T-ASSP/35/8/15135

A Discrete Optimization Method for High-Order FIR Filters with Finite Wordlength Coefficients

K. Nakayama

Reprinted from: IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING,
Vol. ASSP-35, No. 8, AUG. 1987.

A Discrete Optimization Method for High-Order FIR Filters with Finite Wordlength Coefficients

KENJI NAKAYAMA

Abstract—A transfer function is constructed in a cascade form, using a low-order error free function and a high-order function. The high-order function is discretely optimized so that its error spectrum is suppressed by the error-free function. In order to save computing time, the error spectrum is equivalently evaluated in a time domain, and the coefficients are divided into small groups in a discrete optimization procedure.

Manuscript received December 23, 1985; revised January 16, 1987.
The author is with the C&C Systems Research Labs, NEC Corporation, 4-1-1, Miyazaki, Miyamae-Ku, Kawasaki, 213 Japan.
IEEE Log Number 8715135.

I. Introduction

As digital filters have been increasingly applied to various fields, their circuit complexity reduction becomes an important design problem [1]. Among many design aspects for the above problem, this correspondence particularly concerns how to optimize finite wordlength coefficients, using a transfer function approximated with infinite precision as the initial condition. Another approach, for instance, how to find an optimum tradeoff between coefficient wordlengths and a filter order, is not discussed.

Several useful approaches to infinite impulse response (IIR) filters and low-order finite impulse response (FIR) filters have been reported [2]-[6]. On the other hand, for high-order FIR filters, mixed-integer programming techniques [7]-[9], and a local search method [10] have been mainly applied. These approaches, however, still require a large amount of computing time.

This correspondence proposes a new discrete optimization method directed toward saving computing time for high-order FIR filters [11].

II. Discrete Optimization by Error Spectrum Shaping

A. Algorithm

A transfer function $H(z)$ basically has a cascade form structure.

$$H(z) = W(z)\,F(z), \quad z = e^{j\omega T} \tag{1}$$

where T is a sampling period and is assumed to be unity. $W(z)$ and $F(z)$ are a low-order function with prerounded off coefficients and a high-order function to be discretely optimized, respectively.

Let $\Delta F(z)$ be an error function for $F(z)$, caused by quantizing the coefficients, and let it be expressed by

$$\Delta F(z) = F(z) - F_Q(z) \tag{2}$$

where $F_Q(z)$ represents a function with rounded off coefficients. From (1), quantization error for $H(z)$ is expressed by

$$\Delta H(z) = W(z)\,\Delta F(z). \tag{3}$$

In the proposed method, the mean square of $|\Delta H(e^{jw})|$ is employed as an error criterion,

$$E = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\Delta H(e^{j\omega})\right|^2 d\omega \tag{4a}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|W(e^{j\omega})\,\Delta F(e^{j\omega})\right|^2 d\omega. \tag{4b}$$

Minimizing E is equal to shaping $|\Delta F(e^{jw})|$ to be suppressed by $W(z)$. Therefore, the $W(z)$ amplitude response is required to be small in the stopband.

B. Transfer Function Approximation

The transfer function $H(z)$ is first approximated through conventional methods. One approach by the Remez-exchange method [12] is described here.

Letting $D(\omega)$ and $U(\omega)$ be a desired amplitude response and a weighting function for error evaluation, respectively, an approximation process is stated as

$$U(\omega_k)\left[D(\omega_k) - \left|H(e^{j\omega_k})\right|\right] = (-1)^k\delta, \quad k = 0, 1, \cdots, r \tag{5}$$

where $\{\omega_k\}$ is a set of extremal frequencies. From (1), (5) can be rewritten as

$$\left|W(e^{j\omega_k})\right| U(\omega_k)\left[\left|W^{-1}(e^{j\omega_k})\right| D(\omega_k) - \left|F(e^{j\omega_k})\right|\right] = (-1)^k\delta. \tag{6}$$

Since $W(z)$ is fixed, $F(z)$ can be approximated using $|W(e^{jw})|\,U(\omega)$ and $|W^{-1}(e^{jw})|\,D(\omega)$ as a modified weighting function and a desired amplitude response, respectively. The number of the extremal frequencies is equal to the degrees of freedom in $F(z)$. Therefore, the result obtained by solving (6) becomes a near-optimum solution, compared to the direct approximation of $H(z)$. For this reason, a low-order function is desired for $W(z)$.

III. Discrete Optimization Procedure

A. Error Evaluation

Although the error criterion was given by (4), in an actual optimization procedure, E is transformed into a time domain, in order to save computing time. Letting Δh_n be an impulse response for $\Delta H(z)$, E is rewritten as

$$E = \sum_{n=0}^{N-1} \Delta h_n^2 \tag{7}$$

through Parseval's relation [1]. By letting Δf_n and w_n be impulse responses for $\Delta F(z)$ and $W(z)$, respectively, Δh_n is expressed by

$$\Delta h_n = \sum_{m=n_1}^{n_2} w_m \Delta f_{n-m},$$
$$n_1 = \max\{0, n - N_F + 1\}, \quad n_2 = \min\{n, M - 1\} \tag{8}$$

where N_F and M are orders of $F(z)$ and $W(z)$, respectively. From (7) and (8), E is further rewritten as

$$E = \sum_{n=0}^{N-1}\left(\sum_{m=n_1}^{n_2} w_m \Delta f_{n-m}\right)^2. \tag{9}$$

B. Discrete Optimization Procedure

An optimum solution for a set of Δf_n, which minimize E given by (9), is discretely searched for. In this procedure, the number of Δf_n combinations is extremely large. Therefore, a set of Δf_n is divided into small groups. This means E is successively evaluated using the partial sums of Δh_n^2, as follows:

$$E = \sum_{k=1}^{[N/K-K']} E_k, \quad K' < K \tag{10a}$$

$$E_k = \sum_{i=0}^{K-1} \Delta h^2_{k(K-K')-i}. \tag{10b}$$

The partial sum E_k is individually minimized. Adjoining partial sums E_k and E_{k+1} contain common Δh_i, $k(K - K') - K' + 1 \leq i \leq k(K - K')$. In other words, they contain $(K' + M)$ common coefficients Δf_i. Therefore, by optimizing a part of the common coefficients for both E_k and E_{k+1}, a near-optimum solution can be obtained, even though E is separately evaluated.

Search Method: Local and heuristic search methods cannot avoid the risk of falling into a local solution. Since the proposed approach drastically saves the number of assignments, a global search method can be employed.

Search Region: Since the number of Δf_n combinations is exponentially proportional to the number of grids, on which Δf_n is discretely searched for, a moderate search region must be chosen.

C. Modified Weighting Function

Since Δf_n is searched for in a restricted reion, error spectrum shaping is not complete. In other words, an amplitude response for $\Delta F(z)$ is not exactly proportional to that for $1/W(z)$. Therefore, if a mini-max criterion is employed, a weighting function used in the discrete optimization procedure should be modified from that for the transfer function.

D. Number of Computations

Letting the number of Δf_n to be used for E_k minimization be L, all possible combinations of Δf_n become P^L, where P is the number of grids. Furthermore, the number of E_k is $[N_F/(K - K')]$[1]. Hence, the total number of assignments becomes

$$N(E) = P^L[N_F/(K - K')]. \tag{11}$$

The numbers of real multiplications and additions required in E_k calculations are both $(M + 2)K$. Their total numbers are given by

[1] Letting R be a real number, $[R]$ is an integer not exceeding R.

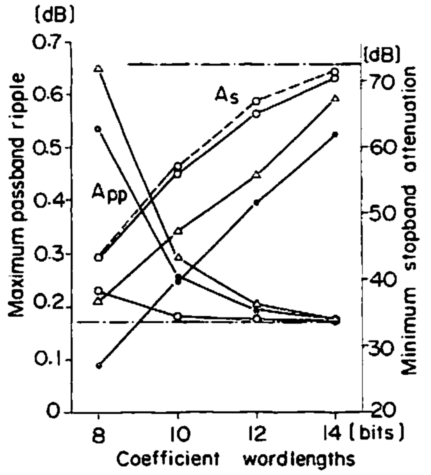

Fig. 1. Frequency response improvements for a 199th-order FIR low-pass filter. Responses for $H(z)$ with infinite precision coefficients are shown by dashed and dotted lines.

$$N_{mult}(E) = N_{Add}(E) = P^L[N_F/(K - K')]\,(M + 2)K. \quad (12)$$

This number is extremely smaller than that required in direct calculation of E and frequency domain evaluation.

IV. Design Examples

A. Filter Responses

The proposed approach was examined by using a 199th-order FIR low-pass filter having 0.17 dB ($p - p$) ripple in the passband 0–0.125 Hz, and 72.5 dB attenuation in the stopband 0.14–0.5 Hz. Sampling frequency used is 1 Hz. The error-free function is the following 4th-order function.

$$W(z) = (1 + 2z^{-1} + z^{-2})(1 + z^{-1} + z^{-2}). \quad (13)$$

$H(z)$ is first approximated through the Remez-exchange algorithm [12]. Differences between approximated filter responses with transfer functions $H(z)$ and $W(z)\,F(z)$ are very small and are mostly negligible. Furthermore, max $\{f_n\}$ is increased from max $\{h_n\}$ by about 15 percent. This is equivalent to improving coefficient wordlengths by 0.2 bit, which is also small.

B. Design Parameters

The filter response error is evaluated by the maximum deviation for an amplitude response. In order to shape $|\Delta F(e^{jw})|$ so that it is approximately proportional to $|1/W(e^{jw})|$, the modified weighting function is chosen to be

$$W^*(z) = (1 + 2z^{-1} + z^{-2})^2. \quad (14)$$

The $F(z)$ coefficients are rounded off into 8, 10, 12, and 14 bits, which do not include a sign bit. The maximum coefficient value is normalized to unity. The number of grids assigned to Δf_n is three or seven.

C. Filter Response Improvements

Maximum passband ripple ($p - p$) App and minimum stopband attenuation As are shown in Fig. 1. Lines indicated by ⊗, △, ○ correspond to $H_Q(z)$, $W(z)F_Q(z)$, and $W(z)F_{QO}(z)$, respectively. $H_Q(z)$ and $F_Q(z)$ have only rounded off coefficients. $F_{QO}(z)$ has the discretely optimized coefficients. Solid and dashed lines correspond to the number of grids for three and seven, respectively. Fig. 1 shows the proposed algorithm sufficiently improves filter responses both in the passband and stopband. These improvements can be rephrased as coefficient wordlengths are reduced by about three bits.

D. Computing Time

The execution time required in the discrete optimization procedure with seven grids was 97 s, using a general purpose computer (NEC ACOS 900). This result obviously allows using the proposed method for high-order FIR filters.

V. Conclusion

A computationally efficient discrete optimization algorithm for high-order FIR filters has been proposed. Through a design example for a 199th-order FIR filter, coefficient wordlength reduction of three bits is obtained with a relatively short computing time.

References

[1] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.
[2] M. Suk and S. K. Mitra, "Computer-aided design of digital filters with finite word lengths," *IEEE Trans. Audio Electroacoust.*, vol. AU-20, pp. 356-363, Dec. 1972.
[3] E. Avenhaus, "On the design of digital filters with coefficients of limited word length," *IEEE Trans. Audio Electroacoust.*, vol. AU-20, pp. 206-212, Aug. 1972.
[4] C. Charalambous and M. J. Best, "Optimum of recursive digital filters with finite word length," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, pp. 424–431, Dec. 1974.
[5] K. Steiglitz, "Designing short-word recursive digital filters," in *Proc. Annu. Allerton Conf. Circuit Syst. Theory*, Oct. 1971, pp. 778-788.
[6] F. Brglez, "Digital filter design with short word-length coefficients," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 1044-1050, Dec. 1978.
[7] Y. Chen, S. M. Kang, and T. G. Marshall, "The optimal design of CCD transversal filters using mixed-integer programming techniques," in *Proc. ISCAS*, May 1978, pp. 748-751.
[8] D. M. Kodek, "Design of optimal finite wordlength FIR digital filters using integer programming techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 304-307, June 1980.
[9] V. B. Lawrence and A. C. Salazar, "Finite precision design of linear-phase FIR filters," *Bell Syst. Tech. J.*, vol. 59, pp. 1575-1598, Nov. 1980.
[10] D. Kodek and K. Steiglitz, "Comparison of optimal and local search methods for designing finite wordlength FIR digital filters," *IEEE Trans. Circuits Syst.*, vol. CAS-28, pp. 28-32, Jan. 1981.
[11] K. Nakayama, "A discrete optimization method for high-order FIR filters with finite wordlength coefficients," in *Proc. ICASSP'82*, 1982, pp. 484-487.
[12] T. W. Parks and J. H. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," *IEEE Trans. Commun. Technol.*, vol. CT-19, pp. 189-194, Oct. 1971.

0096-3518/87/0800-1217$01.00 © 1987 IEEE