BANDPASS ADAPTIVE POLE-ZERO FILTERING

John J. Shynk Bernard Widrow

Information Systems Laboratory
Department of Electrical Engineering
Stanford University
Stanford, CA 94305

Abstract - An adaptive pole-zero filter comprised of a bank of bandpass filters is presented. The bandpass filters permit the adaptive filter to be realized in a parallel form of first-order sections. Simple monitoring of the filter poles during adaptation is therefore possible so that stability can be ensured. This paper focuses on bandpass filters which are implemented by a frequency-sampling structure; the use of other types of bandpass filters is briefly discussed. An application in system identification is described and computer simulation results are given.

INTRODUCTION

Most adaptive pole-zero filters discussed in the literature have been direct-form realizations [1,2]. Recently, a *parallel-form* adaptive pole-zero filter implemented in the frequency domain was introduced [3,4]. The frequency-domain adaptive pole-zero filter (FDAF) uses a discrete Fourier transform (DFT) to split the input signal into several (approximately) orthogonal signals. These signals are then independently filtered by a bank of adaptive first-order filters which permit simple monitoring of the adaptive filter poles without the large complexity generally required by direct-form realizations.

The DFT of the FDAF essentially operates as a *bank of bandpass filters*. This interpretation is easily seen when a frequency-sampling structure is used to implement the DFT [5]. It turns out that the bandpass filtering is fundamental to the operation of the FDAF. Consequently, other types of bandpass filters could be used so that the FDAF can be generalized to a class of parallel-form adaptive pole-zero filters. We will refer to this class of adaptive filters as bandpass adaptive pole-zero filters.

BANDPASS ADAPTIVE POLE-ZERO FILTER

Filter Description

The bandpass adaptive pole-zero filter (BPAF) is comprised of a bank of N bandpass filters as shown in Fig. 1. Each bandpass filter is nonadaptive and each has a different transfer function denoted by $F_k(z)$. Observe that the input signal $x(n)$ is split into N signals $u_k(n)$ by the bandpass filters. The signals $u_k(n)$ are then filtered by a bank of adaptive subfilters $H_k(n,z)$ to produce N intermediate output signals $y_k(n)$. The intermediate output signals are added to give the overall filter output $y(n)$, which is compared with the desired response $d(n)$ to generate an error signal $e(n)$. An adaptive algorithm uses this error to adjust the coefficients of the adaptive subfilters $H_k(n,z)$ so that the mean-square-error (MSE) is minimized.

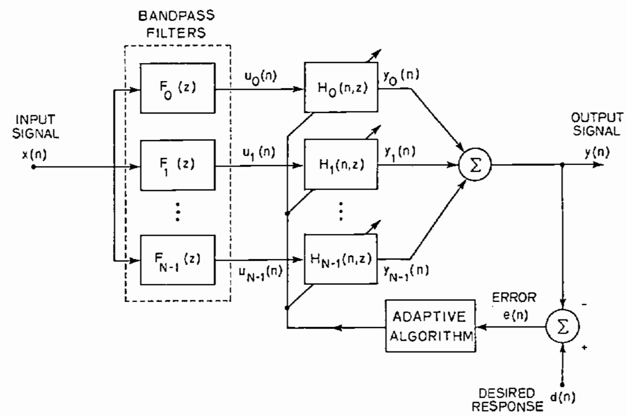

Fig. 1. Bandpass adaptive pole-zero filter.

The adaptive subfilters are comprised of a single pole and a single zero as follows:

$$H_k(n,z) = \frac{b_k(n) + c_k(n)z^{-1}}{1 - a_k(n)z^{-1}} , \qquad (1)$$

where $\{a_k, b_k, c_k\}$ are the adaptive coefficients. Since each $H_k(n,z)$ has only one pole, it is simple to monitor the adaptive filter poles for stability. If a pole attempts to move outside the unit circle, the adaptive update for that pole is simply ignored. The update of all stable poles can be performed however so that the algorithm is less likely to lock up [2].

The purpose of the bandpass filters is to separate the energy of the input signal into N approximately nonoverlapping frequency bins. As a result, the signals $u_k(n)$ are (approximately) orthogonal so that it

ICASSP 86, TOKYO CH2243-4/86/0000-2107 $1.00 © 1986 IEEE

is possible to adjust the coefficients of the adaptive subfilters independently of each other and still have satisfactory convergence of the adaptive algorithm. In effect, the bandpass filters transform a wideband adaptive filter into several narrowband adaptive filters which operate independently to minimize the common error $e(n)$.

If a frequency-sampling (FS) structure is used to implement the bandpass filters, then $F_k(z)$ is given by

$$F_k(z) = \frac{1}{N}\,\frac{1 - \beta^N W_N^{k_0 N} z^{-N}}{1 - \beta W_N^{k+k_0} z^{-1}} \qquad (2)$$

where $W_N = e^{j2\pi/N}$. The transfer function of (2) is comprised of N zeros equally spaced around a circle of radius β; a pole at $z = \beta W_N^{k+k_0}$ exactly cancels one of these zeros. As a result, each bandpass filter (theoretically) has a finite impulse response. The radial contraction coefficient β, where $0<\beta<1$, insures that the bandpass filter pole (and zeros) lie inside the unit circle so that (2) is stable. The angular shift factor k_0, where $-1/2 \le k_0 \le 1/2$, rotates the frequency bin centers so that no bin centers lie on the real axis. This permits all of the coefficients $\{a_k, b_k, c_k\}$ to be complex valued (see [4] for further details).

The frequency response $|F_k(e^{j\omega})|^2$ of (2) is shown in Fig. 2 for $N=8$, $k=k_0=0$, and three different values of β. For ease of comparison, each curve has been normalized by the corresponding area of $|F_k(e^{j\omega})|^2$. From the figure, we see that decreasing the radial contraction coefficient β increases the overlap between frequency bins. This in turn increases the correlation between the signals $u_k(n)$ so that the adaptive algorithm will generally converge more slowly.

If we assume that $x(n)$ is a white random process with unit variance, then the magnitude of the correlation coefficient between the signals $u_k(n)$ and $u_l(n)$ is given by

$$|\rho_{kl}| = \frac{1-\beta^2}{(1 - 2\beta^2\cos(2\pi|k-l|/N) + \beta^4)^{1/2}} \qquad (3)$$

Equation (3) is plotted in Fig. 3 for $|k-l|=1$ and three different values of N. As expected, the correlation between frequency bins increases as β decreases toward zero. Furthermore, $|\rho_{kl}|$ becomes more sensitive to changes in β as N is increased. Typically β would be chosen close to one so that ρ is nearly zero; a possible exception to this is discussed later in this section. Nevertheless, varying β can be a useful means of controlling the amount of correlation to observe the effect on convergence of the adaptive algorithm. Some results of this are given in the system identification application section.

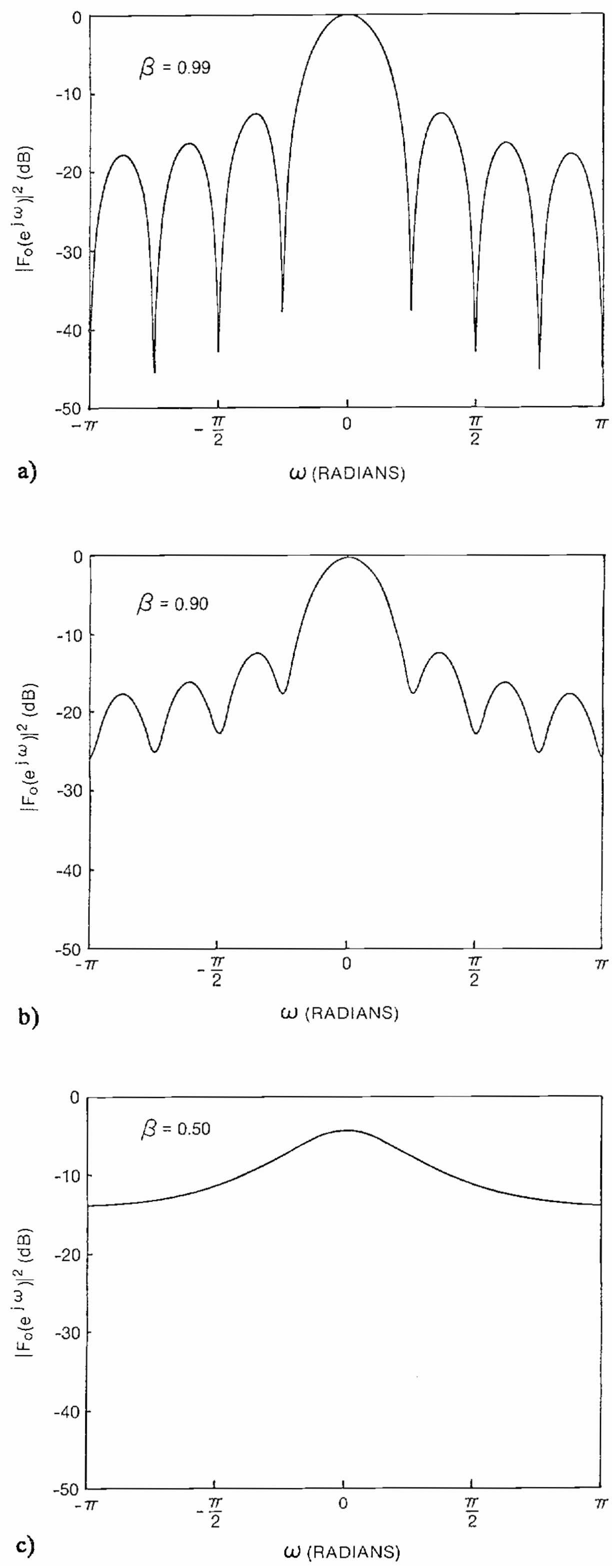

Fig. 2. Bandpass filter frequency response for a) $\beta = .99$, b) $\beta = .90$, c) $\beta = .50$.

40. 5. 2

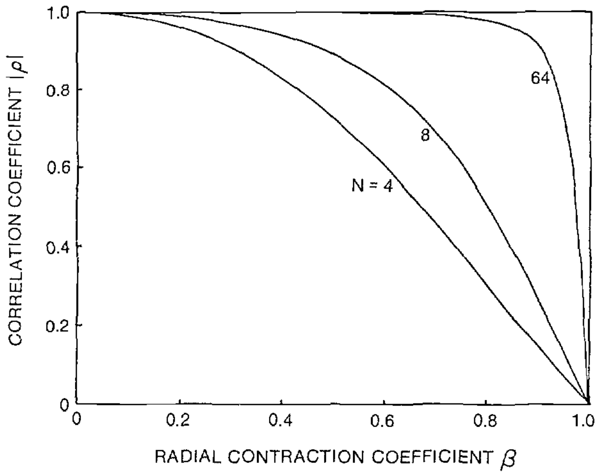

Fig. 3. Correlation coefficient for $|k-l| = 1$.

Adaptive Algorithm

The adaptive algorithm is shown in Table 1. The superscripts * and H refer to complex conjugate and conjugate transpose, respectively. Observe that the signal $w(n)$ is common to all bandpass filters as given by (2). The matrix $R_k(n)$ is the Hessian matrix for the k^{th} subfilter; it is initialized to δI, where δ is a scalar approximating the power of $u_k(n)$ and I is the identity matrix. The scalar α controls the algorithm convergence rate; it is generally chosen so that $1-\alpha$ is close to one. A more complete description of the adaptive algorithm can be found in [4,6].

Imperfect Pole-Zero Cancelation

When implementing (2), we cannot expect perfect pole-zero cancelation. As a consequence, the impulse response corresponding to (2) is actually infinite. This could be undesirable since a significantly long impulse response might adversely affect the convergence rate of the adaptive algorithm. If we assume that β of the numerator of (2) is replaced by $\beta+\epsilon$ where ϵ is the deviation from β, then the impulse response of (2) is (approximately) given by

$$|f_k(n)| \approx \frac{\beta^n}{N}\Big[s(n) - s(n-N)\Big] \qquad (4)$$

$$+\ \epsilon\, \beta^{n-1} s(n-N)\ ,$$

where $s(n)$ is the unit step function. The first term of the right side of (4) has finite duration and corresponds to the impulse response of (2) if exact pole-zero cancelation were possible. The second term of the right side of (4) has infinite duration and is introduced by the pole-zero mismatch. Observe however that this term is proportional to ϵ. Since ϵ is typically much less than one, we would expect little effect from imperfect pole-zero cancelation, except possibly for large N and for β very close to the unit circle. Thus, it may be desirable to choose β somewhat less than one (e.g., $\beta = 0.90$). The time constant τ for the decay of the infinite part of the impulse response is approximately given by $\tau = 1/(1-\beta)$. Choosing $\beta = 0.90$ results in $\tau = 10$ iterations as opposed to $\tau = 100$ iterations for $\beta = 0.99$.

Table 1 Adaptive Algorithm

INITIALIZATION:

$$a_k(0) = b_k(0) = c_k(0) = 0$$

$$u_k(-1) = y_k(-1) = u_k^f(-1) = y_k^f(-1) = 0$$

$$R_k(0) = \delta I$$

VECTOR DEFINITIONS:

$$\theta_k(n) = [\, a_k(n)\,,\, b_k(n)\,,\, c_k(n)\,]^T$$

$$\phi_k(n) = [\, y_k(n-1)\,,\, u_k(n)\,,\, u_k(n-1)\,]^T$$

$$\psi_k(n) = [\, y_k^f(n-1)\,,\, u_k^f(n)\,,\, u_k^f(n-1)\,]^T$$

FOR EACH NEW INPUT $x(n)$, $d(n)$; $n \geq 0$:

$$w(n) = \frac{1}{N}[\, x(n) - (\,\beta W_N^{k_o})^N x(n-N)\,]$$

FOR $k = 0, 1, \ldots, N-1$:

$$u_k(n) = w(n) + \beta W_N^{k+k_o}\, u_k(n-1)$$

$$u_k^f(n) = u_k(n) + a_k^*(n)\, u_k^f(n-1)$$

$$y_k(n) = \theta_k^H(n)\, \psi_k(n)$$

$$y_k^f(n) = y_k(n) + a_k^*(n)\, y_k^f(n-1)$$

$$e(n) = d(n) - \sum_{k=0}^{N-1} y_k(n)$$

FOR $k = 0, 1, \ldots, N-1$:

$$R_k(n+1) = (1-\alpha)R_k(n) + \alpha\psi_k(n)\psi_k^H(n)$$

$$\theta_k(n+1) = \theta_k(n) + \alpha R_k^{-1}(n+1)\, \psi_k(n)\, e^*(n)$$

Extensions

From the previous discussion, it is clear that other types of bandpass filters could be used in the BPAF. For example, bandpass filters with less overlap would result in the signals $u_k(n)$ being less correlated so that the adaptive algorithm would converge more rapidly. Employing such filters would undoubtedly require more complexity. There is consequently a tradeoff between the algorithm convergence rate and the complexity of the bandpass filters.

There is also a modeling issue which may be a consideration in some applications. It was shown in [3] that the FS structure permits the BPAF to model an arbitrary rational system. This property, which would be important for example in system identification applications, is obviously not satisfied by all types of bandpass filters.

One family of bandpass filters which does satisfy this property is based on the Lagrange structure [5], of which the FS structure is a special case. The transfer function of a Lagrange-type filter is comprised of N *arbitrarily-placed* zeros and a pole which cancels one of these zeros. This structure would be useful, for example, if it is known a priori that the spectrum of $d(n)$ is concentrated in some region $|\omega| < \omega_c$. In this case, it would be more efficient to use bandpass filters based on the Lagrange structure instead of those based on the FS structure.

APPLICATION IN SYSTEM IDENTIFICATION

In system identification, the desired response signal and the adaptive filter input are often assumed to be generated as follows [7]:

$$d(n) = G(z)x(n) + v(n) \quad , \tag{5}$$

where $x(n)$ and $v(n)$ are zero-mean, mutually uncorrelated signals, and $G(z)$ is the system to be identified. The particular system to be identified was

$$G(z) = \frac{K(1-0.85z^{-1}+0.36z^{-2})(1+1.39z^{-1}+0.64z^{-2})}{(1-0.87z^{-1}+0.25z^{-2})(1+0.75z^{-1}+0.56z^{-2})} \tag{6}$$

which has four complex poles at $p_{1,2} = 0.50\angle\pm 30°$ and $p_{3,4} = 0.75\angle\pm 120°$, and four zeros at $z_{1,2} = 0.60\angle\pm 45°$ and $z_{3,4} = 0.80\angle\pm 150°$. The input signal was a real white random process and the constant K was chosen such that $G(z)x(n)$ had unit variance. The additive noise signal was also a real white random process with a variance of 10^{-4}. The adaptive algorithm of Table 1 was used with the following parameters: $\alpha = 0.01$, $N = 4$, $\delta = 0.25$, $k_0 = 0.5$, and three different values of β. They were $\beta = 0.50$, 0.90, and 0.99.

Fig. 4 displays the MSE learning curves obtained by averaging $|e(n)|^2$ over 200 independent simulation runs. Only the curves for $\beta = 0.50$ and 0.90 are shown since the curve for $\beta = 0.99$ was nearly indistinguishable from that of $\beta = 0.90$. Observe that the curve for $\beta = 0.90$ decreases rapidly to the noise floor of -40 dB, converging (approximately) by iteration 2000. On the other hand, the curve for $\beta = 0.50$ converges more slowly. This slower convergence is due to the increased correlation between frequency bins as shown in Fig. 3.

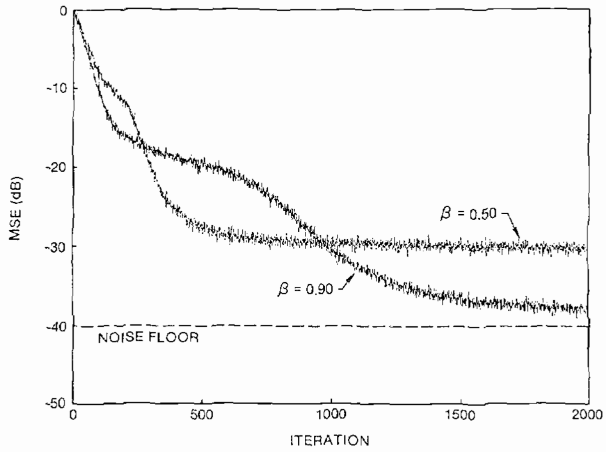

Fig. 4. MSE learning curves.

CONCLUSIONS

We have presented a parallel-form adaptive pole-zero filter which is comprised of a bank of bandpass filters (Fig. 1). A frequency-sampling structure was used to implement the bandpass filters, although other types of bandpass filters are possible. There generally is a tradeoff between the complexity of the bandpass filters and the convergence rate of the adaptive algorithm. The bandpass adaptive pole-zero filter is always stable and, when using a frequency-sampling structure, it demonstrates rapid convergence in system identification applications (Fig. 4).

REFERENCES

[1] B. Widrow and S.D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1985.

[2] C.R. Johnson, Jr., "Adaptive IIR filtering: Current results and open issues," *IEEE Trans. Inform. Theory*, vol. IT-30, no. 2, pp. 237-250, Mar. 1984.

[3] J.J. Shynk and R.P. Gooch, "Frequency-domain adaptive pole-zero filtering," *Proc. IEEE*, vol. 73, no. 10, pp. 1526-1528, Oct. 1985.

[4] J.J. Shynk, R.P. Gooch, and B. Widrow, "A frequency-domain adaptive pole-zero filter with applications," *Proc. 19th Asilomar Conf. Circuits, Systems, Computers*, Pacific Grove, CA, Nov. 6-8, 1985.

[5] L.R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1975.

[6] J.J. Shynk, "A complex adaptive algorithm for IIR filtering," *IEEE Trans. Acoust., Speech, Signal Proc.*, accepted for publication.

[7] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*, Cambridge, MA: MIT Press, 1983.