ADAPTIVE BLIND NOISE SUPPRESSION IN SOME SPEECH PROCESSING APPLICATIONS

Georgi Iliev and Nikola Kasabov
Department of Information Science, University of Otago
P.O. Box 56, Dunedin, New Zealand
giliev@infoscience.otago.ac.nz, nkasabov@otago.ac.nz

Abstract – **In many applications of speech processing the noise reveals some specific features. Although the noise could be quite broadband, there are a limited number of dominant frequencies, which carry the most of its energy. This fact implies the usage of narrow-band notch filters that must be adaptive in order to track the changes in noise characteristics. In present contribution, a method and a system for noise suppression are developed. The method uses adaptive notch filters based on second-order Gray-Markel lattice structure. The main advantages of the proposed system are that it has very low computational complexity, is stable in the process of adaptation, and has a short time of adaptation. Under comparable SNR improvement, the proposed method adjusts only 3 coefficients against 250-450 for the conventional adaptive noise cancellation systems. A framework for a speech recognition system that uses the proposed method is suggested.**

I. INTRODUCTION

The noise existence is inevitable in real applications of speech processing. It is well known that the additive noise affects negatively the performance of the speech codecs designed to work with noise-free speech [1],[2],[3], especially codecs based on linear prediction coefficients (LPC). Another application strongly influenced by noise is related to the hands-free phones where the background noise reduces the signal to noise ratio (S/N) and the speech intelligibility [4],[5]. Last but not least, is the problem of speech recognition in a noisy environment. A system that works well in noise-free conditions, usually shows considerable degradation in performance when background noise is present [6].

It is clear that a strong demand for reliable noise cancellation methods exists that efficiently separate the noise from speech signal. The endeavors in designing of such systems can be followed some 20 years ago [7].

The core of the problem is that in most situations the characteristics of the noise are not known a priori and moreover they may change in time. This implies the use of adaptive systems capable of identifying and tracking the noise characteristics. This is why the application of adaptive filtering for noise cancellation is widely used [7],[8],[9],[10].

The classical systems for noise suppression rely on the usage of adaptive linear filtering and the application of digital filters with finite impulse response (FIR). The strong points of this approach are the simple analysis of the linear systems in the process of adaptation and the guaranteed stability of FIR structures. It is worth mentioning the existence of relatively simple and well investigated adaptive algorithms for such kind of systems as least mean squares (LMS) and recursive least squares (RLS) algorithms [7],[10].

The investigations in the area of noise cancellation reveal that in some applications the nonlinear filters outperform their linear counterparts. That fact is a good motivation for a shift towards the usage of nonlinear systems in noise reduction [11],[12],[13].

Another approach is based on a microphone array instead of the two microphones, reference and primary, that are used in the classical noise cancellation scheme [6].

A brief analysis of all mentioned approaches leads to the conclusion that they try to model the noise path either by a linear or by a nonlinear system. Each of these methods has its strengths and weaknesses. For example, for the classical noise cancellation [1] with two microphones this is the need of reference signal; for the neural filters - the fact that as a rule they are slower than classic adaptive filters and they are efficient only for noise suppression on relatively short data sequences [14], which is not true for speech processing and finally for microphone arrays - the need of precise space alignment [6].

In present contribution the approach is slightly different. The basic idea is that in many applications, for instance, hands-free cellular phones in car environment [4], howling control in hands-free phones, noise reduction in an office environment, the noise reveals specific features that can be exploited. In most instances although the noise might be quite wide-band, there are always, as a rule, no more than two or three regions of its frequency spectrum that carry most of the noise energy and the removal of these dominant frequencies results in a considerable improvement of S/N ratio. This brings the idea to use notch adaptive filters capable of tracking the noise characteristics. In this paper a modification of all-pass structures is used [15]. They are recursive, and at the same time, are stable during the adaptive process. The approach is called “blind” because there is no need of a reference signal.

II. CLASSICAL ADAPTIVE NOISE CANCELLATION

One of the most wide spread applications of adaptive filtering is adaptive noise cancellation. Fig. 1 shows the popular scheme for adaptive noise cancellation using digital FIR filter.

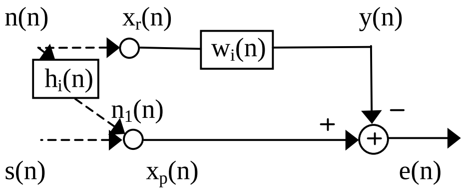

Fig. 1. Adaptive noise cancellation.

The basic prerequisite for this realization is the availability of the two inputs called primary and reference. The primary signal consists of speech s(n) plus noise $n_1(n)$ while the reference signal consists of noise n(n) alone. The two noise signals $n_1(n)$ and n(n) are correlated and $h_i(n)$ is the impulse response of the noise path from the noise source to the primary microphone. Assuming that the signals are discrete-time and the sampling period is T=1, the primary input can be written as

$$x_p(n) = s(n) + n_1(n) \tag{1}$$

where speech signal s(n) and noise signal $n_1(n)$ must be uncorrelated.

Going trough the scheme of Fig. 1 and all mentioned above it is clear that here noise cancellation is simply the joint process estimation problem. The system is to reduce the effect of the noise in the primary input using the correlation between the two noise signals n(n) and $n_1(n)$. This can be implemented by minimizing the mean-square error $E[e^2(n)]$ where

$$e(n) = x_p(n) - y(n). \tag{2}$$

In (2) y(n) is the output signal of the adaptive filter

$$y(n) = \sum_{i=0}^{N} w_i(n)\, x_r(n) \tag{3}$$

where N is the filter order and $w_i(n)$ is the ith coefficient of the adaptive filter.

Having in mind that

$$E[e^2(n)] = E[s^2(n)] + E[n_1(n) - y(n)]^2 \tag{4}$$

the minimization of $E[e^2(n)]$ is equivalent to the minimization of the difference between the signal on the adaptive filter output y(n) and the noise signal $n_1(n)$ present on the primary input. Obviously the better replica of $n_1(n)$ y(n) is, that is, the better the adaptive filter is modeling the impulse response $h_i(n)$ of the noise path, the smaller the difference.

The minimization of $E[e^2(n)]$ can be achieved by updating of the adaptive filter coefficients. Most often the LMS and NLMS algorithms are used, the latter having the advantage that the step size is relatively independent of the amplitude of the input signal. According to the scheme in Fig. 1 the updating equations for LMS and NLMS algorithms are given by

LMS:

$$w_i(n+1) = w_i(n) + \mu e(n) x_r(n) \tag{5}$$

NLMS:

$$w_i(n+1) = w_i(n) + \mu \frac{e(n)\, x_r(n)}{\sum_{i=0}^{N} x_r^2(n-i)} \tag{6}$$

$$0 \le i \le N$$

where μ is a step size.

III. ADAPTIVE BLIND NOISE SUPPRESSION (ABNS) SCHEME

As mentioned in the introduction, the specific features of the noise in some speech processing applications suggest the usage of narrow-band notch filters. They have to meet the following requirements:

- to adapt as fast as possible to the changes in the noise which might be quite rapid, for example car engine noise;
- the cancelled portions of the spectrum should be as narrow as possible in order to prevent speech signal distortions.

Both requirements could be met much easier using IIR adaptive filters instead of FIR adaptive filters. IIR filters are usually avoided because they create a lot of stability problems. To overcome this problem we use a realization based on second order Gray-Markel lattice circuit [15] - Fig.2. Using this circuit it becomes possible to implement a second order notch/bandpass section [16] - Fig. 3.

What are the advantages of such a realization? First, it has extremely low pass band sensitivity that means resistance to quantization effects. Second, it is very convenient for realization of adaptive notch filters because it is possible to control independently the notch frequency and the bandwidth.

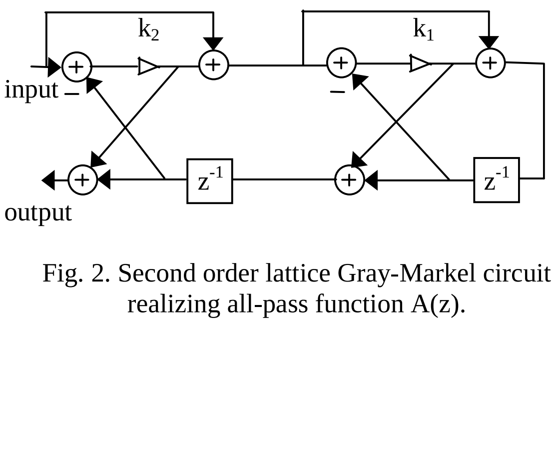

Fig. 2. Second order lattice Gray-Markel circuit realizing all-pass function A(z).

Thus if the all-pass function A(z) is

$$A(z) = \frac{k_2 + k_1(1+k_2)z^{-1} + z^{-2}}{1 + k_1(1+k_2)z^{-1} + k_2 z^{-2}} \quad (7)$$

then k_1 controls the notch frequency ω_0 while k_2 is related to the bandwidth BW via

$$k_1 = -\cos\omega_0 \quad (8)$$

$$k_2 = \frac{1-\tan(BW/2)}{1+\tan(BW/2)}. \quad (9)$$

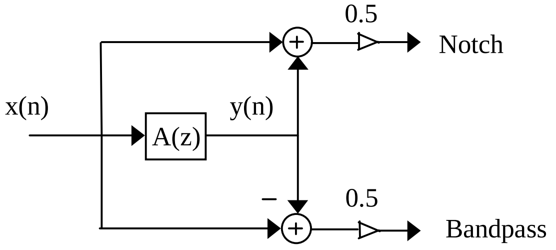

Fig. 3. Second order notch/bandpass section.

But, on the other hand, BW is directly connected to the distance from the pole to the unity-circle. So if we use the structure of Fig. 3 as an adaptive filter we may fix BW and thus fixing k_2 we make constant the distance from the pole to the unity-circle which means that with this constraint we obtain an adaptive IIR filter free of stability problems. Adapting k_1 we may shift the notch frequency around the unity-circle.

Using the basic structure of Fig. 3 and the constraint mentioned above, the final arrangement of our system is shown in Fig. 4. The system will work in the following manner: each section will remove one of the dominant frequencies using an appropriate adaptive algorithm. As shown in Fig. 4 we propose to update only the coefficients k_{11}, k_{12},…, k_{1M}, while k_2 is a priori determined from equation (9). Thus we can reduce considerably the number of computations and can guarantee the stability of the adaptive structure. The number of sections is determined by the application. Here we introduce the NLMS algorithm for adjusting the filter coefficients as

$$e_i(n) = 0.5[e_{i-1}(n) + y_i(n)] \quad (10)$$

for i = 1 to M and $e_0(n) = x(n)$

$$k_{1i}(n+1) = k_{1i}(n) - \mu \frac{e_i(n)\, y'_i(n)}{[y'_i(n)]^2} \quad (11)$$

$$y'_i(n) = \frac{d\, y_i(n)}{d\, k_{1i}(n)}$$

for $1 \le i \le M$

where M is the number of sections, $e_i(n)$ is the error signal, μ is the step size and $y'_i(n)$ is the derivative of $y_i(n)$ with respect to the coefficient subject of adaptation.

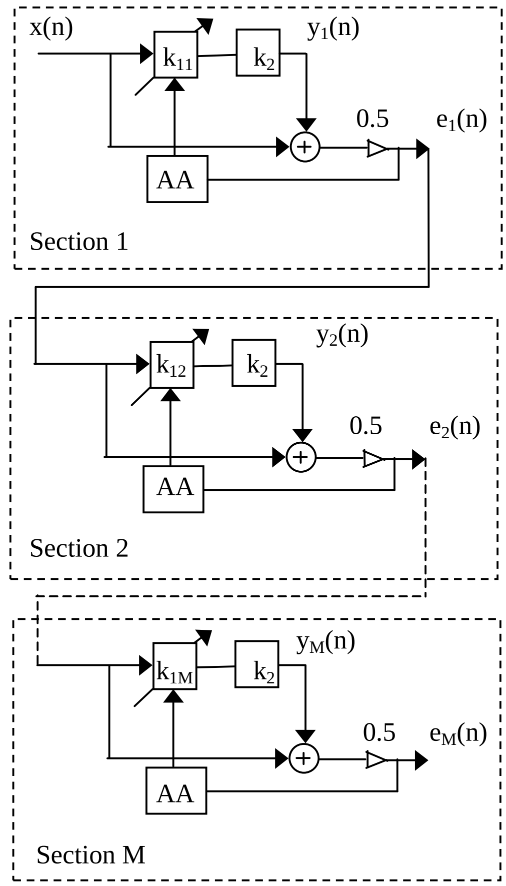

Fig. 4. Adaptive system for noise suppression.

IV. TEST RESULTS

The performance of the ABNS method for noise suppression is assessed by computer simulations. Fig. 5 shows the original speech. The speech is corrupted with noise from computer cooling fan that is most often encountered in office environment and the resultant signal is depicted in Fig. 6. The process of noise suppression is shown in Fig .7. Here the system is composed of 3 sections each of them adapting its coefficient to one of the dominant frequencies in the noise spectrum. Fig.8 presents the trajectories of the filter coefficients. In this experiment the capability of the system to track the changes in noise signal is tested as the dominant frequencies shift from 0.1, 0.2 and 0.4 at the beginning, to 0.14, 0.23 and 0.36. Here the system does not have information about the dominant frequencies and adjusts its coefficients to them, as it works.

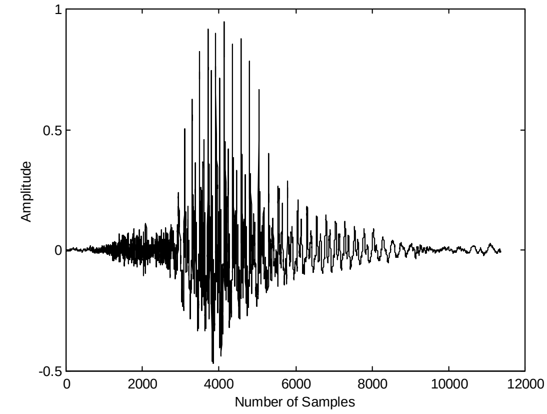

Fig. 5a. Original speech - the word “home”.

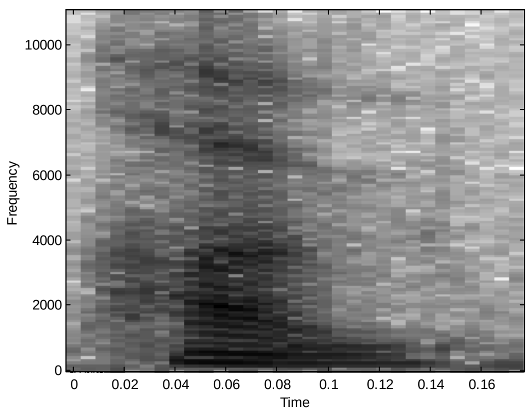

Fig. 5b. Spectrogram - the word “home”.

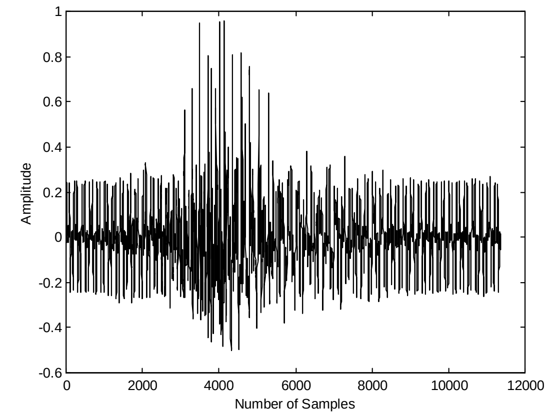

Fig. 6. Noise-contaminated speech.

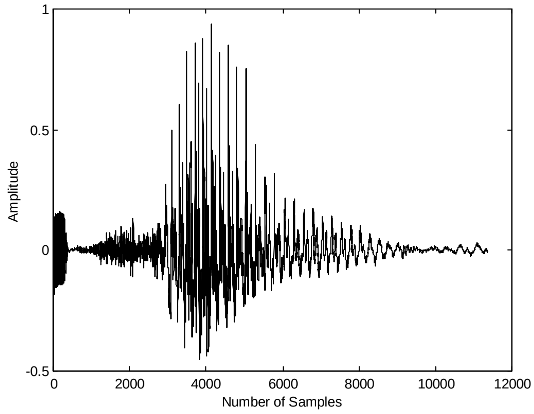

Fig. 7. Speech after noise suppression.

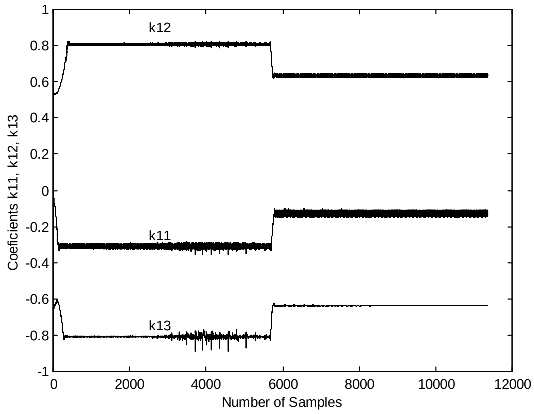

Fig. 8. Trajectories of the filter coefficients.

Table 1 shows the improvement in SNR as a result of the application of the proposed system. The obtained results are comparable to these of the conventional adaptive noise canceller (ANC), for example these reported in [1]. The proposed system is much faster and simpler to implement.

Table 1. SNR before and after noise suppression.

ABNS	Before	After	SNR Gain
SNR (dB)	0	11.9	11.9
SNR (dB)	3.5	14.3	10.8
ANC	Before	After	SNR Gain
SNR (dB)	0	10	10

V. A FRAMEWORK FOR A SPEECH RECOGNITION SYSTEM USING ABNS

A block diagram of a speech recognition system is given in Fig. 9. It consists of the following modules:

adaptive blind noise suppression (ABNS), endpoint detection (EPD), acoustic feature extraction (AFE), feature normalization (FN) and speech recognition module (SRM).

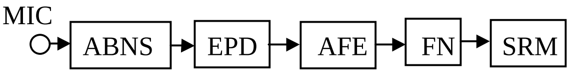

Fig. 9. Block diagram of a speech recognition system using ABNS.

Short-time energy and zero-crossing rate are combined to detect the speech utterance boundaries. Acoustic features of the input speech are extracted over 20 ms frames. Hamming windows having an overlap of 10 ms are used to calculate Mel Frequency Scale Cepstral Coefficients and log-energy. Here the speech recognizer can be implemented on the base of adaptive evolving fuzzy neural networks (EFuNNs) [17]. Since the input layer of the networks has fixed size, while the segments (words) are made up of a variable number of frames, a technique for normalization is needed. A discrete cosine transform (DCT) is applied to the whole segment, retaining as many parameters as it is necessary. Several application-oriented systems for automatic dialing and robot control are under development.

VI. CONCLUSIONS

A very efficient adaptive system based on IIR structures for noise suppression is proposed in this contribution. The main advantages of the present realization are:

- the adaptive system has a short time of adaptation - about 100 iterations;
- the system is very simple and flexible, for comparison, here we adjust only 3 coefficients against 250-450 for conventional adaptive noise cancellation system;
- the second-order lattice structures are stable during the adaptation that defines the high stability of the whole system.

The proposed system for noise suppression may be applied in many situations where the noise reveals the specific features mentioned in the previous sections and the application of this system could considerably improve the speech intelligibility.

A demonstration program is available on the WWW from: "http://divcom.otago.ac.nz/infosci/KEL/CBIIS.html".

REFERENCES

[1] S. Ikeda, A. Sugiyama, "An adaptive noise canceller with low signal distortion for speech codecs," IEEE Trans. Signal Processing, vol. 47, pp. 665-674, Mar. 1999.
[2] J. Gibson, B. Koo, S. Gray, "Filtering of colored noise for speech enhancement and coding," IEEE Trans. Signal Processing, vol. 39, pp. 1732-1742, Aug. 1991.
[3] W. Harrison, J. Lim, E. Singer, "A new application of adaptive noise cancellation," IEEE Trans. Acoust., Speech, Signal Processing, vol. 34, pp. 21-27, Jan. 1986.
[4] S. Kuo, H. Chuang, P. Mallela, "Integrated automotive signal processing and audio system," IEEE Trans. Consumer Electronics, pp. 522-531, Aug. 1993.
[5] G. Mirchandani, R. Zinser, J. Evans, "A new adaptive noise cancellation scheme in the presence of crosstalk," IEEE Trans. Circuits Syst., vol. 39, pp. 681-694, 1992.
[6] D. Van Compernolle, W. Ma, F. Xie, M. Van Diest, "Speech recognition in noisy environments with the aid of microphone arrays," Speech Communication, vol. 9, pp. 433-442, 1990.
[7] B. Widrow, S. Stearns, Adaptive Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[8] J. Candy, Signal Processing: The Modern Approach. New York: McGraw-Hill, 1988.
[9] K. Astrom, G. Goodwin, P. Kumar, Adaptive Control, Filtering, and Signal Processing. New York: Springer-Verlag, 1995.
[10] G. Goodwin, K. Sin, Adaptive Filtering, Prediction and Control. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[11] L. Yin, J. Astola, Y. Neuvo, "A new class of nonlinear filters-neural filters," IEEE Trans. Signal Processing, vol. 41, pp. 1201-1222, Mar. 1993.
[12] W. Knecht, M. Schenkel, G. Moschytz, "Neural network filters for speech enhancement," IEEE Trans. Speech and Audio Processing, vol. 3, pp. 433-438, Nov. 1995.
[13] J. Lin, T. Sellke, E. Coyle, "Adaptive stack filtering under the mean absolute error criterion," IEEE Trans. Acoust., Speech, Signal Processing, vol. 38, pp. 934-954, June 1990.
[14] K. Mehrotra, C. Mohan, S. Ranka, Elements of Artificial Neural Networks. MA: MIT Press, 1997.
[15] P. Regalia, S. Mitra, P. Vaidyanathan, "The all-pass filter: a versatile signal processing building block," Proc. IEEE, pp. 19-37, Jan. 1988.
[16] P. Regalia, "An improved lattice-based adaptive IIR notch filter," IEEE Trans. Signal Processing, pp. 2124-2128, Sep. 1991.
[17] N. Kasabov, "Evolving connectionist and fuzzy-connectionist systems: methods, tools, applications," in Neuro-Fuzzy Techniques for Intelligent Information Processing, N. Kasabov and R. Kozma, eds., Heidelberg, Physica Verlag 1999.