Speech Denoising Using Functional Link Artificial Neural Network

[1]Niraj Kumar Shukla, [2]Silpa Nayak, [3]Ravi Kant Jha

Department of Electronics and Communication Engineering
GIET, GUNUPUR

***Abstract*: Speech processing is a major field of use in present day. When a Speech sent from source to destination, it gets noisy due to different reasons. Mainly this noise added at the time of acquisition, processing and at the time of transmission of the speech signal. Low light level, sensor temperatures, sensor noise, poor illumination, high temperature of the surroundings are the major factor affecting the speech Again when this speech transmitted to the destination, it corrupted due to interference in the channel use for transmission. Speech signal Image transmitted from in wireless network might be corrupted as a result of lighting or other atmospheric disturbances. Different noises added in this image are additive white Gaussian noise, salt and pepper noise, Both Gaussian and salt and pepper noise, Rayleigh Noise, Erlang Noise, Exponential Noise, Uniform Noise, and Periodic Noise etc. In this paper various noise conditions are studied and efficient adaptive filters based on Functional link artificial neural network are designed to suppress Gaussian noise. The developed filters may use for offline or for online applications interference in the channel use for transmission. In this paper various noise conditions are studied and efficient adaptive filters based on Functional link artificial neural network are designed to suppress Gaussian noise.**

***Keywords:* Neural network, FLANN, SS, Sensor, ASR, etc.**

1 INTRODUCTION

Speech enabled applications in public areas such as train stations, airports or tourist information centers might serve the customer with answers to their spoken query. Physically handicapped or elderly people might also be able to access services in a more natural way, since the use of a keyboard is not required. Speech enabled applications in public areas such as train stations, airports or tourist information centers [1] might serve the customer with answers to their spoken query. Physically handicapped or elderly people might also be able to access services in a more natural way, since the use of a keyboard is not required. There have been various efforts already done towards Automatic Speech Recognition (ASR)[2] in Indian languages like Oriya,Tamil, Telgu, Marathi and Hindi. The recognition performance, in all cases, is acceptable in noise free condition, but it degrades dramatically in the presence of noise. Investigation of robust features for recognition of noisy speech is still an active area of research [1]. If the recognition system is noises [3] such as aircraft noise, speech of other persons etc. are present, then the performance of system comes down because the system is trained with clean speech [4] and testing is performed in noisy environment. So there is need to analyze the performance of the system in noisy as well as noise free conditions [5]. The analysis approach used in this paper first adds controlled amounts of various types of noises to clean training data with specified signal to noise ratio. The recognition performance is analyzed with noise suppression techniques [6] applied to reduce the noise from the corrupted noisy signal.
Different noise suppression techniques [7] are used. The Spectral Subtraction (SS)[4] method is one of the well-known techniques for noise reduction from the signal. The SS estimates the power spectrum of clean speech by explicitly subtracting the noise power spectrum from the noisy speech power spectrum, resulting in a signal with better signal-to-noise ratio.
In Speech waveform there is non-stationary component manifests as spectral peaks and valleys whose width and temporal duration depends on the noise. To reduce noise knowledge Based Spectral Subtraction (KBSS)[8] method is used. KBSS is the combination of spectral subtraction and the speech enhancement methods. The efficiencies of three speech enhancement techniques [2] (SS, NSS, KBSS) for speech recognition are studied in this work. The baseline speech recognition system, against which the evaluation experiments were carried out, was developed using Sphinx [9] , a software toolkit for speech recognition. In this paper we carried out an analysis of the performance of a speaker independent, continuous speech recognition system in the task of recognizing spoken Hindi under additive noises of various types.

2 NOISE IN DIGITAL SPEECH

When an speech taken, stored, processed and sent to destination, it gets noisy due to different reasons at the time of acquisition, processing and at the time of transmission. Speech acquired through modern sensors may be contaminated by a variety of noise sources. By noise we refer to stochastic variations as opposed to deterministic distortions such as shading or lack of focus. A speech signal gets corrupted with noise during , transmission, storage and retrieval processes. Acquisition noise is usually very low variance Gaussian noise. The acquisition noise is quite negligible except certain particular application like remote sensing, biomedical instrumentation etc. Some noise added at the time of digital processing of the speech signal image.

2.1 NOISE MODEL

The principal source of noise in digital speech is arises from speech acquisition and transmission. Light level and sensor temperature are the major factor affecting the amount of noise in the speech. Speech signal is corrupted during transmission due to interference in the channel use for transmission. Speech transmitted in wireless network might be corrupted as a result of lighting or other atmospheric disturbances. Except the spatial periodic noise, noise is independent of spatial co-ordinate. It is uncorrelated with respect to image itself.

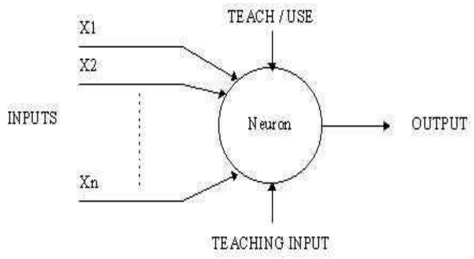


Fig 3.1: A simple neuron

3 ATRIFICIAL NEURAL NETWORK AND FLANN

The fixed filters are good for offline applications where speech is filtered by using computer and some software algorithm. Here the expert knows the type of noise and the noise power level. So he/she can apply a specific filtering operation depending upon the requirement. He may change the filtering operation depending upon the requirement. But such human decision cannot be taken for an online and real time operation. For example when data is transmitted in channel the noise added varies from time to time. Again it changes in fraction of second. A human expert can't take decision to choose a filter at that small time. To avoid different limitations of fixed filters, adaptive filters are designed that adapt themselves to the changing conditions of signal and noise. The filter characteristics change as the signal statistics or noise type or noise power level vary from time to time. One of the broad categories of adaptive speech filters developed for efficient noise suppression and presented in this paper is (i) FLANN based filter.

3.1 Neural Network

Artificial Neural Networks are being touted as the wave of the future in computing. They are indeed self-learning mechanisms which don't require the traditional skills of a programmer. Yet these networks are simply tools and as such the only real demand they make is that they require the network architect to learn how to use them.

3.2 A simple neuron

An artificial neuron is a device with many inputs and one output. The neuron has two modes of operation; the training mode and the using mode. In the training mode, the neuron can be trained to fire (or not), for particular input patterns. In the using mode, when a taught input pattern is detected at the input, its associated output becomes the current output. If the input pattern does not belong in the taught list of input patterns, the firing rule is used to determine whether to fire or not.

In mathematical terms, the neuron fires if and only if; X1W1 + X2W2 + X3W3 + ... > T ….. 3.1

The addition of input weights and of the threshold makes this neuron a very flexible and powerful one. How the input and output data is represented, or encoded, is a major component to successfully instructing a network.

4 FUNCTIONAL LINK ARTIFICIAL NEURAL NETWORK

The FLANN, which is initially proposed by Pao, is a single layer artificial neural network structure capable of forming complex decision regions by generating nonlinear decision boundaries. In a FLANN the need of hidden layer is removed. In this functional link it acts on an element or the entire pattern itself by generating a set of linearly independent functions. Here the functional expansion block comprises of exponential series or subset of Chebyshev polynomials. Let input to this structure is $X = [x_1 x_2]'$. An enhanced pattern obtained by using functional expansion is given by $X = [1 x_1 T_2(x) ....x_2 T_2(x_2)............]'$ …. . 4.1

5. THE CHEBYSHEV NEURAL NETWORK

The Chebyshev polynomials are a set of orthogonal polynomials defined as the solution to the Chebyshev differential equation and denoted as Tn(x). The structure of a chebyshev neural network is shown in . It is similar to FLANN, except that the input vector is expanded using the Chebyshev polynomials to form the enhanced high dimensional space.The first few Chebyshev polynomials are given by $T_0(x) = 1.0$, $T_1(x) = x$ and $T_2(x) = 2x\char`^2 - 1$. The higher order Chebyshev polynomials may be generated by the recursive formula given by:

$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$. The first few Chebyshev polynomials are given by

T0(x) = 1

T1(x) = x

T2(x) = 2x2-1

T3(x) = 4x3-3x

5.1 DEVELOPMENT OF ADAPTIVE FILTER USING FLANN

Let input to this structure is $X = [x_1 x_2]'$. An enhanced pattern obtained by using functional expansion is given by

$X = [1x_1 T_2(x).......x_2 T_2(x_2)...........]'$ ………. 5.1

This enhanced pattern can be used for classification or estimation purposes. For functional expansion of the input pattern, the trigonometric, power series, exponential and Chebyshev polynomials is chosen individually and compare the output individually.

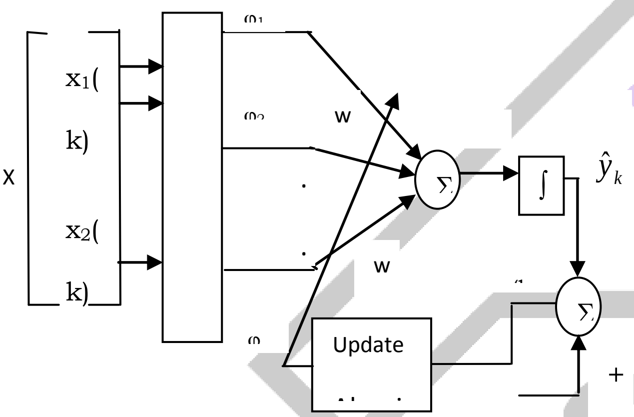


Fig 5.1 Proposed network with reference to speech

5.2 Algorithm for FLANN

In the FLANN a set of basic functions Φ and a fixed number of weight parameters $'W'$ are used to represent $f_w(x)$. With a specific choice of a set of functions the problem is then to find the weight parameters $'W'$ that provides the best possible approximation of $'f'$ on the set of input-output examples. This can be obtained by recursively updating W. Detail theory on the FLANN may be found in [2]. Let the training pattern be denoted by $\{X_{k,} Y_k\}$ and the weight metrics by $W(k)$. Discrete time index k is given by $k = k + \lambda K$, for $k = 1,2,,.....K and \lambda = 0,12,...$ where K is the total number of training patterns. At k th instant, the $n-$dimensional input pattern and the m-dimensional FLANN output are given by

$X_k = [x_1(k)x_2(k).........x_n(k)]^T$ and $\hat{Y}_k = \left[\hat{y}_1(k)\,\hat{y}_2(k)...........\hat{y}_m(k)\right]^T$ respectively… 5.2

Its corresponding target pattern is represented by

$Y_k = [y_1(k)y_2(k)..............y_m(k)]^T$ ………. 5.3

5.3 Chebyshev Expansion

In this paper we used Chebyshev polynomial for functional expansion. These polynomials are easier to compute than that of trigonometric polynomials. Here in our study, we found superior performance by using CFLANN. The FLANN structure consider for denoising purpose using Chebyshev functional expansion is depicted here.

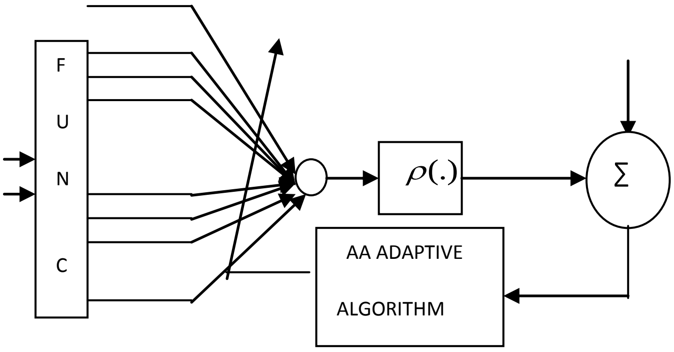


Fig.5.2 Chebyshev filter

The first few Chebyshev polynomials are given by $T_0(x)=1.0$, $T_1(x)=x$ and $T_2(x)=2x\char`^2-1$. The higher order Chebyshev polynomials may be generated by the recursive formula given by:

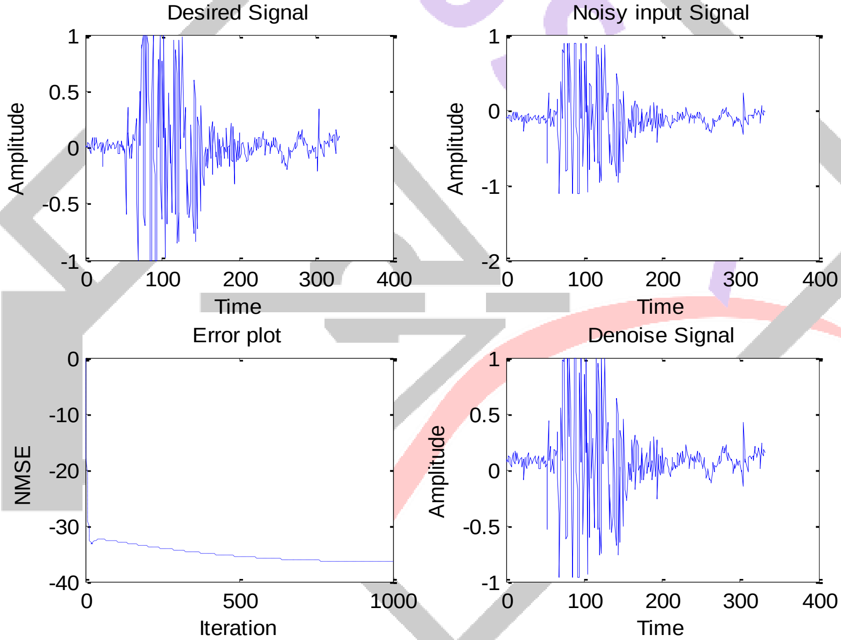
$$T_{n+1}(x)=2xT_n(x)-T_{n-1}(x).$$

(i) Forward calculations to find the activation value of all the nodes of the entire network.

(ii) Back –error propagation for calculation of square error derivatives.

(iii) Updating weights of the entire network.

Result



6 CONCLUSION

Here we have proposed use of single layer FLANN structure which is computational efficient for denoising of speech corrupted with Gaussian noise. The functional expansion may be thought of analogous to the nonlinear processing of signals in the hidden layer of an MLP. This functional expansion of the input increases the dimension of the input pattern. In the FLANN structure proposed for denoising of speech, the input functional expansion is carried out using the trigonometric, exponential or Chebyshev polynomials. The prime advantage of the FLANN structure is that it reduces the computational complexity without any sacrifice on its perrformance. Simulation results indicate that the performance of FLANN is better than MLP for Gaussian noise suppression from an speech. From these work it is clear that FLANN having Chebyshev Functional expansion is better for Gaussian noise suppression than other FLANN structure. The FLANN structure having Chebyshev functional expansion may be used for online speech processing application due to its less computational requirement and satisfactory performance. The new nonlinear adaptive filter FLANN shown satisfactory results in its application to speech with additive noise. Its adaptive capacity to different parameters when generating the speech with Gaussian noise has to be studied. Generalization of this filter applicable to other types of Denoising process. A comparison of the computational requirements in one iteration of algorithm for the FLANNs is provided in this table

Table 6.1

Nos.	FLANN
Wts	$n_l(n_0+1)$
Adds	$2n_l(n_0+1)+n_l$
Multi	$3n_l(n_0+1)+n_0$
Tanh()	n_l

Where,
Wts be the weight,
Add be the addition
Multi be the multiplication and
tanh represents tanh (.) respectively

References

[1] P.J.Antsaklis,"Neural network in control system."IEEE, Control.Syst.Mag.PP. 3-5. April.1990.
[2] S. Haykin. Neural Networks. Ottawa.ON.Canda Maxwell Macmillan. 1994.
[3] R Grino.G.Cembrano and C.Torres."Nonlinear system identification using additive dynamic neural networks two on line approaches."IEEE Trans Circuits SystemIvol47 150-165.Feb 2000.
[4] T.Poggio and F.Girosi."Networks for approximation and learning."Proc. IEEE ,vol 78,pp.1481-1497,sep1990.
[5] Q.Zhang and A.Benvenister."Wavwlet networks ."IEEE Trans Neural Network vol 3pp 889-898Mar 1992.
[6] K.H.Chon and R.J Cohen. "Linear and nonlinear ARMA model parameter estimation using an artificial neural neural network."IEEE Trans Biomed.Engg Vol 44.pp168-174,Mar 1997.
[7] Patra,J.C. Pal ,R,N. Chatterji,B,N. Panda,G. "Identification of nonlinear dynamic systems using functional link artificial neural networks"**IEEE Transactions , Systems, Man and Cybernetics, Part B,** Vol 29 ,pp 254 - 262,April-1999
[8] A.Namatame and N.Ueda."Pattern classification with Chebyshev neural networks," Ind.J.Neural Networks ,Vol 3,pp23-31,Mar. 1992.
[9] Patra, J.C.; Kot, A.C. "Nonlinear Dynamic System Identification Using Chebyshev Functional Link Artificial Neural Networks", *IEEE Transactions*, Vol 34,pp: 1627- 1627, June 2004
[10] Teck Hou Teng , Patra,J.C,Ee-Luang Ang "Identification and tracking of dim moving targets in FLIR using artificial neural networks" **Seventh International Symposium ,Signal Processing and Its Applications, 2003,**vol.2,**pp**223-226,2003
[11] Weng,W.-D , Yen,C.T,"Reduced-decision feedback FLANN nonlinear channel equaliser for digital communication systems"**IEE Proceedings-**Vol4pp305-311Aug.2004