Proc. Canadian Conf. Electrical, Computer Engineering (Toronto, ON), pp. WM9.24.1-WM9.24.4, Sept. 1992

BACKWARD-ADAPTIVE PREDICTION CASCADED WITH FORWARD FORMANT AND PITCH FILTERS

Qian Yasheng [2] Peter Kabal [1,2]

[1]Dept. of Electrical Eng.
McGill University
Montreal, Quebec
Canada H3A 2A7

[2]INRS-Telecommunications
Université du Québec
Verdun, Québec
Canada H3E 1H6

Abstract

Two kinds of cascaded backward-adaptive predictor (forward formant–backward formant–forward pitch and backward formant-forward pitch) are investigated in this paper. We have analyzed and tested two important parameters for the backward-adaptive formant predictor in these configurations: the update rate of the linear prediction coefficients and the analysis frame length. We have found that if the analysis frame length of the backward-adaptive formant predictor is shorter than the pitch period, the backward prediction gain degrades rapidly. We have found that the average prediction gain for the slower update rates is close to the fast update one. The slower the update rate, the fewer the computations. Particularly, the backward predictor with slower update rate behaves more like a linear filter. These new results provide a useful platform to explore the applications of backward adaptive prediction to low bit-rate speech coders, in which the backward-adaptive formant predictor is cascaded with a forward pitch predictor or the forward formant–backward formant–forward-adaptive pitch predictor is used [1], [2].

1. Introduction

Backward-adaptive linear prediction has been used in medium rate speech coders with high quality and low delay, such as 8–16 kb/s LD-CELP coders (delay less than 2 ms) [1], [3], [4]. A forward formant–backward formant–forward pitch predictor has been incorporated into a 3 kb/s Single-Pulse Excitation/Code-Excited Linear Prediction(SPE/CELP) coder to improve the speech quality while maintaining almost the same bit-rate [2]. Backward-adaptive filters form their updates from the reconstructed signal(available to both the coder and decoder). No side-information need be transmitted. Forward-adaptive filters base their updates on the input signal and, hence, need explicit transmission of the filter parameters.

This research was supported by a grant from the Canadian Institute for Telecommunications Research under the NCE program of the Government of Canada

In a conventional forward-adaptive formant predictor, the shorter the analysis frame length, the higher the prediction gain. However, the shorter the analysis frame length, the more the side-information to be transmitted. A fast update forward-adaptive formant predictor (sample-by-sample update) performs better than a slow update predictor. The update rate or the analysis frame length are determined by the available bits for the side-information transmission.

By contrast, if the analysis frame length of the backward-adaptive formant predictor is shorter than the pitch period, the backward prediction gain decreases rapidly. Therefore, the analysis frame length must be larger than the pitch period.

We have compared the prediction gain of the backward-adaptive formant predictor in two configurations with different update rates of the coefficients varying from sample-by-sample to a block length of 40 samples. We have found that the average prediction gain for the slower update rates can be close to the fast update one. The slower the update rate, the fewer the computations. Particularly, the backward predictor with the slower update rate behaves more like a linear filter. These new results provide us a useful platform to explore the applications of backward adaptive prediction to low bit-rate speech coders [1],[2].

In the following, we first introduce the analysis model of the backward-adaptive formant predictor. Then, we will describe the analysis window length and the update rate of the backward-adaptive formant predictor. Finally, we give the pitch prediction gain of these configurations: forward formant-backward and forward pitch predictor and backward formant-forward pitch predictor.

2. The backward-adaptive formant predictor

There are well-known algorithms for the backward-adaptive predictor, such as the Least Mean Square (LMS) Method, etc. [5]. The backward adaptation algorithm used in our study is different from LMS. An analysis model for calculating the prediction coefficients of the backward-adaptive formant predictor with a transversal implementation is illustrated in Fig. 1. The input signal

$x(n)$ is multiplied by a data window $w_d(n)$ to give $x_w(n)$. The signal $x_w(n)$ is predicted from a set of its previous samples to form an error signal,

$$e(n) = x_w(n) - \sum_{k=1}^{N_p} c_k x_w(n-k) \quad (1)$$

The final step is to multiply the error signal by a error window $w_e(n)$ to obtain a windowed error signal $e_w(n)$ where $e_w(n) = w_e(n)e(n)$. The mean square error is defined by,

$$\varepsilon^2 = \sum_{n=\infty}^{\infty} e_w^2(n) \quad (2)$$

The coefficients c_k are computed by minimizing ε^2. This leads to a linear system of equations which can be written in matrix form, $\Phi c = \alpha$

$$\begin{bmatrix} \phi(1,1) & \cdots & \phi(1,N_p) \\ \phi(2,1) & \cdots & \phi(2,N_p) \\ \vdots & \vdots & \vdots \\ \phi(NP,1) & \cdots & \phi(N_p,N_p) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{N_p} \end{bmatrix} = \begin{bmatrix} \phi(0,1) \\ \phi(0,2) \\ \vdots \\ \phi(0,N_p) \end{bmatrix} \quad (3)$$

where

$$\phi(i,j) = \sum_{n=-\infty}^{\infty} w_e^2(n) x_w(n-i) x_w(n-j) \quad (4)$$

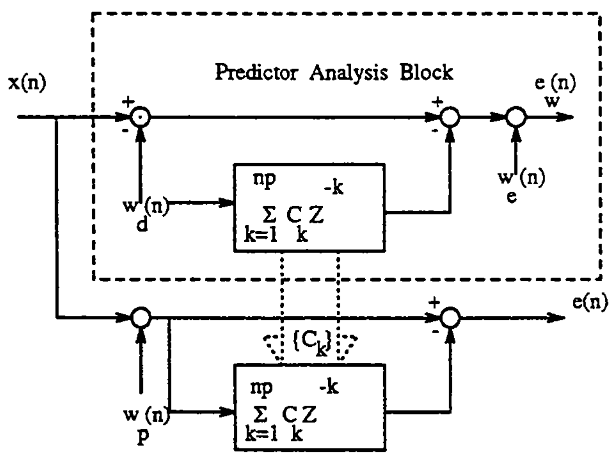

Fig. 1 Backward-Adaptive Formant Predictor

There are two well-known analysis methods: autocorrelation method and covariance method. The autocorrelation method results if $w_e(n) = 1$ for all n. The data window $w_d(n)$ is typically time-limited (rectangular, Hamming or others). The covariance method results if $w_d(n) = 1$ for all n and the error window is rectangular, $w_e(n) = 1$ for $0 \le n \le N-1$. Since the data window $w_d(n)$ of a conventional forward-adaptive prediction coefficients analysis overlaps the predictor window $w_p(n)$, the error signals are included in the data window. In the sense of the Minimum Mean Square Error(MMSE), the forward-adaptive formant predictor is an optimum predictor. On the contrary, the data window $w_d(n)$ of the backward-adaptive formant predictor precedes the predictor window $w_p(n)$, that is, the estimated or error signals are not included in the data analysis window. Therefore, the backward-adaptive predictor is not an optimum predictor in the sense of the MMSE. The prediction coefficients are derived by autocorrelation or covariance algorithm from the data analysis window, as shown in Fig. 2. The coefficients then multiply the samples within the predictor window to predict the last sample in the predictor window. Note that causality requires that the last sample of the predictor window lies to the right of the data window. The prediction coefficients are updated sample-by-sample or block-by-block, that is, the backward-adaptive formant predictor has a set of new prediction coefficients each sample or each block.

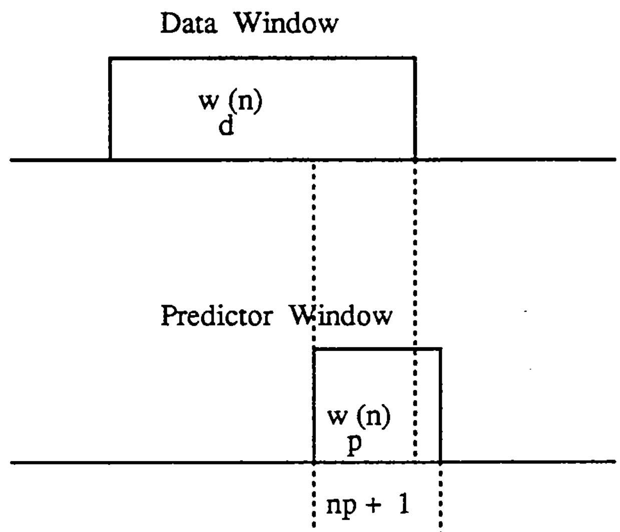

Fig. 2 Windows for Backward-adaptive Predictors

3. Analysis window length of the backward-adaptive formant predictor

A backward-adaptive formant predictor cascaded with a conventional forward-adaptive predictor, as shown in Fig. 3, is first investigated.

The first forward-adaptive predictor is a conventional 10-th order linear predictor using Durbin's recursive algorithm. The residuals of the first predictor is then fed to the second backward-adaptive formant predictor. The backward-adaptive predictor window with one sample ahead of the data analysis window will not achieve to an optimum forward prediction gain, unless the correlation matrix of the backward-adaptive predictor Φ in (4) is identical to the correlation matrix of the forward-adaptive predictor.

As to a periodical signal, the correlation matrix is identical to the one-sample or one block shifted version, if the data analysis window is larger than the period. There

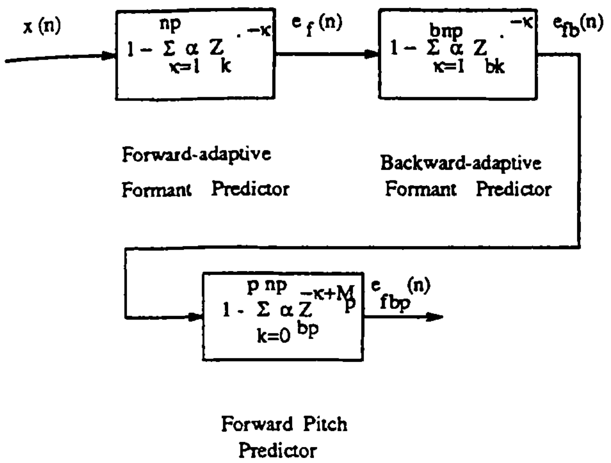

Fig. 3 The Forward Formant-Backward Formant-Formant Pitch Predictor

are a series of pitch pulses in the residuals of the first forward formant predictor for the voiced segments. If the data analysis window is shorter than the pitch period, the difference between these two correlation matrixes could be quite large. Therefore, the prediction gain of the backward predictor falls down sharply.

Several speech segments of male and female speakers have been tested to calculate the backward prediction gains. The forward prediction gains are also noted for comparison. Both of the backward and forward prediction gains for different data analysis window lengths are listed in Table 1. The pitch for the tested male voiced segments is about 77 samples. The backward prediction gain is 5.03 dB for 10th order backward predictor of 80 sample window length. The corresponding forward prediction gain is 5.99 dB. However, the backward prediction gains degrade rapidly, when the window length is shorter than the pitch period of 77 samples for male speech. The backward prediction gains go down to −10.7 dB for the male speech and to −0.64 dB for the female speech (pitch period of 36 samples) with the data analysis window of 12 samples. However, the forward prediction gain reaches to the highest values of 11.4 dB and 12.8 dB respectively. For these tests, the backward predictors use covariance algorithm to produce the prediction coefficients.

In the backward formant-pitch filter the input signal for the backward-adaptive formant predictor is the original speech signal. We have tested the prediction gains using different data analysis window lengths with sample-by-sample update rate. The backward prediction gain goes as high as 21.5 dB for female speech and 19.8 dB for male speech, if the analysis window has 110 samples and the backward predictor is 10-th order. The backward prediction gains degrade rapidly down to 11.5 dB and 11.4 dB respectively, when the window length is shorter than the corresponding pitch period. For these tests, the backward predictors use autocorrelation algorithm to produce the prediction coefficients.

Prediction Gain (dB) Male			
		Backward	Forward
Order	Framesize	PGain(dB)	PGain(dB)
10	110	5.54	5.31
10	80	5.03	5.99
10	50	1.67	7.24
10	20	−3.90	9.53
10	12	−10.7	11.4

Table 1 Prediction Gain of the Backward Formant Predictor for LP-10 Residuals (Male speech). Each row shows the average prediction gain for voiced frames

		Male Speech	Female Speech
Order	Framesize	PGain(dB)	PGain(dB)
10	110	19.8	21.5
10	80	17.8	21.3
10	50	15.7	19.7
10	22	14.2	16.5
10	12	11.4	11.5

Table 2 Prediction Gain of the Backward-adaptive Predictor for Male and Female Speech with different data analysis window length.

4. Update rate of the backward-adaptive formant predictor

We have tested the backward prediction gains with different update rates of the prediction coefficients. The fastest update rate calculates the prediction coefficients, based on sample-by-sample. Each time the data analysis window moves one sample forward. The slowest update rate computes the prediction coefficients, based on block-by-block, with block length of 40 samples. The results in Table 3 have shown that the prediction gains between the fastest update rate and the slowest are less than 0.8 −2.5 dB.

Speech	Male	Female
Update Rate	PGain (dB)	PGain (dB)
1	19.8	22.4
10	20.1	22.5
20	20.9	21.0
40	19.0	19.9

Table 3 Prediction Gain of the Backward-Adaptive Predictor of different update rates. The data analysis window has 110 samples.

5. Cascaded backward-adaptive formant predictor with pitch filter

Two kinds of cascaded backward-adaptive formant

Update Rate (samples)	Backward PGain (dB)		Pitch P Gain(dB)	
Speech	Male	Female	Male	Female
1	19.2	21.5	8.84	11.9
10	20.1	22.5	6.51	8.32
20	20.9	21.0	6.62	8.80
40	19.0	19.9	6.97	9.93

Table 5 Pitch prediction gains with different update rates of the Backward Formant-Forward Pitch predictor configuration.

with pitch filter have been used in the low-bit rate speech coder, as shown in Fig. 3 and 4. The Backward Formant-Forward Pitch predictor, as shown in Fig. 4, is employed in a Low-Delay 8 kb/s speech coder[1]. We have studied the pitch prediction gains for both configurations with different update rates of the backward-adaptive formant predictor and fixed data analysis length of 110 samples. The pitch prediction coefficients and pitch lag are determined by minimizing the mean square error of the pitch predictor[6]. Pitch parameters are updated with 20 samples (2.5 ms). Results for the voiced segments are shown in Table 4 and 5. We have found that the pitch prediction gains for both configurations are still very high, because the backward-adaptive formant predictor removes only the short term redundancy. Therefore, both configurations can be successfully applied to a Code-Excited Linear Prediction coder.

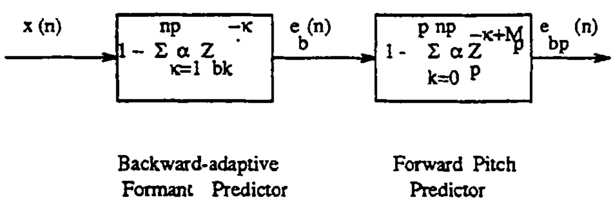

Fig. 4 The Backward Formant-Forward Pitch Predictor

Update Rate (samples)	Backward PGain (dB)		Pitch P Gain(dB)	
Speech	Male	Female	Male	Female
1	4.25	8.28	4.00	7.18
10	4.75	8.36	4.03	6.73
20	4.82	8.64	3.90	6.92
40	4.70	8.60	3.87	7.11

Table 4 Pitch prediction gains with different update rates of the Forward Formant-Backward Formant-Forward Pitch predictor configuration.

6. CONCLUSIONS

We have analyzed and verified that if the data analysis window of the backward-adaptive formant predictor is shorter than the pitch period of the speech signal, the prediction gain of the backward predictor decreases rapidly. If the data analysis window is longer than the pitch period, the backward formant predictor can approach to an optimum forward predictor. This is an important parameter for the design of the backward-adaptive formant predictor. Since the characteristics of speech signal is quasi-stationary in short segments (5 ms), the backward-adaptive formant predictor can reach high prediction gain at update rate of 20–40 samples (2.5–5 ms). Because of the nature of the backward adaptation, the fastest update rate of the prediction coefficients gives better performance. However, it increases the computation load greatly. Therefore, the slower update rate of the predictor parameters is a good compromise between the computation load and prediction gain for the backward-adaptive formant predictor.

Since a 10-th order forward-adaptive formant predictor cannot remove all redundancy of the speech signal, a second cascaded backward-adaptive formant predictor can obtain addition prediction gain with no extra side-information as a non-linear predictor. The performance of the second backward-adaptive predictor is close to a non-linear predictor [7]. The proposed configuration can be applied to a low bit-rate speech coder to improve its quality. We have incorporated the backward- adaptive formant synthesizer into a 3 kb/s SPE/CELP speech coder. It has shown that it can gain $SNR_{seg,bf}$ improvements of 1.02 to 2.60 dB for some voiced frames[2].

References

1. J. Chen and M. Rauchwerk,"An 8 kb/s Low-Delay CELP Speech Coder," *IEEE Global Telecommun. Conf.* (Phoenix, AZ), Dec. 2–5, 1991, pp.53.7.1–53.7.5
2. Y. Qian, Y. Cheng and P. Kabal, "Backward Adaptation for Single-Pulse Excitation Coder", *Proc. Int. Conf. on Communications Technology* Beijing, China, Sept.16–18 1992
3. V. Iyengar and P. Kabal,"A low delay 16 kbits/s speech coder", *IEEE Trans. Signal Processing,* Vol. 39, No. 5, May 1991, pp. 1049–1057
4. J. Chen, "A robust low-delay speech coder at 16 kb/s", *IEEE Global Telecommun. Conf.* (Dallas, TX), Nov. 27–30, 1989, pp. 1237–1241.
5. N. Jayant and P. Noll, "Digital Coding of Waveforms", *Prentice-Hall, Inc.* (Englewood Cliff, NJ), 1984, pp. 303–306
6. R. Ramachandran and P. Kabal, "Pitch Prediction Filters in Speech Coding", *IEEE Trans. on Acoustics, Speech, Signal Processing,* Vol.37, No. 4, April 1989, pp.467–478
7. B. Townshend, "Non-linear prediction of speech",*Proc. IEEE Conf. on Acoustics, Speech, Signal Processing,*(Toronto, Ont.), May 1991, pp.425–428