

## Speech Enhancement Using Wavelet Neural Network with Sub-Band Adaptive Matched Filter

Dan Yang<sup>1, a</sup>, Bin Xu<sup>2, b</sup>, Linlin Ye<sup>1</sup> and Xu Wang<sup>1, c</sup>

<sup>1</sup> School of Information Science and Engineering, Northeastern University, Shenyang City, Liaoning Province, China, 110819

<sup>2</sup> Computing Center of Northeastern University, Shenyang City, Liaoning Province, China, 110819

<sup>a</sup>yangdan@ise.neu.edu.cn, <sup>b</sup>xubin@mail.neu.edu.cn, <sup>c</sup>wangxu@ise.neu.edu.cn

**Keywords:** Wavelet Neural Network, time-frequency analysis, Speech Enhancement, sub-band adaptive matched

**Abstract.** Wavelet neural network (WNN) is a time-frequency analysis method, which detects the subtle small changes in the signal frequency domain. Adaptive filter provides a kind of simple and applied method for processing signals in noise. In this paper, we proposed a new speech enhancement technique which is based on wavelet neural network using adaptive matched filter adjusting weight. We choose the signal with noise pollution as the input signal and then put it to the trained wavelet neural network. Wavelet decomposition and wavelet neural network weights processing adopt signal sub-band adaptive matched filter, the output signal of wavelet neural network is an approximation form of original signal. The results show that the WNN is a quite effective method for the speech enhancement and improving the ration of signal to noise.

### Introduction

The speech signal is often accompanied by the environment noise [1]. Hence speech enhancement plays an important role in the recognition or compression of speech signals system performances. Some well-known speech enhancement methods are Winner filter, linear prediction and spectral subtraction, achieved by the use of linear processing techniques [2]. Recently, more effort has been done on nonlinear techniques for speech enhancement utilizing like wavelet transform. The main idea of these methods is to decompose the input noisy speech signals into critical sub-band domains, which is designed to match the psychoacoustic model and to improve the performance of speech processing as human perception.

Wavelet neural network (WNN) is a feed forward neural network, which combine the capability of artificial neural networks in learning processes and the capability of wavelet decomposition [3,4,5,6]. It offers efficient functional representations that build on the time-frequency localization property of wavelets. Much research has been done on applications of WNN, such as identification and control of dynamic systems, background deduction and recovery of characteristic information.

This paper presents a comprehensive investigation of practicality of using WNN and sub-band adaptive matched filter to extract clean speech signals from the noisy environment.

In this paper, we propose a new approach for speech enhancement. The used approach is to achieve the clean speech with wavelet neural network and sub-band adaptive matched filter. We discuss how to construct the structure of WNN based on sub-band adaptive theory and how to train the weights of WNN. The proposed techniques were successfully applied on the detection of speech signals in the noisy environment.

This paper is organized as follows. Section 2 describes the structure of the WNN-based on a sub-band adaptive matched filter along. Some implementation issues concerned with the weight training scheme are presented in Section 3. Section 4 presents the experimental results, which are finally followed by some concluding remarks and future work proposals outlined in Section 5.

### WNN with sub-band adaptive matched filter

A wavelet network has been proposed for approximating arbitrary nonlinear function, which absorbs the advantages of high resolution of wavelets and the advantages of learning and feedforward of neural networks[7]. The typical output of wavelet neural network can be written as :

$$y_j = \sigma(\sum_{h=1}^H w_{hj} \sum_{n=1}^N x(n) \psi(\frac{t_n - b_h}{a_h})) \quad (1)$$

In the proposed WNN with sub-band adaptive matched filter, the training function of wavelet node can be partitioned into two categories: One is a scaling function  $\phi_{LK}(x)$ , which is an orthonormal basis function under different offsets in the resolution  $L$ , forming the approximation to expected input under the roughest resolution; the other is a wavelet function  $\psi(x)$ , which is detailed orthonormal basis function. The weight's coefficient of output cell can be expressed as the following approximate form, which sampling points per unit length in input space with the lowest resolution  $L$

$$F(x) \approx F_L(x) = \sum_k a_{Lk} \phi(2^L x - k) \quad (2)$$

If we need to improve the approximation, we can decompose the space  $V_L$  by scale. We can get orthonormal project  $F_{L-1}(x)$  of  $F(x)$  under the space  $V_{L-1}$ , then

$$F(x) \approx F_{L-1}(x) = \sum_{k=1}^n a_{Lk} \phi(2^L x - k) + \sum_{k=1}^n d_{Lk} \psi(2^L x - k) \quad (3)$$

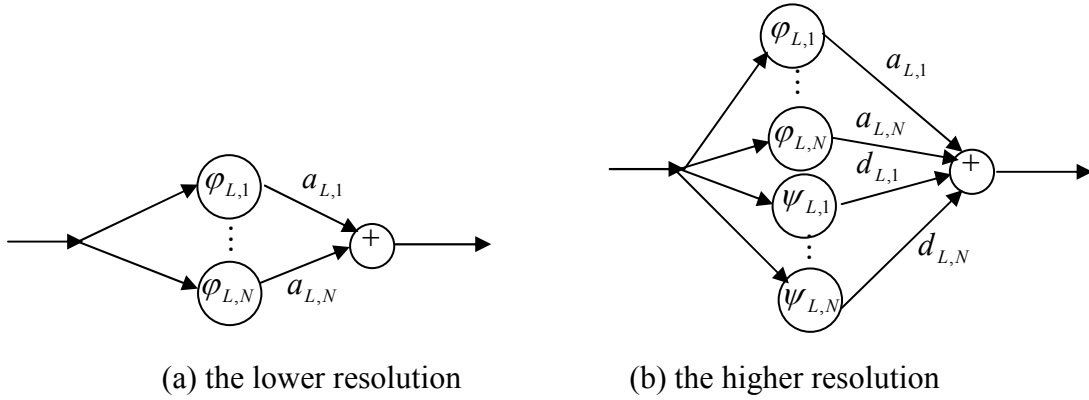


Fig. 1. Wavelet neural network with different resolution

The structure of the wavelet neural network with different resolution is shown as Fig.1. During training stage, we do not need to change the coefficient learnt before, only need to learn the coefficients of new detail. The location and width of node's functional functions are determined by sampling points and sampling cycles. The typical relationship of the input and output of network can be written as follows:

$$F(x) \approx f(x) = \sum_{m,k} C_{mk} \phi_{mk}(x) \quad (4)$$

### The proposed Algorithm Description

At first, we trained wavelet neural network with the noise signals. The noise signal is decomposed into different frequency domains by wavelet transform. The sub-band signals of the noise signal at the different scales are used as the input signals of the WNN. The corresponding outputs are denoted as  $y_i(k)$ . The expected output of the different sub-band signal is nearly zeros. The weights of the WNN are adjusted by the variances of the output signals. The training goal is to make the error of each sub-band energy minimum. Then we trained another wavelet neural network with the speech signals. The object of weights training is that the input original speech signal does not change. The WNN may be an all-pass filter for the input original speech signal. Now the two trained WNNs are combined into the one, which is all-pass filter for the speech signals and for the noise signals is stop-band filter. When the two networks with different hidden layer node, the hidden layer nodes and

weights for the original speech signals are used as the new wavelet neural network hidden layer nodes and weights. If the two networks have the same hidden layer node, take the weight of the smaller hidden layer nodes, added to the new wavelet neural network hidden layer node.

For the trained wavelet neural network, the polluted speech signal can be filtering on line, online filtering phase diagram as shown Fig.2.

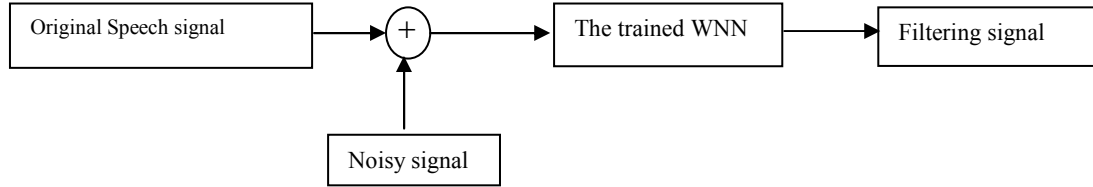


Fig. 2. Filtering structure of wavelet neural network

The polluted speech signal by additive noise is input signal of the trained WNN. The noisy signal polluted the original speech signal is related with the noise training WNN, but has nothing to do with the original signal. The expression of the described processing is denoted as

$$\begin{aligned}
 y(k) &= x(k) + n(k) = \sum_{i=0}^N w_i [x(k-i) + n(k-i)] \\
 &= \sum_{i=0}^N w_i x(k-i) + \sum_{i=0}^N w_i n(k-i) \\
 &= x'(k) + n'(k)
 \end{aligned} \tag{5}$$

where  $y(k)$  is the output of WNN,  $x(k)$  is the original speech signal,  $n(k)$  is the additive noise signal,  $w_i$  is the weights of the trained WNN,  $n(k)$  is noise signal related with the training noise signal,  $x'(k)$  is the output after filtering with the trained WNN,  $n'(k)$  is denoted as the noise output passed from the trained WNN. In general,  $n'(k) \rightarrow 0$ .

## Experiment Results and Discussion

We choose bior4.4 wavelet function to extract the weak speech signal in the noisy environment. By using scaling function and wavelet function as the excitation function of wavelet neural network, it do adaptive filter in the corresponding sub-band and finally take error as the system output to inverse wavelet transform. In this paper, we compared wavelet neural network with BP network on the performance of speech enhancement.

We use the initial 20,000 sampling point of voice signal and noise signal as the training set. At the same and set the error of training goal parameter 0.01. Wavelet neural network are designed as 10 scale decomposition. The number of the hidden layer of BP network is 10, the activity function of which is “tansig”. Its activity function of the output layer is “purelin”. After 32385 times' studying, the error of convergence goal parameters of BP is 0.00999787. The various parameters of wavelet neural network after convergence is shown in Table1.

We can see that under the same training objective parameter approximation error conditions, the training time of wavelet neural network is much shorter than BP network. After wavelet neural network and the BP neural network trained well, we can use them for filtering. The signals of being detected by two network are shown in Fig3.

Tabel.1 Parameters of wavelet neural network

order	Learning number	Learning time	Converge weights	Converge thesholds	Converge errors
1	742	5.8280	4.1473	-0.4105	0.00999783
2	735	5.5000	2.7622	0.9177	0.00999003
3	825	5.5630	6.8021	-0.5655	0.00999418

4	777	6.1560	4.7032	-0.7053	0.00998794
5	770	5.4680	13.4465	-0.2848	0.00996786
6	827	5.5940	-13.4852	0.2596	0.00999430
7	729	4.9220	9.0757	0.4405	0.00995316
8	835	5.5620	2.7310	-0.2079	0.00997580
9	811	5.6410	7.1821	0.3921	0.00998392
10	828	5.5160	-7.7904	0.3492	0.00999930

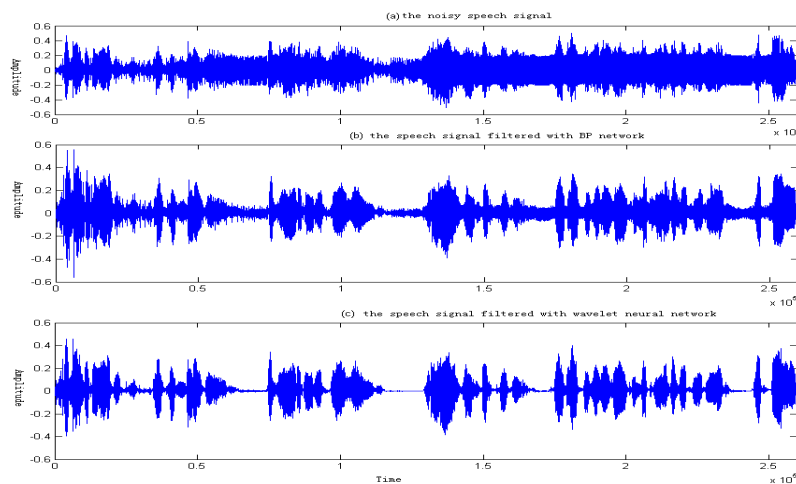


Fig. 3. (a)noisy speech (b)filtered by BP(c) filtered by WNN

## Conclusion

In this paper we proposed a sub-band adaptation filtering algorithm based on Wavelet Transform to deal with the problem of speech enhancement in strong noisy environment. Our passage method depends on the theory of wavelet transform and adaptive filtering. The experimental results show that the method can greatly enhance the speech signal and acquire a good result. Comparing with BP network, the proposed WNN converge faster and more stable.

We gratefully acknowledge the support from the Fundamental Research Funds for the Central Universities (No. N100304008), and also thank all the colleagues for their contributions to experiments and data analysis.

## References

- [1]HacıTasmaza,Ergun,Ercelebi., speech enhancement based on undecimated wavelet-perceptual filterbanks and MMSE-STSA estimation invarious noise environments[J],Digital Signal Processing, 2008.18:797–812
- [2]Mohammed Bahoura, Jean Rouat, Wavelet speech enhancement based on time-scale adaptation[J], Speech Communication,2006.48(12):1620-1637.
- [3]Xu J,Ho DWC. A basis selection algorithm for wavelet neural Networks[J]. Neurocomputing 2002;48:681–689.
- [4]Subasi A, Alkan A, Koklukaya E, Kiymik MK. Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing[J],Neural Networks 2005;18(7):985–97
- [5]Xuegin Zhao, Jinaming Lu, Putranto, W.P.A.,Yahagi, T.Nonlinear time series prediction using wavelet networks with Kalman filter based algorithm[C], IEEE conference on Industrial technology,2005,1226-1230.
- [6]Wei Zhang, Xu Wang, Linlin Ge et al. Simulation of weak signal detection based on wavelet analysis[C],The Proceedings of the 11th International Conference on Industrial Engineering and Engineering Management(IEE),2005-04:652-656.
- [7] Qinghua Zhang, Alber Benvensite. Wavelet Networks[J],IEEE transactions on neural networks, 1992.3(6):889-899