Contemporary Engineering Sciences, Vol. 1, 2008, no. 2, 81 - 90

Enhancement of Nonlinear Spectral Subtraction Method with Applying LPC Analysis

Ehsan Nadernejad and Ali Behrad

Department of Computer Engineering, Faculty of Engineering
Mazandaran Institute of Technology
P.O. Box: 744, Babol, Iran
ehsan_nader@yahoo.com, ali_behrad761@yahoo.com

Abstract

The spectral subtraction is one of the best methods for elimination of stationary noise from degraded speech signal. Here we turn to research about the nonlinear spectral subtraction standard method and improved model. After studying the nonlinear method we turn to this matter that whether it can improve the quality of enhanced speech signal or not, with applying the LPC analysis.

Keywords: signal processing, spectral subtraction, LPC analysis

I. INTRODUCTION

Using a suitable method is a very important goal to elimination of stationary noise from degraded speech signal and enhancing its quality in speech communication systems. Therefore, many methods are assessed such as linear spectral subtraction (LPC analysis) and nonlinear spectral subtraction.

Many researchers assumed the linear and nonlinear spectral subtraction methods alone and achieved desirable results [3],[4],[5] .

In this paper, we want be predict if combination of nonlinear and LPC method, will enhanced the results of nonlinear method or not.

II. Nonlinear spectral subtraction

The basic of new nonlinear spectral subtraction method is a combination of two important theories [1].

- The noise improvement model is used where obtains in the course of a speech paused.

- The nonlinear subtraction is used when a frequency dependent signal to noise ratio (SNR) is obtained. This means that in spectral subtraction a minimal factor is used in high SNR in turn.

III. The standard nonlinear spectral subtraction

In standard method, the spectral subtraction is doing in the following equations and eliminate a standard short-term noise spectral from degraded speech signal that in this condition [1],[2]:

$$Y_i(\omega) = H_i(\omega).X_i(\omega) \quad (1)$$

So in this equation:

$$H_i(\omega) = \frac{D_i(\omega)}{\left|\hat{X}_i(\omega)\right|} \quad (2)$$

$$D_i(\omega) = \left|\hat{X}_i(\omega)\right| - \left|\hat{B}_i(\omega)\right| \quad (3)$$

In above equations $y_i(\omega)$ is estimation of enhanced speech signal. In frame i, $x_i(\omega)$ is estimation of short-term spectral speech in frame i, $|\hat{B}_i(\omega)|$ is estimation of the noise magnitude in frame i, $|\hat{x}_i(\omega)|$ is estimation of the degraded speech signal magnitude in frame *i* and $\hat{B}_i(\omega)$ is estimation of short-term noise in frame *i*.

We can compute the magnitude of $|\hat{x}_i(\omega)|$ and $|\hat{B}_i(\omega)|$ with these equations:

$$\left|\hat{B}_i(\omega)\right| = \lambda_{\hat{B}} \left|\hat{B}_{i-1}(\omega)\right| + (1-\lambda_B)\left|B_i(\omega)\right| \quad (4)$$

$$\left|\hat{X}_i(\omega)\right| = \lambda_{\hat{X}} \left|\hat{X}_{i-1}(\omega)\right| + (1-\lambda_X)\left|X_i(\omega)\right| \quad (5)$$

So the quantity of λ_x and λ_B with implementation determines in this interval:

$$0.1 < \lambda_B < 0.3 \quad (6)$$

$$0.3 < \lambda_x < 0.7 \quad (7)$$

So in this interval the best quantities are $\lambda_B = 0.2$ and $\lambda_x = 0.5$.

IV. The improvement nonlinear spectral subtraction method by $\alpha(\omega)$ factor

An improvement of noise model in nonlinear spectral subtraction method determines by $|\hat{B}_i(\omega)|$ and $\alpha(\omega)$. $\alpha(\omega)$ is an overestimated factor that is jointing with frequency and we can estimate it when pauses speech compound with noise [1]. Here we can estimate $\alpha(\omega)$ by this equation:

$$\alpha_i(\omega) = \max\left(\left|\hat{B}_i(\omega)\right|\right) \tag{8}$$

that usually $\alpha_i(\omega)$ is averaging amount for all frames i and is in this like:[1]

$$\alpha_i(\omega) = 1.5|\hat{B}_i(\omega)| \tag{9}$$

Therefore we can use from this equation for nonlinear spectral subtraction:

$$D_i(\omega) = \left|\hat{X}_i(\omega)\right| - \phi(\rho_i(\omega), \alpha_i(w), \left|\hat{B}_i(\omega)\right|) \tag{10}$$

That in the (10) equation amount of signal to noise ratio $\rho_i(\omega)$ is determined by this equation:

$$\rho_i(\omega) = \frac{\left|\hat{X}_i(\omega)\right|}{\left|\hat{B}_i(\omega)\right|} \tag{11}$$

ϕ is a nonlinear function that it determines the magnitude of spectral in basic amount of signal to noise ratio $\rho_i(\omega)$ and its amount is in this interval :

$$\left|\hat{B}_i(\omega)\right| \le \phi(\rho_i(\omega), \alpha_i(\omega), \left|\hat{B}_i(\omega)\right|) \le 3\left|\hat{B}_i(\omega)\right| \tag{12}$$

and amount of it determine by using of this equation and depended to the $\rho_i(\omega)$ factor:

$$\phi\left(\rho_i(\omega), \alpha_i(\omega), \left|\hat{B}_i(\omega)\right|\right) = \frac{\alpha_i(\omega)}{1 + \gamma\, \rho_i(\omega)} \tag{13}$$

In this equation γ is an important factor that is depended on variation range of SNR $\rho_i(\omega)$.

V. The using of linear prediction coding (LPC) analysis for improvement of nonlinear method

This method in addition to decreasing the musical noise cause to increase the quality of enhanced speech signal. This method is included in two separate parts:

1. noise estimation part
2. spectral subtraction part

5-1: Noise estimation part

In common spectral noise is estimated from pause frames, but in this method is used from LPC analysis for estimation of noise. This approach includes applying the LPC analysis on the degraded speech signal; we calculate the number of coefficients named LP coefficients. In fact these coefficients include specifications of speech signal, because in calculating the LP coefficients we use Auto correlation function and just as we know the auto correlation function applies the act of averaging on the signal. Now if these signals to be degraded speech signal, noise eliminated nearly with the act of averaging. For this reason we can say that LP coefficients only are includes the properties of speech signal. Now, we use from this coefficients for generation of inverse LPC filter:

$$A(z) = 1 - \sum_{k=1}^{p} a_k z^{-k} \tag{14}$$

and with applying the degraded speech signal to it, we see that output signal is very similar to noise signal and we suppose it as estimation of noise in the degraded speech.[3]

5.2 Spectral subtraction part

In this part estimated noise in prior part, eliminated by spectral subtraction from speech signal and we will see the enhanced speech signal in output.

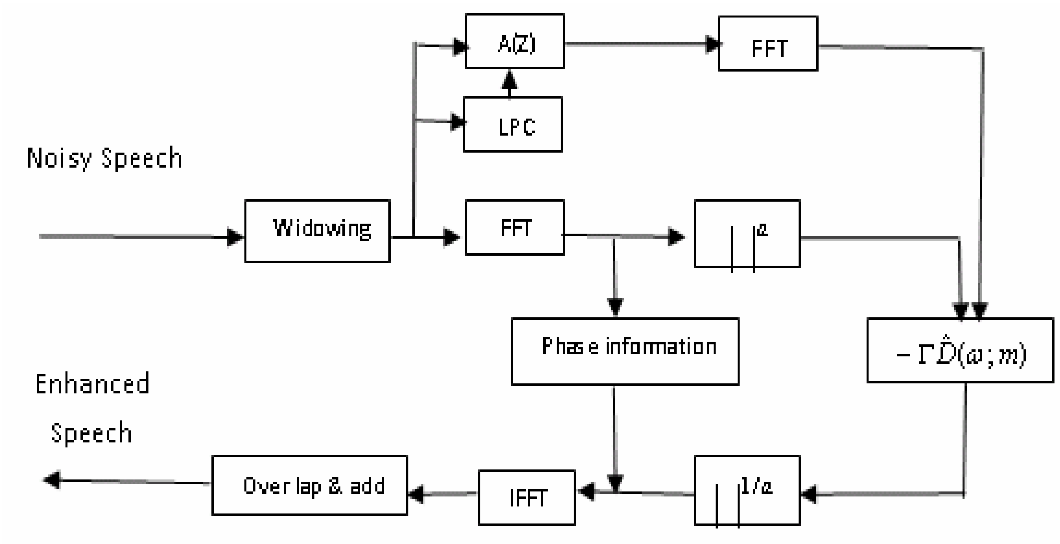


Figure 1 : Block diagram of LPSS method

Implementation

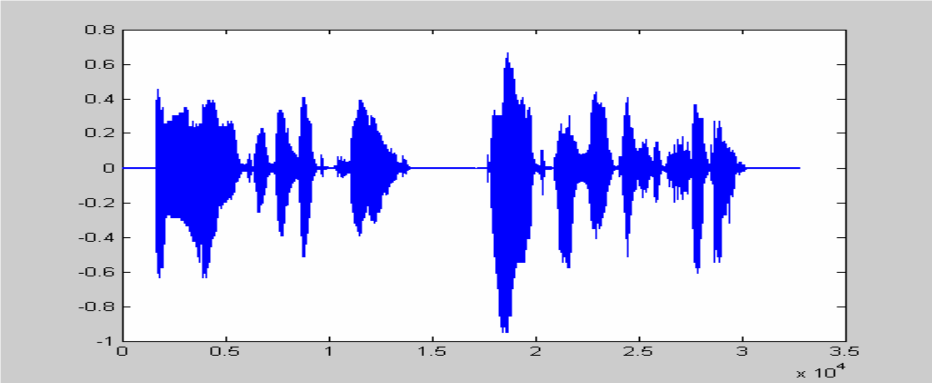
In this part we will study the said methods. We operate in this form, first select the clean speech signal from *TIMITS* database, and then we produce by using white Gaussian noise generated by the degraded speech signal. After generating the degraded speech signal we apply them to the standard and optimized by α factor nonlinear systems as input. Then we compare the generated outputs for the purpose of signal to noise ratio (SNR), transient state graph and spectrogram.
In the next step we apply the generated outputs of prior nonlinear methods as input of LPC system. Then we compare the generated output with the results of prior part.

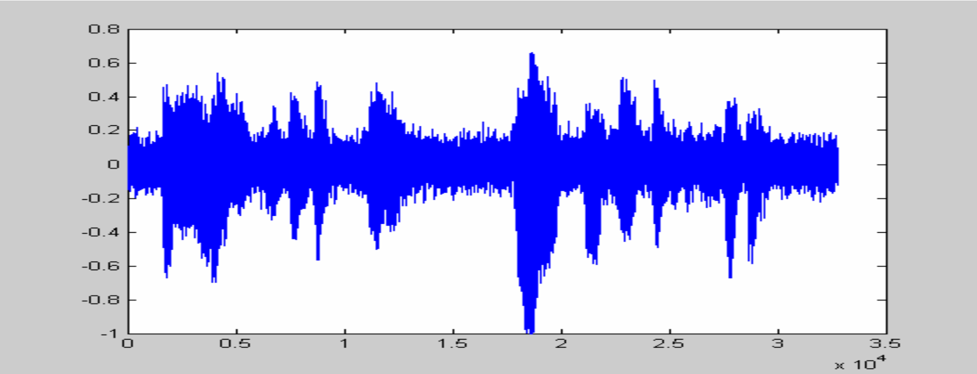
Table 1: Test results without LPC

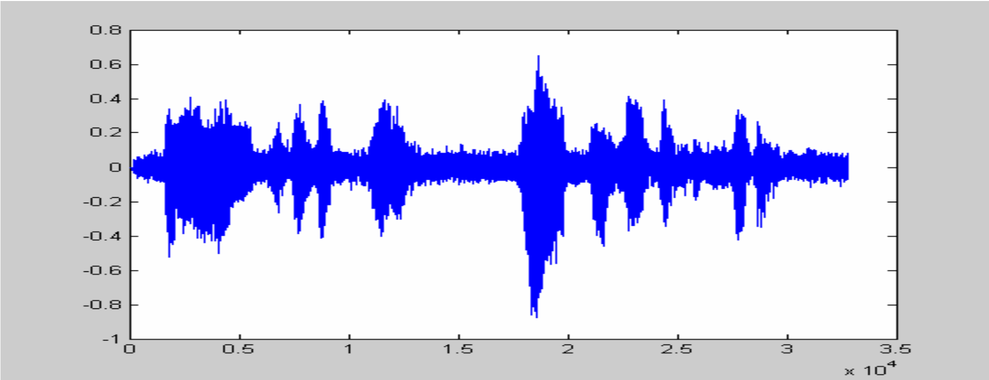
Measurement type: signal to noise ratio (db) Clean Signal: sampled clean speech signal from *TIMIT* database Added noise to clean signal : White Gaussian noise .			
Input SNR / Enhancement Method	5 db	10 db	15 db
NSS	7.5525	11.9915	15.6749
NSS With alpha	1.6459	12.3650	16.4670

Table 2: Test results with LPC

Measurement type : signal to noise ratio (db) Clean signal :sampled clean speech signal from *TIMIT* database Added noise to clean signal : White Gaussian noise			
Input SNR / Enhancement Method	5 db	10 db	15 db
NSS	7.7897	12.1325	16.1436
NSS With alpha	8.4943	12.5360	16.9756


0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
0.5
1
1.5
2
2.5
3
3.5
x 10^4


0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
0.5
1
1.5
2
2.5
3
3.5
x 10^4


0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
0.5
1
1.5
2
2.5
3
3.5
x 10^4

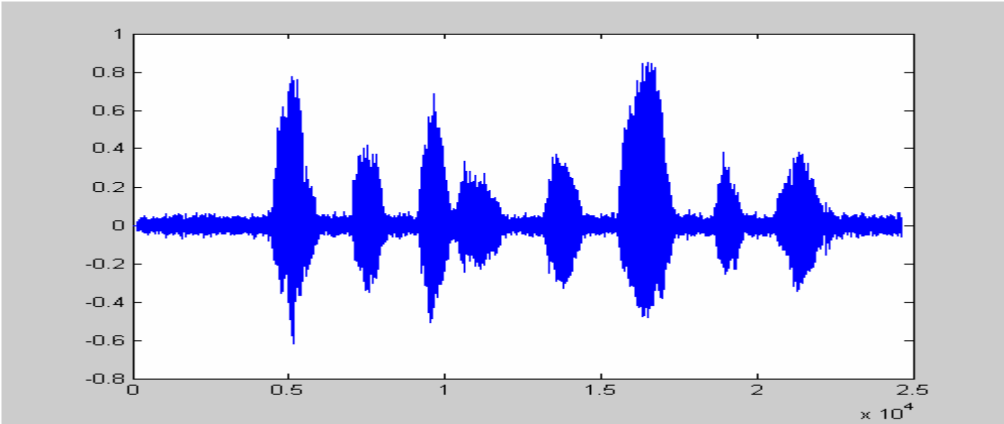
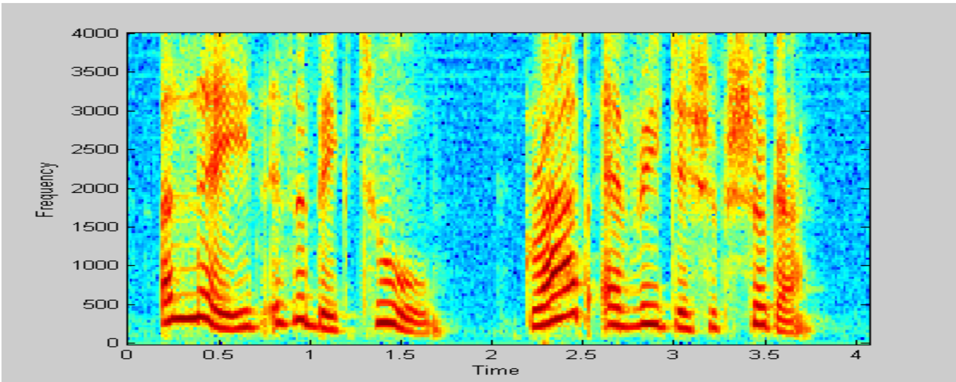
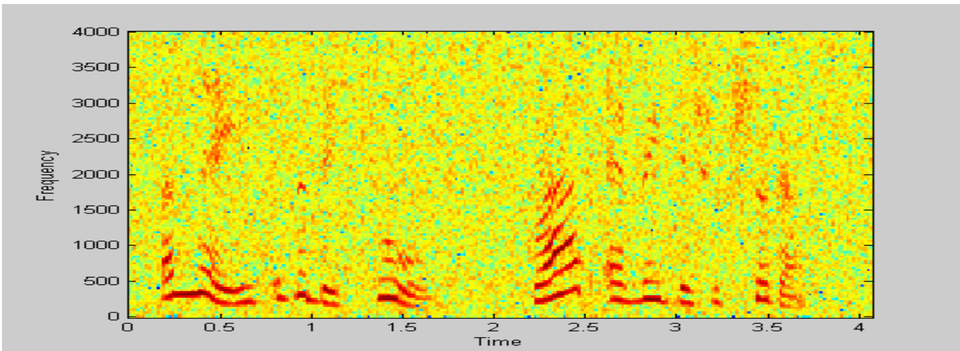


Figure 2: Transient state graph a) clean speech signal b) degraded speech signal (SNR=5db) c) Enhanced speech signal with NSS method d) Enhanced speech signal with LPNSS method





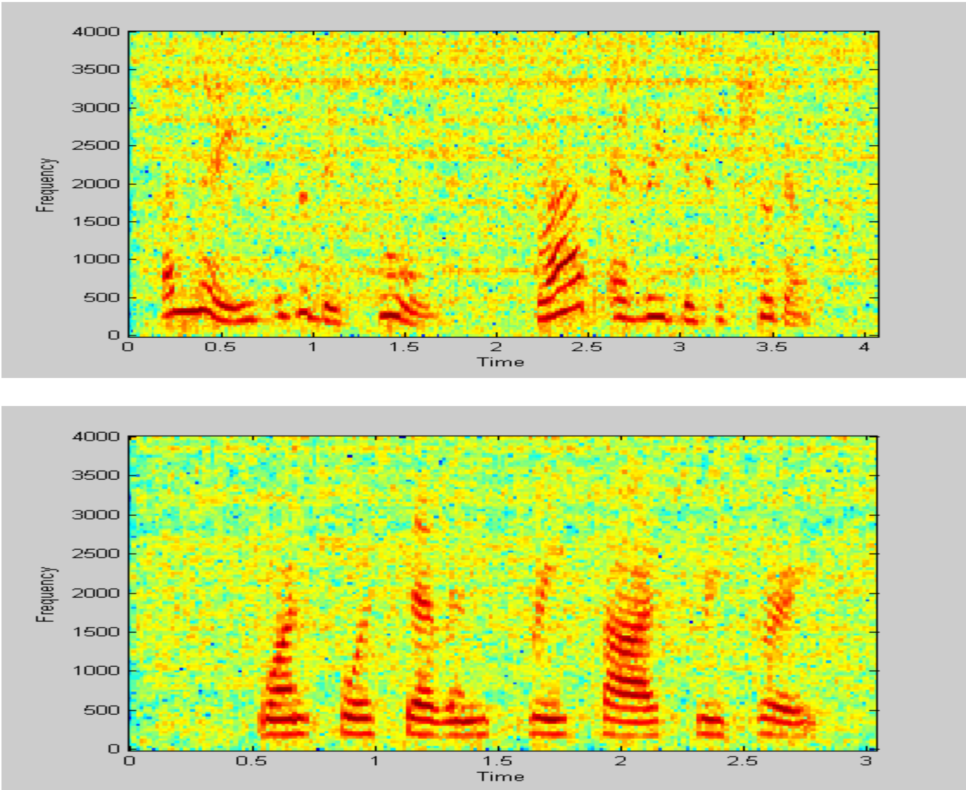


Figure 3: Spectrogram a) clean speech signal b) degraded speech signal (SNR=5db) c) Enhanced speech signal with NSS method d) Enhanced speech signal with LPNSS method

Conclusion

With paying attention to the SNR quantities and generated transient state graph and spectrogram in part (5) we will see that improvement of nonlinear model has better results relative to the standard state.
With applying the LPC analysis to each of the prior methods you will see that the results of each of them have improved.

Therefore, we can use LPC analysis as an effective method for improvement of nonlinear spectral subtraction methods.

References

[1] P. LOCKWOOD , J. BOUDY, "Experiments With a Nonlinear Spectral Subtractor (NSS), Hidden Markov Models and projection, For Robust Recognition in car" *Speech communication*, vol . 11 , pp:215-228 , June 1992 .

[2] J. PORUBA , "Compartion of Speech Detection Algorithm" . 24th International Conference Telecommunications and Signal Processing , Brno 2001.

[3] J.Tierney , "A study of LPC analysis of speech in additive noise" , IEEE trans. Acoust. Speech and signal process., ASSP-28,4,PP.379- 389(Aug.1980) .

[4] M.R Sambur, and N.S Jayant., "LPC analysis/synthesis from speech inputs containing quantizing noise or additive white noise" , IEEE Trans . Acoust . Speech and signal process ., ASSP-24, 6,pp.488-494(Dec.1976) .

[5] JIRI PORUBA , " *S*peech Eenhancement Based On Nonlinear Spectral Subtraction " , Fourth IEEE International Caracas Conference on Devices , Circuits and Systems , Aruba , April 17-19 , 2002 .

Received: February 19, 2008