are derived from the wiener filter which is the base of all adaptive filters [2]. Adaptive filter like LMS, NLMS, RLS has used for speech processing [2]. In LMS (least mean square) the mean square error is adaptively used to enhance the speech performance. Although LMS gives good result in linear noise but it fails in non-linear noise, therefore one can use Kalman filter to remove the noise when the noise is non-linear.

The Kalman Filter is an estimator for what is called the *"*linear- quadratic Problem", which focuses on estimating the instantaneous"state"of a linear dynamic system perturbed by white noise [3].

Fuzzy logic deals with reasoning that is approximate rather than fixed and exact. Fuzzy logic variables may have a truth value that ranges between 0 and 1 [3]

Artificial Neural Network includes training, learning and generalization. Training is the process by which these connection weights (W) are assigned [4]. Learning process is based on comparison, between networks computed output and the correct expected output, generating an error. The error generated is used to change network parameter that result improved performance. The NN possesses the capability to generalize. They can predict new outcomes from past trends. The NN is said to generalize well when it sensibly interpolates input patterns that are new to network. When input-output mapping computed by network is correct [5].

2. SPEECH SIGNAL AND THEIR PROCESSING

Speech is a form of communication in everyday life. It is essential to know how we produce and perceive it and how speech technology may assist us in communication [3]. Figure1 shows the general diagram of a human speech production system. In [6] the different parts of the human speech production system have given. Processing of speech is very much important in communication. We always need to process the speech before its application. The processing of speech and its modeling is described below:

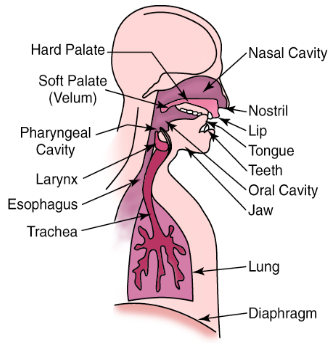

Figure 1:Human vocal system [6]

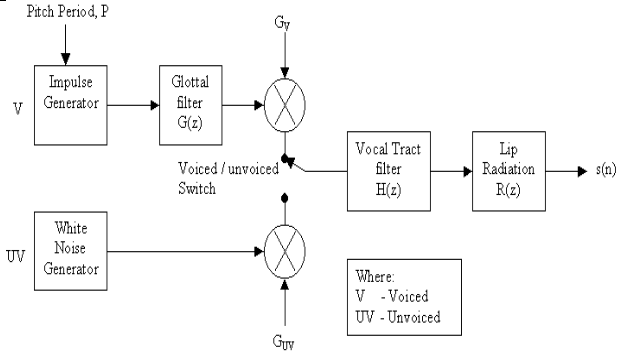

Figure 2 - A general discrete time model for speech production [6]

Analyzing Figure 2, we may derive a system transfer function for voiced and unvoiced speech which is represented by equation (1) and equation (2) respectively.

$$S(z) = E(z).G(z).H(z).R(z) \qquad (1) \qquad [6]$$

$$S(z) = E(z).H(z)R(z) \qquad (2) \qquad [6]$$

3. ADAPTIVE FILTER

Adaptive filters are very much useful in speech filtering. There are many kinds of adaptive filters are available for speech filtering like LMS, NLMS and RLS adaptive filter etc. Although adaptive filter are best for speech filtering but in some cases the above mentioned filters fails to meet the desired performance [7], so in that case we need some more sophisticated and more powerful filters which lead us to the Kalman filter. Kalman filter is also an adaptive filter but it performs well in the cases where LMS fails. The LMS (least mean square) algorithm was developed by Widrow and Hoff in 1960 and it is a member of the stochastic gradient algorithms [8].In [8] Jan Vaňuš, Vitězslav Styskala Gives the optimal setting of the parameter of the adaptive noise canceller with the conventional LMS, with the Signed-Regressor LMS algorithm, with the Sign LMS algorithm and with the Sign-Sign LMS algorithm. The conventional LMS adaptive filter weights update rule is modified into sign-sign LMS algorithm which is the combination of the sign LMS and signed-regressor LMS algorithm.

Mean square error (MSE) used for comparison between original signal (without noise) and recovered signal [7]

The conventional LMS adaptive filter weight update rule is given by equation (3)

$$w(n+1) = w(n) + 2\mu e(n)x(n) \qquad (3)$$

The modified sign-sign LMS algorithm rule is given by equation (4)

$$w(n+1) = w(n) + 2\mu \,\mathrm{sign}(e(n))\,\mathrm{sign}(x(n)) \qquad (4)$$

Where

$w(n)$ Is *M*-tap weight vector

$w(n+1)$ Is *M*-tap weight vector update

$e(n)$ Is the estimated error and

$x(n)$ Input signal

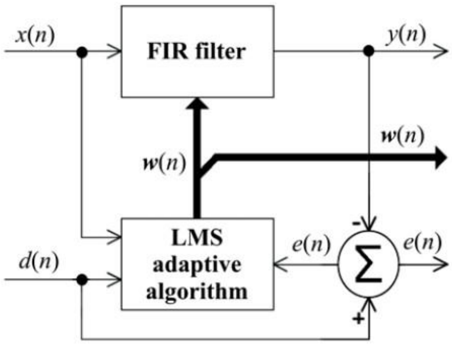

Figure 3*:* Block diagram of the LMS adaptive filter. [8]

The block diagram of an LMS adaptive filter is shown in the Figure 3. Now setting the value of the parameter step size µ, we can have an optimal performance of the LMS adaptive filter. To maximize the convergence speed of the LMS algorithm, a big step size is required. A large value of the µ can quickly lead to the optimal settings of the LMS algorithm for speech signal processing [8]. Where as a high value of µ may increase an estimate error of the speech signal and a small value of µ can ensures the stability and the convergence of the LMS algorithm [8]. But the LMS adaptive noise cancellation suffers response degrades and slow convergence rate under low Signal-to-Noise ratio (SNR) condition. The performance of LMS adaptive algorithms is highly dependent on their filter order and signal condition $w(n+1) = w(n) + 2\mu e(n)x(n)$

Where, *x(n)* is the input vector of time delayed input values, *w(n)* is the coefficients of the adaptive FIR filter tap weight vector at time *n* and µ is known as the step size. Selection of a suitable value for μ is imperative to the performance of the LMS algorithm, if the value is too small, the time adaptive filter takes to converge on the optimal solution will be too long; if μ is too large the adaptive filter becomes unstable and its output diverges [10],[11].An adaptive filter comprises of two basic components, these are a digital filter and an adaptive algorithm. The digital filter produces an output in response to an input signal and the adaptive algorithm is responsible for adjusting the coefficients of the digital filter. The adaptive filter is more effective when the carrier frequency of the narrowband interference is offset from the carrier of the spread spectrum signals [8]. The major advantages of LMS filter are:

- It is very effective in rejecting the narrowband interference when the ratio of the narrowband interference bandwidth to the spread spectrum bandwidth is small.
- It is simple and powerful [9].

The disadvantages are:

- High computational complexity,
- Highly depends on the step size and fails to meet the desired performance in nonlinear situation [9].

To overcome the disadvantages of the LMS adaptive filter we used a more powerful filter i.e. kalman filter which give more accurate estimation of the noise. Kalman filter is a linear quadratic problem estimator which gives the optimal estimation of the speech signal mixed in noise [3]. Considering a model of noisy speech signal given by

$$y(n) = s(n) + w(n) \qquad (5)$$

where n=1,2,3…and y(n), s(n) and w(n) denote discrete time samples of noisy speech, clean speech and noise, respectively [12]. Now from [13], [12] the model for speech is given by equation 1

$$s(n) = \sum_{k=1}^{q} a(n,k)s(n-k) + e(n) \qquad (6)$$

where e(n), the excitation signal, is generated by a zero-mean white Gaussian process with variance $\sigma^2_{e(n)}$, a (n,k)'s are the adaptive filter coefficients, q is the filter order, and s(n) is the clean output speech. Now the above equation the excitation e(n) is modified to represent the both voiced and unvoiced speech as well as the silence i.e. e(n)=b(n, p_n)e(n-p_n)+d(n) where $d(n)$ is generated by zero-mean white Gaussian process with variance $\sigma^2_{d(n)}$, p_n is the instantaneous pitch period and b(n, p_n) is the periodicity of of the speech. Now with the help of these 3 above model kalman filter gives the best estimate of the noise [12]. Similar work has done in [14],[3] with AR model and it is found that kalman filter gives the best estimation of the clean speech mixed with noise signal and also it gives the good channel estimation than LMS. The author zenton goh in [12] states that the standard algorithm for Kalman filtering involves multiplications of very large matrices and thus required high computational cost; they improve the efficiency quite significantly by exploiting the sparsity of the matrices concerned. Here the computational complexity is reduced by 1/900 times. Kalman filter normally assumes the process noise and observation noise are both uncorrelated and have normal distributions. This implies that the Kalman based method is best suitable for reduction of white Gaussian noise. The Kalman filter is derived based on two assumptions, linearity and Gaussian noise [15],[16]. Figure 4 shows a generalize block diagram of kalman filter application.

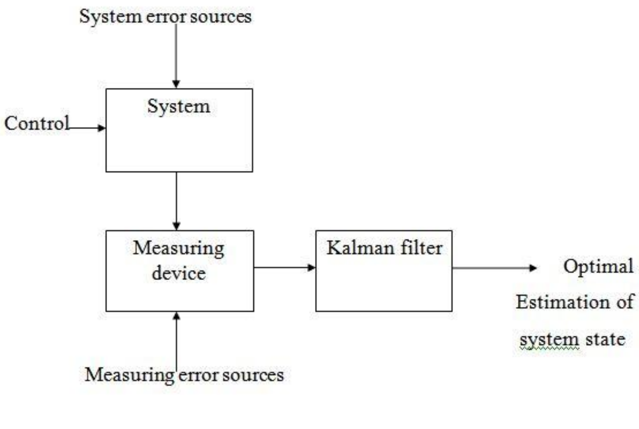

Figure 4:Block Diagram of typical Kalman Filter application. [3]

4. NEURAL NETWORK

The adaptive filtering techniques like LMS, NLMS and Kalman filter gives a optimal results unless the noise is stationary but in case of a non-stationary noise performance the above filtering techniques degrades and this disadvantages is over come by the use of Artificial neural networks [17]. Application of ANN in noise reduction problem is increasing day by day and the main design goal of these Neural Networks (NNs) was to obtain a good approximation for some input output mapping [18]. Neural networks are composed of simple elements operating in parallel. Neural networks have been trained to perform complex functions in various fields of application including pattern recognition, identification, classification, speech, and vision and control systems [10]. Vartika Anand, Shalini Shah, Sunil Kumar proposed an FIR digital filter in [19] to train a neural network [19]. The experimental results shows that using neural networks in noise separation produce a more robust and powerful separation of speech and noise than other LMS and Kalman filter. In their experiment the neural networks such as Elman, Radial Basis Function and Perception networks are trained with different training algorithms and compared with the performance of FIR digital filter including its computational complexity and it is found that the algorithm use to train the neural network is very important and has effect on the final results. Using neural networks in noise classification produce a more robust and powerful separation of speech and noise than other traditional algorithms. In [17] four ANN model s namely Function Fitting (FitNet), Nonlinear Auto Regressive (NARX), Recurrent Neural Network (RNNs), and Cascaded-Forward Net were developed to filter a noise mixed speech signal. The four models were trained separately on stereo (noisy and clean) audio signals to produce the clean signal. These four ANN models constructed with 3 layers (input, hidden and output) for speech signal enhancement to remove noise from speech signal as shown in Figure 5. Two neurons in input layer to represent speech signal and its associated noise. The output layer includes one neuron that represents the enhanced signal after removing noise. Different numbers of hidden neurons were used for each model (10, 20 and 30). Three training algorithms (GD, GDM and LM) were used to train each one of the four models separately to become a filter. It was found that FitNet and NARAX models produces best results respectively and Training algorithm LM is the best training algorithm in this case [17].

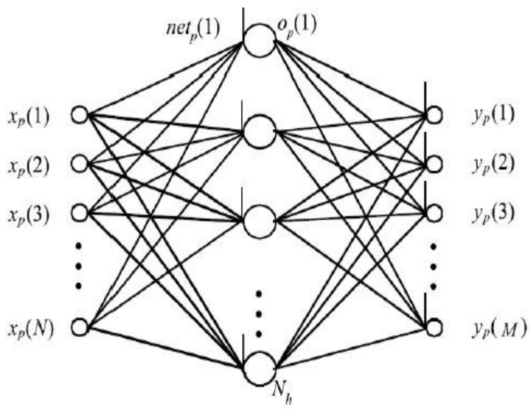

Figure 5: Multilayer feed-forward ANN [17]

Simillar work has been done in paper [10], where particle filters are combined with neural networks presents a solution to computational cost constraint of particle filter by minimizing no. of particles required by preprocessing the speech.

5. NEURO-FUZZY SYSTEM

To enhance the speech signal many algorithm and techniques are used, but the hybrids technique like the combination of fuzzy logic and neural networks gives more powerful techniques for adaptive signal processing. And then the adaptive neural fuzzy filter (ANFF) algorithm was developed [20].

ANFIS is a Neuro-fuzzy system that combines the learning capabilities of neural networks, with the functionality of fuzzy inference systems. The ANFIS is functionally equivalent to fuzzy inference systems. It's easy to implement and fast convergence [20].

In paper [4] an adaptive neuro-fuzzy filtering scheme has given using the artificial neuro-fuzzy inference system (ANFIS) for noise reduction in speech. In Figure 7, a 5^{th} order ANFIS structure is shown to cancel the noise from the speech and the fuzzy system has used two rules which are

1. If x is A] and y is B], then fl =pl x+ql y+rl
2. If x is A2 and y is B2, then f2=p2 xA+q2 y+r2

And finally the ANFIS hybrid-learning algorithm is applied to determine the other parameter which is required in noise cancellation. Thus the ANFIS scheme is quite acceptable for non-linear noise cancellation [4]. The process of using ANFIS for solving the adaptive noise cancellation problem is given in Figure 6.

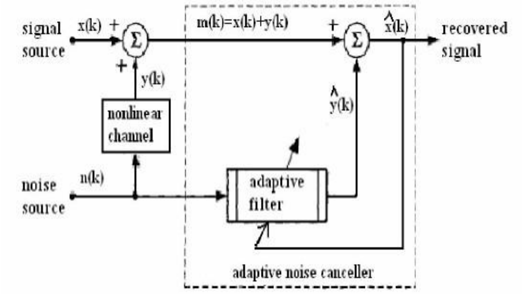

Figure 6: Schematic diagram of adaptive noise cancellation [4]

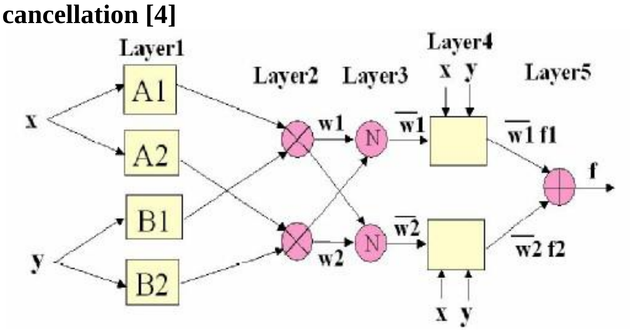

Figure 7: ANFIS structure [4]

Similarly in [21] noise canceling problem is solved for a long distance communication using the LMS and ANFIS filter and it was found that the ANFIS provides better result in low SNR environment i.e. it is more efficient to eliminate noise and also offers faster convergence time. In [22] a filter was developed combining the NLMS, Kalman filter and Fuzzy logic to solve the nonlinear problem of speech enhancing termed as normalized fuzzy logic kalman filter (NFLKF). The fuzzy system was used, which starts from the establishment of If-Then rules and processes the signal via the optimal fuzzy kalman filter based on consistent improvement and self adaptive computation adjustment in the fuzzy system and kalman filter parameter modulation [22].

6. CONCLUSION

In this review paper we have investigated different speech processing and different noise reduction techniques that have been used in the speech recognition system. The techniques that were discussed in this paper are Adaptive filter, LMS adaptive filtering, Kalman filter, Neural network and adaptive neuro-fuzzy filter (ANFIS) and it was found that, although LMS adaptive filter perform well in linear condition and depends on the order of the filter but in nonlinear case it performance degrades where as in Kalman filter gives better result in comparison to the LMS. The neural network and adaptive neuro-fuzzy filter gives the optimal results in non stationary noise.

ACKNOWLEDGEMENT

The authors would like to acknowledge Assam Don Bosco University, Guwahati for providing the resources necessary for compiling this review article.

REFERENCES

[1] V.J. Naveen, T.prabakar, J.V. Suman and P.Devi Pradeep, "Noise suppression in speech signals using adaptive algorithm", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 3, No. 3, September, 2010.

[2] Monson H. Hayes, "statistical digital signal processing and modeling", Wiley India pvt ltd.

[3] V.Murikipudi, K.Phani at all, "Design and analysis of speech processing using kalman filtering", Journal of Theoretical and Applied Information Technology, © 2005 - 2011 JATIT & LLS.

[4] J.Thevaril and H. K.Kwan, "Speech Enhancement using Adaptive Neuro-Fuzzy Filtering", Proceedings of 2005 International Symposium on Intelligent Signal Processing and Communication Systems, December 13-16, 2005 Hong Kong.

[5] P.Bactor and A.Garg, "Different Techniques for the Enhancement of the Intelligibility of a Speech Signal", International Journal of Engineering Research and Development Vol 2, Issue 2 (July 2012), PP. 57-64.

[6] K.Dutta and M.Sarma, "Assamese Numeral Corpus for Speech Recognition using ANN", gauhati university, June 2010.

[7] F.Kosser, Sumaya, G.M.Choudhary and A.Garg, "A Brief Review of Speech Enhancement Algorithms", International Journal of Research in Information Technology, ISSN 2001-5569.

[8] J.Vaňuš, V.Styskala, "Application of variations of the LMS adaptive filter for voice communications with control system", ISSN 1330-3651.

[9] S.B.Marge, P.V.Janse and R.R.Deshmukh, "A Review on Feature Extraction and Noise Reduction Technique", International Journal of Advanced Research in Computer Science and Software Engineering, Vol 4, Issue 2, February 2014.

[10] V.Anand, S.Shah, S.Kumar,"Intelligent Adaptive Filtering For Noise Cancellation",International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 2, Issue 5, May 2013.

[11] R.Singh and A.Mishra, “Speech enhancement using practical filters: A critical review”, Indian Journal of Computer Science and Engineering (IJCSE).

[12] Z.Goh, K.Tan and B.T.G. Tan, “Kalman-Filtering Speech Enhancement Method Based on a Voiced-Unvoiced Speech Model”, IEEE transaction on speech and audio processing, vol.7, No.5, September 1999.

[13] Cai H, Grivele E. and Najim M., “Speech enhancement using AR model driven by white noise with time-varying variance”, ESI, UMR LAP5131, ENSEIRB-Université Bordeaux 1 BP 99, F 33402 talence Cedex, France.

[14] T.Paul, P.Karmakar and T. Paul, “Performances of Kalman filter and LMS Algorithm in Channel Estimation”,International Journal of Engineering Research &Technology (IJERT) Vol. 1 Issue 5, July - 2012 ISSN: 2278-0181.

[15] K.S.Panchal, “Noise reduction from the speech signals using WP coefficient and modified thresholding”, IJIRST –International Journal for Innovative Research in Science & Technology, Vol 1, Issue 3 August 2014 ISSN : 2349-6010.

[16] Omaima N. A. AL-Allaf, Al-Zaytoonah, “Removing Noise from Speech Signals Using Different Approaches of Artificial Neural Networks”, I.J. Information Technology and Computer Science, 2015, 07, 8-18 Published Online June 2015 in MECS .

[17] Widrow, B. et, al, “Adaptive Noise Cancelling: Principles and Applications”, Proc. IEEE, 63(12), pp: 1692-1716, 1975.

[18] L. Badri, “Development of Neural Networks for Noise Reduction”, the International Arab Journal of Information Technology, Vol. 7, No. 3, July 2010.

[19] S.K.Hasnain and Dr.P.Akhtar, “A Fair Play Comparison of Neural Networks Algorithms in Speech Signal Processing”.

[20] P.Bactor and A.Garg, “Review of adaptive noise cancellation techniques using Fuzzy-Neural networks of speech”, International journal of computing and corporate research, vol 2 issue 4 July 2012.

[21] Mohammed Hussein Miry Ali Hussein Miry Hussain Kareem Khleaf, “Adaptive Noise Cancellation for speech employing fuzzy and neural network”, Iraq J. Electrical and Electronic Engineering, Vol.7 No.2, 2011.

[22] Mao-lin Chen, Hung-Yan Gu and Ching-Long Shih, “A new speech enhancing scheme combing NLMS, Fuzzy Logic and Kalman filtering”, 2008 IEEE International Conference on Fuzzy Systems.

Authors’ Profile

Minajul Haquehas completed his B.Tech degree from the Department of electronics and telecommunication engineering Of Konark Institue of Science and Technology, Bhubaneswar, Odisha India in 2014. He is currently pursuing his M.Tech degree in the department of ECE with specialization in Digital Signal Processing in School Of Technology, ADBU, Guwahati, India. His area of interest is Speech Processing and Communications.

Kaustubh Bhattacharyya is currently serving as an Assistant Professor in the Electronics and Communication Engineering Department of Assam Don Bosco University, Assam, India. Prior to this he held research positions at Gauhati University, Assam, India. He completed his Bachelor of Science in Electronics, Master of Science in Electronics Science, Master of Philosophy in Electronics Science and Master of Technology in Electronics and communication Technology from Gauhati University, Assam, India in 2004, 2007, 2009 and 2012 respectively. His research interest is in high frequency communication and soft computing.