Proceedings of ITC-CSCC 2000, Pusan, Korea

A New Noise Reduction Method Based on Linear Prediction

Arata Kawamura†, Kensaku Fujii††, Yoshio Itho†, Yutaka Fukui†

†Faculty of Engineering, Tottori University
4-101 Koyama-Minami, Tottori 680-8552 JAPAN
TEL/FAX:0857-31-5690
E-MAIL:kawamura@dacom2.ele.tottori-u.ac.jp

††FUJITSU LABORATORIES LTD
Wireless Communication Systems lab. (L41)
4-1-1 Kamikodanaka, Nakahara-ku, Kawasaki-shi, 211-8588

Abstract: A technique that uses linear prediction to achieve noise reduction in a voice signal which has been mixed with an ambient noise (Signal to Noise (S-N) ratio = about 0dB) is proposed. This noise reduction method which is based on the linear prediction estimates the voice spectrum while ignoring the spectrum of the noise. The performance of the noise reduction method is first examined using the transversal linear predictor filter. However, with this method there is deterioration in the tone quality of the predicted voice due to the low level of the S-N ratio. An additional processing circuit is then proposed so as to adjust the noise reduction circuit with an aim of improving the problem of tone deterioration. Next, we consider a practical application where the effects of round off errors arising from fixed-point computation has to be minimized. This minimization is achieved by using the lattice predictor filter which in comparison to the transversal type, is known to be less sensitive to the round-off error associated with finite word length operations. Finally, we consider a practical application where noise reduction is necessary. In this noise reduction method, both the voice spectrum and the actual noise spectrum are estimated. Noise reduction is achieved by using the linear predictor filter which includes the control of the predictor filter coefficient's update.

1. Introduction

In recent years, research on methods of noise reduction from a voice which has been mixed with noise is actively being done by the use of microphone array[1], spectrum subtraction[2], etc. Imperfection can be seen in the method of the noise reduction using two microphones which can be considered as a directional microphone with a blind spot in the incident direction of noise. When many noise sources exist, an increase in number of microphones cannot be avoided. It is therefore important to develop a noise reduction method which uses a single microphone, and which can cancel multiple noise sources. In the systems with only one microphone, extracting a voice signal which has been mixed with noise requires the use of spectrum subtraction method for noise reduction. One of the spectrum subtraction method [2] improves the S-N ratio at the expense of signal distortion and musical tones that arise due to the residual noise. In order to improve on these negative effects, a new method which involves the knowledge of psychoacoustical phenomena has been proposed [3,4]. However, simulation result of the method shows that the quality of extracted voice is still not satisfactory for low S-N ratio. The problem of distortion of extracted voice is also not solved.

In this paper, we propose a new method of noise reduction that extracts voice signal emerged in white noise (S-N ratio = about 0dB). The proposed method employs one microphone and the theory of linear prediction which is also normally used in speech analysis. This noise reduction method which is based on the linear prediction estimates the voice spectrum while ignoring the spectrum of the white noise. However, with this method there is deterioration in the tone quality of the predicted voice due to the low level of the S-N ratio. An additional processing circuit is then proposed so as to adjust the noise reduction circuit with an aim of improving the problem of tone deterioration. Next, we consider a practical application where the effects of round off errors arising from fixed-point computation has to be minimized. This minimization is achieved by using the lattice predictor filter which in comparison to the transversal type, is known to be less sensitive to the round-off error associated with finite word length operations[5]. Finally, we consider a practical application where noise reduction is necessary. In this noise reduction method, both the voice spectrum and the actual noise spectrum are estimated. Noise reduction is achieved by using the linear predictor filter which includes the control of the predictor filter coefficient's update.

2. Principle of noise reduction method

2.1. Noise Reduction Using Adaptive Transversal Filter

Linear prediction is usually used in speech analysis and synthesis, pre-whitening in system identification, among many areas. In the linear forward prediction, the current input sample is estimated using the weighted sum of the immediate past M samples of the input signal [5]. Figure 1 shows the noise reduction circuit using transversal filter as linear predictor(LP). The prediction error e(n) whose mean square value is to be minimized is defined as e(n)=x(n)-y(n), where y(n) is an estimate of the current input signal x(n). y(n) is given by

$$y(n)=\sum_{k=1}^{M}h_k(n)x(n-k) \tag{1}$$

When the input signal is white, the coefficients of the linear predictor will hardly change from their initial zero value. On the other hand, when the input is a stationary colored signal, the coefficients of the linear predictor will converge such that the error signal becomes white, assuming that the order of the filter is sufficiently large. When the input is a voice, the linear predictor in the noise reduction circuit estimates the voice spectrum. The proposed method removes the noise elements other than voice spectrum by passing the noise through linear predictor.

2.2. The Adaptation Algorithm

In order to take in to account the non-stationary of voice signal the linear predictor must be updated using a fast algorithm. The adaptive algorithm used in updating the coefficient of the linear predictor is the Sub-RLS[6] algorithm which has a high convergence speed for comparatively little number of calculation. The Sub-RLS algorithm is given by

$$\mathbf{h}(n+1)=\mathbf{h}(n)+\beta\,(\mathbf{q}(n)-\mathbf{R}(n)\mathbf{h}(n)) \qquad (2)$$

where, β is the step size of adaptation, $\mathbf{q}(n)$ is the cross-correlation vector of the input signal and predicting signal $y(n)$, $\mathbf{R}(n)$ is auto-correlation matrix of the input signal, and $\mathbf{h}(n)$ is tap coefficient vector at time n. $\mathbf{q}(n)$ and $\mathbf{R}(n)$ are given by

$$\mathbf{q}(n)=[q_1(n),q_2(n),\ldots,q_M(n)]^T \qquad (3)$$

$$\mathbf{R}(n)=\begin{bmatrix} R_{1,1}(n) & R_{1,2}(n) & \cdots & R_{1,M}(n) \\ R_{2,1}(n) & R_{2,2}(n) & \cdots & R_{2,M}(n) \\ \vdots & \vdots & \ddots & \vdots \\ R_{M,1}(n) & \cdots & \cdots & R_{M,M}(n) \end{bmatrix} \qquad (4)$$

where

$$q_m(n)=\alpha\, q_m(n)+y(n)x_m(n) \qquad (5)$$

$$R_{i,m}(n)=\alpha\, R(n-1)+x_i(n)x_m(n) \qquad (6)$$

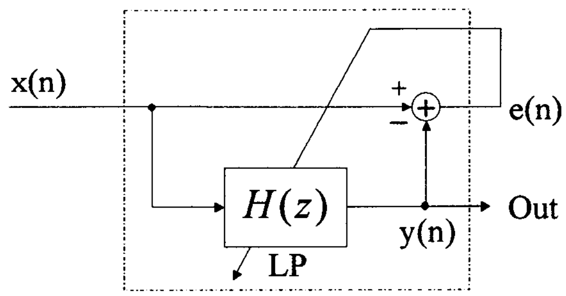
α : forgetting coefficient

When the noise power is very large, the correlation matrix and vector cannot be calculated accurately. Furthermore, due to the presence of the non-predictable portion within the voice signal, it is not possible to completely predict the voice signal. As a result of these, there is a deterioration in the quality of the sound extracted.

2.3. Improvement of Tone Quality

In this section we propose a method which take into account the portion of the voice signal which could not be predicted. Because the voice element which was not able to be predicted is included in the prediction error, tone quality will be improved by adding part of the prediction error to the predicting signal. The re-processing improved the quality of the extracted voice at an expense of a reduction in the S-N ratio. In order to minimize this problem we propose a method which involves cascading the predictor filter and the re-processor as shown in figure 2(where $0 \leqq K < 1$). In this figure both the LPs have the same initial parameter value. The improvement in performance was verified using computer simulation.

Table 1 summarizes the parameters used in this simulation while figure 3 shows the results that were obtained. In this figure, (c) and (d) represent the extracted signal without and with the cascading, respectively. From this result, it can be seen that an improvement in both the S-N ratio and the extracted voice quality was obtained.



LP : Linear Predictor

Figure 1. Noise Reduction Circuit using transversal filter as LP.

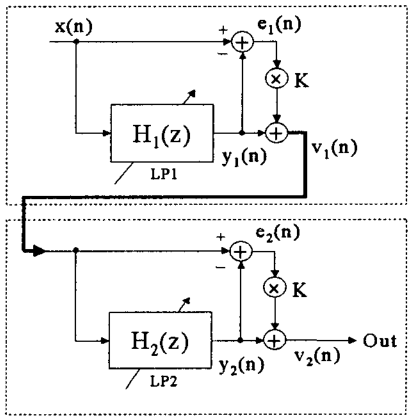


Figure 2. Structure of Cascade Connection.

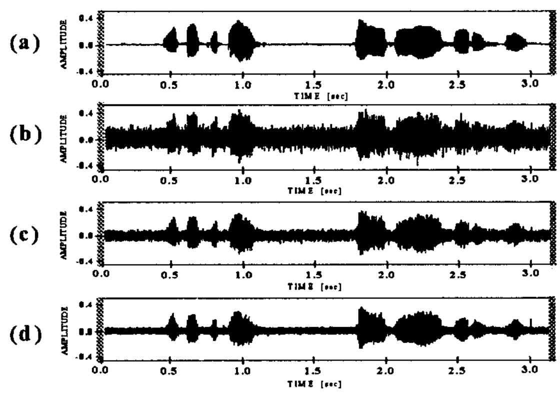


Figure 3. (a) Non-corrupted voice signal.
(b) Voice signal mixed with white noise.
(c) Voice signal extracted before cascading (v_1).
(d) Voice signal extracted after cascading (v_2).

Table 1. Parameters used in the simulation of Fig.5.

S-N ratio of The Input Signal [dB]	0
The Number of Tap Coefficient	64(×2)
Forgetting Coefficient α	0.985
Step Size β	0.01
K	0.5

3. Lattice Composition

Next, we consider a practical application where the effects of round off errors arising from fixed-point computation has to be minimized. This minimization is achieved by using the lattice predictor filter which in comparison to the transversal type, is known to be less sensitive to the round-off error associated with finite word length operations[5]. Figure 4(a) shows the lattice filter, where,

x(n) : input signal at time n
f(n) : forward prediction error at time n
b(n) : backward prediction error at time n
$\gamma^{(f)}(n)$: forward reflection coefficient at time n
$\gamma^{(b)}(n)$: backward reflection coefficient at time n

Figure 4(b) shows the noise reduction circuit using lattice predictor filter. The extracted voice signal is given as x(n)-f_M(n). The adaptive algorithm used in the update of the coefficients is order-update recursion algorithm[5]. This algorithm is given as follows.

$$\gamma_m^{(f)}(n) = -\frac{k_{m-1}(n)}{g_{m-1}^{(b)}(n-1)}, \quad \gamma_m^{(b)}(n) = -\frac{k_{m-1}(n)}{g_{m-1}^{(f)}(n)} \tag{7}$$

$$k_{m-1}(n) = \lambda k_{m-1}(n-1) + f_{m-1}(n)b_{m-1}(n-1) \tag{8}$$

$$g_m^{(f)}(n) = g_{m-1}^{(f)}(n) - \frac{k_{m-1}^2(n)}{g_{m-1}^{(b)}(n-1)} \tag{9}$$

$$g_m^{(b)}(n) = g_{m-1}^{(b)}(n-1) - \frac{k_{m-1}^2(n)}{g_{m-1}^{(f)}(n)} \tag{10}$$

$$f_j(m) = f_{m-1}(n) + \gamma_m^{(f)}(n)b_{m-1}(n-1) \tag{11}$$

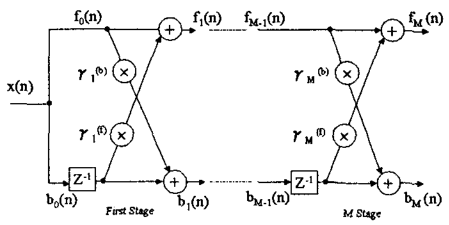
$$b_m(n) = b_{m-1}(n-1) + \gamma_m^{(b)}(n)f_{m-1}(n-1) \tag{12}$$

$$m = 1,2,\ldots,M$$

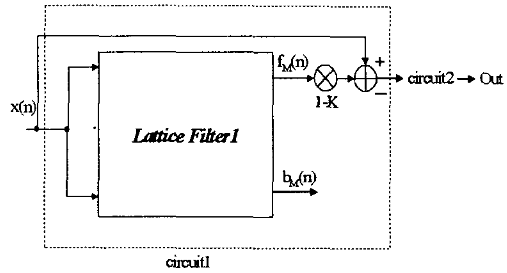
We performed computer simulation using the structure shown in figure 4(b). The parameters used in this simulation are shown in table 2. Figure 5 shows the result obtained. We observe from this result that the performance is almost the same as the results obtained using the transversal filter.

4. Actual Noise Reduction using The Control of The Predictor Filter Coefficients

In this section, we consider a practical application where noise reduction is necessary. In this noise reduction method, both the voice spectrum and the actual noise spectrum are estimated. Noise reduction is achieved by



(a) Lattice Filter.

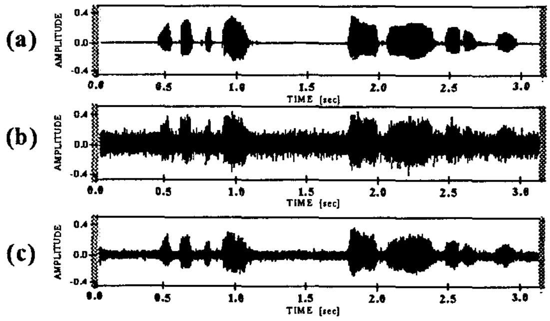


(b) Noise Reduction Circuit using Lattice Filter.

Figure 4

Table 2 Parameters used in the simulation of Fig.5.

S-N ratio of the input signal [dB]	0
The Number of Coefficient	64(×2)
Forgetting value λ	0.998
K	0.5



(a) Non-corrupted voice signal.
(b) Voice signal mixed with white noise.
(c) Voice signal extracted with lattice filter.

Figure 5

using the linear predictor filter which includes the control of the predictor filter coefficient's update. Since the change in the statistical properties of the noise is small in comparison to the voice signal, it is logical to assume that the change in the convergence value of the linear predictor's coefficients will be small when only the noise exists. The control procedure of the linear predictor's coefficients for each of the two proposed methods of noise reduction is given as follows: If only the noise signal exists, the coefficients of linear predictor are controlled so as to reduce the output. Thus, the voice spectrum is easier to estimate when compared with the noise spectrum. We therefore introduce the following condition for control of the linear predictor's coefficients.

(Noise Reduction using Transversal Filter)
When

$$T_{min} > |\, h(n+1)-h(n)\, | \tag{12}$$

h(n+1) is

$$h(n+1) = C_{at} \times h(n) \tag{13}$$

where

T_{min} : the threshold of the change in the value of the Transversal Filter's coefficients

C_{at} : the attenuation constant for Transversal Filter's coefficients

(Noise Reduction using Lattice Filter)

$$g_0^{(f)}(n) = G_{fix}$$
$$g_0^{(b)}(n) = G_{fix}$$

where

G_{fix} : the control constant for Lattice Filter's coefficients

The noise reduction using lattice filter is most effectively done when the estimate value $g_m(n)$ is corresponding to the noise variance. However, $g_m(n)$ is small in comparison to the noise variance, because actual noise spectrum is estimated by linear predictor. Then, it is assumed that the noise spectrum does not change, and fixes $g_m(n)$ by G_{fix}. The improvement in the ability to reduce noise can be expected if $g_0(n)$ is fixed near the variance of all noise.

We performed computer simulation using the two kinds of control of the linear predictor's coefficients. The new parameters used in these simulations are shown in Table 3. Figures 6 and 7 show the results obtained. In these figures, (c) and (d) represent the extracted signal without and with the control of the Linear Predictor's coefficients, respectively. We observe from this result (d) that there is considerable reduction in the actual noise in comparison to the results obtained using the linear predictor filter without the control of the coefficients update.

5. Conclusion

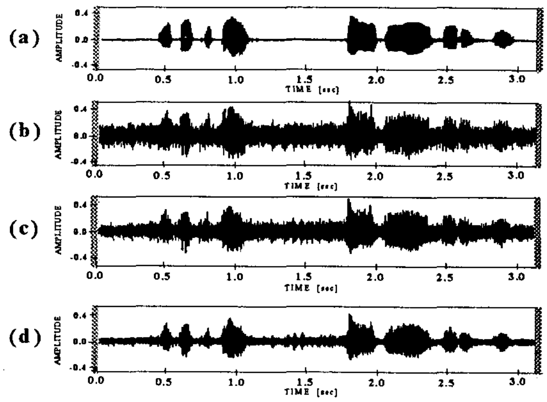
In this paper, we have proposed a new method of noise reduction and voice extraction from a voice signal mixed in noise by using the linear predictor. From the computer simulation results, it was observed that there was a decrease in noise from the extracted voice signal. Possible areas which need further research, involves performance evaluation in a test product, improving further the tone of extracted voice signal and research on a method which automatically detects the optimum value of the parameters used in the noise reduction.

Table 3 Parameters used in the simulation of Fig.6.

T_{min}	0.0005
C_{at}	0.7

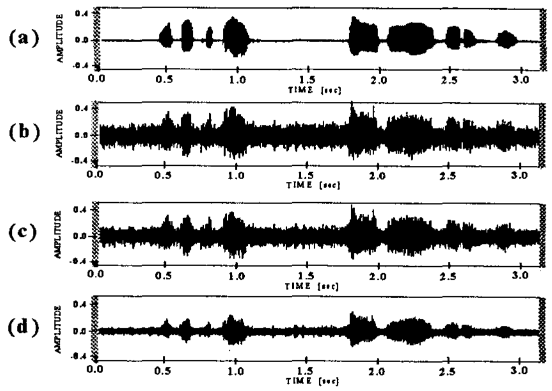
Parameter used in the simulation of Fig.7.

G_{fix}	14



(a) Non-corrupted voice signal.
(b) Voice signal mixed with actual noise.
(c) Voice signal extracted without the control of coefficients update.
(d) Voice signal extracted with the control of coefficients update.

Figure 6 Noise Reduction using Transversal Filter



(a) Non-corrupted voice signal.
(b) Voice signal mixed with actual noise.
(c) Voice signal extracted without the control of the Lattice Filter's coefficients.
(d) Voice signal extracted with the control of the Lattice Filter's coefficients.

Figure 7 Noise Reduction using Lattice Filter

REFERENCES

[1] J. L. Flanagan, J.D. Johnston, R. Zahn, and G. W. Elko. "Computer-stresses microphone arrays for sound transduction in large room," J. Acoust. Soc. Am., 78, pp. 1508-1518, 1985.

[2] S. F. Boll. "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-27, no. 2, pp. 113-120, April 1979.

[3] N. Viraf "Speech Enhancement Based on Masking Properties of the Audiory System," Proc. ICASSP, pp. 796-799, 1995

[4] P. Dreiseitel and H. Puder. "Speech Enhancement for Mobile Telephony based on Non-Uniformly Spaced Frequency Resolution," Proc. EUSIPCO, pp. 965-968, 1998

[5] S.Haykin. Adaptive Filter Theory, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[6] K. Fujii and J. Ohga. "Sub-RLS algorithm with an extremely simple update equation," IEEE International Conference on Acoustics, Speech, and Signal Processing. Munich, Germany, April 1997, vol. 3, pp. 2325-2328.