Design of Cascade-Form IIR Filters with Powers-of-Two Coefficients Using Mixed Integer Linear Programming

Hyuk Jun Oh, Woo Jin Oh, and Yong Hoon Lee

Department of Electrical Engineering
Korea Advanced Institute of Science and Technology
373-1 Kusong-dong, Yusong-gu, Taejon, 305-701 KOREA
E-MAIL : yohlee@eekaist.kaist.ac.kr

Abstract - We propose an optimal design method for digital IIR filters with powers-of-two coefficients. This method is based upon the formulation of a linear optimization problem that minimizes filter's complexity for given specifications. It is shown that by taking the logarithm to the transfer function of cascade-form IIR filters, the design problem becomes linear and can be solved by mixed integer linear programming (MILP). Design examples are presented to demonstrate the efficiency of the proposed method.

I. INTRODUCTION

One approach to multiplier free digital filter design is to choose filter coefficients from a discrete set consisting of powers-of-two and sums of a few powers-of-two terms. Design methods based on this approach have been developed for FIR filters [1]-[5] and 2-D IIR filters [6],[7].

In this paper, we consider the design of cascade-form 1-D IIR filters with powers-of-two coefficients. Our objective is to develop an optimal procedure for designing such filters with *minimal* complexity, for the specification at hand. By taking the logarithm to the transfer function of a cascade-form IIR filters, we formulate an optimization problem with a linear objective function and linear constraints. Then this problem is solved by MILP. Design examples demonstrate that this method leads to very efficient cascade-form IIR filters with powers-of-two coefficients, which are particularly suitable for VLSI implementation.

II. PROBLEM FORMULATION

Consider the following cascade-form transfer function

$$H(z) = K\frac{\prod_{p=1}^{P}(1 \pm z^{-1})^{l_p}\prod_{q=1}^{Q}(1 + a_q z^{-1} + z^{-2})^{m_q}}{\prod_{r=1}^{R}(1 + c_r z^{-1} + d_r z^{-2})^{n_r}} \qquad (1)$$

where a_q, c_r, d_r are sum of powers-of-two coefficients and l_p, m_q and n_r are nonnegative integers. To simplify notation, we define $L_p(z) = (1 \pm z^{-1})$, $N_q(z) = (1 + a_q z^{-1} + z^{-2})$ and $D_r(z) = (1 + c_r z^{-1} + d_r z^{-2})$. Since the word length of a_q, c_r and d_r is finite, the number of all possible $N_q(z)$ and $D_r(z)$ are finite. Given the desired filter specifications, we determine sets of eligible polynomials by examining each possible $L_p(z)$, $N_q(z)$ and $D_r(z)$ based on the following criteria: (i) Roots of $L_p(z)$ and $N_q(z)$ should lie in stopbands. (ii) Roots of $D_r(z)$ should be located within the unit circle, and should not lie in stopbands. In (1), the number of eligible $L_p(z)$, $N_q(z)$ and $D_r(z)$, respectively, are denoted by P, Q and R. Our objective is to find optimal l_p, m_q and n_r values for all eligible $L_p(z)$, $N_q(z)$ and $D_r(z)$. Note that all zeros of $H(z)$ lie on the unit circle. This fact leads to efficient implementation at the expense of some limitation in reducing passband ripple size [8].

Let α_q and β_r be the number of adders required to implement the polynomials $N_q(z)$ and $D_r(z)$, respectively. For example, when the coefficients a_1, c_1 and d_1 are sum of *two* powers-of-two terms, we get $\alpha_1 = 3$, $\beta_1 = 4$. Note that the number of adders for realizing $L_p(z)$ is always one. Now an IIR filter, represented by (1), with minimal complexity may be designed by solving the following

0-7803-3073-0/96/$5.00 ©1996 IEEE

the following optimization problem. For a given passband ripple δ_p and stopband ripple δ_s:

$$\begin{aligned}
&\text{Minimize} && v = A_{num} + A_{den} && (\textit{measure of complexity})\\
&\text{Subject to} && \big||H(\omega)| - H_d(\omega)\big| < \delta_p && (\textit{in passbands})\\
& && |H(\omega)| < \delta_s && (\textit{in stopbands})\\
& && \sum_{p=1}^{P} l_p + \sum_{q=1}^{Q} \alpha_q \bullet m_q = A_{num} && (\textit{No. of adders in numerator})\\
& && \sum_{r=1}^{R} \beta_r \bullet n_r = A_{den} && (\textit{No. of adders in denominator})
\end{aligned} \quad (2)$$

where A_{num} and A_{den} denote the total number of adders in the numerator and the denominator, respectively, and $H_d(\omega)$ is the desired magnitude response. This design problem requires some nonlinear optimization technique due to the nonlinearity of $H(\omega)$. However, by taking the logarithm on $H(\omega)$, this can be converted into a linear problem. We define

$$\begin{aligned}
H_{dB}(\omega) &= 20log|H(\omega)|\\
&= K_{dB} + \sum_{p=1}^{P} l_p L_{pdB}(\omega)\\
&\quad + \sum_{q=1}^{Q} m_q N_{qdB}(\omega) - \sum_{r=1}^{R} n_r D_{rdB}(\omega)
\end{aligned} \quad (3)$$

where $K_{dB} = 20log|K|$, $L_{pdB}(\omega) = 20log|L_p(\omega)|$, $N_{qdB}(\omega) = 20log|N_q(\omega)|$ and $D_{rdB}(\omega) = 20log|D_r(\omega)|$. Now the optimization in (2) can be rewritten as

$$\begin{aligned}
&\text{Minimize} && v = A_{num} + A_{den} && (\textit{measure of complexity})\\
&\text{Subject to} && 20log(1 - \frac{\delta_p}{H_d(\omega)}) < H_{dB}(\omega) - 20logH_d(\omega) && (\textit{in passbands})\\
& && H_{dB}(\omega) - 20logH_d(\omega) < 20log(1 + \frac{\delta_p}{H_d(\omega)}) && (\textit{in passbands})\\
& && H_{dB}(\omega) < 20log\delta_s && (\textit{in stopbands})\\
& && \sum_{p=1}^{P} l_p + \sum_{q=1}^{Q} \alpha_q \bullet m_q = A_{num} && (\textit{No. of adders in numerator})\\
& && \sum_{r=1}^{R} \beta_r \bullet n_r = A_{den} && (\textit{No. of adders in denominator})
\end{aligned} \quad (4)$$

This problem can be solved by MILP, treating K_{dB}, l_p, m_q, and n_r as variables. The proposed design algorithm is summarized as follows:

Step 1. Determine the sets of eligible polynomials following the procedure stated below (1).

Step 2. Determine K_{dB}, l_p, m_q, n_r by solving the optimization problem in (4).

III. DESIGN EXAMPLES

To compare the proposed design methods with the previous ones, our method is applied to the filter design problems considered in [9] and [10]. The design problems were solved in a Sparc 20 workstation using the commercial general purpose MILP package in [11].

Example 1(Lowpass filter): The specifications in normalized frequency are:

passband: $F \in [0,\ 0.2]$, stopband: $F \in [0.3,\ 0.5]$

ripple: $\delta_p \le 0.1dB$ in passband,

$\delta_s \ge 33dB$ in stopband.

In [9] an IIR filter with 6 bits coefficients, requiring multipliers for its implementation, was designed under these specifications. In our design, we assume that filter coefficients are sum of two powers-of-two terms and that each powers-of-two term consists of 6 bits including sign bit. The number of eligible polynomials are: P=1, Q=10, R=280. The design time was approximately 10 minutes. Fig. 1 shows the magnitude response of the resulting filter, and Table I compares the results of the proposed method with those of [9]. It is seen that our design reduced 15 multipliers and 2 adders, while resulting in smaller passband ripple and greater stopband attenuation. The proposed technique, which leads to multiplier-free implementation, can yield better and simpler IIR filters than can the method in [9] which requires multipliers.

Example 2(Bandpass filter): The specifications in normalized frequency are:

passband: $F \in [0.125,\ 0.175]$,

stopband: $F \in [0,\ 0.105] \cup [0.195,\ 0.5]$

ripple: $\delta_p \le 0.2dB$ in passband,

$\delta_s \ge 40dB$ in stopband.

In [10] an IIR filter employing cyclotomic numerator polynomials, which leads to multiplierless numerator implementation, and denominator polynomials with infinite precision coefficients, was designed under these specifications. In this case, we assume that filter coefficients are sum of *three* powers-of-two terms. Again, each powers-of-two term consists of 6 bits. The number of eligible polynomials are: P=2, Q=31, R=142. The design time was approximately 15 minutes. Fig. 2 shows the magnitude response of the resulting filter, and the last two columns of Table I compares the results of the proposed method with those in [10]. Our design eliminated the need of 9 infinite precision multipliers at the expense of 17 additional adders. The proposed technique leads to simpler IIR filters requiring only shifters and 6 bits adder than can the method in [10] which requires infinite precision multipliers and adders.

REFERENCES

[1] Y. C. Lim and S. R. Parker, "FIR filter design over a discrete Powers-of-Two coefficient space," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 583-591, June 1983.

[2] Y. C. Lim and A. G. Constantinides, "Linear phase FIR digital filter without multipliers," in Proc. *IEEE Int. Symp. on Circuits and Systems*, Tokyo, Japan, pp. 185-188, July 1979.

[3] H. Samueli, "An improved search algorithm for the design of multiplierless FIR filters with Powers-of-Two coefficients," *IEEE Trans. Circuits and Syst.*, vol. 36, pp. 1044-1047, July 1989.

[4] Y. C. Lim and B. Liu, "Design of cascade form FIR filters with discrete valued coefficients," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1735-1739, Nov. 1988.

[5] H. K. Kwan and C. L. Chan, "Circularly symmetric two-dimensional FIR digital filter design using an enhanced McClellan transformation," *Proc. Inst. Elect. Eng., pt G, Circuits, Devices and Systems*, vol. 136, no. 3, pp. 129-134, June 1989.

[6] S. Varoufakis and A. Venetsanopoulos, "A special class of multiplierless 2-D recursive digital filters," *J. Franklin Inst.*, vol. 318, pp. 105-121, Aug. 1984.

[7] H. K. Kwan and C. L. Chan, "Design of multiplierless spherically symmetric and constant group delay recursive digital filters with Powers-of-Two-coefficients," *IEEE Trans. Circuits and Syst.*, vol. 37, pp. 1027-1035, Aug. 1990.

[8] T. Saramaki, "Design of optimum recursive digital filters with zeros on the unit circle," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 31, pp. 450-458, Apr. 1983.

[9] J. Radecki, J. Konrad and E. Dubois, "Design of finite-wordlength IIR filters with prescribed magnitude, group delay and stability properties using simulated annealing," *Proc. IEEE Int. Conf. Acoust. Speech Signal Processing.*, pp. 1637-1640, May 1991.

[10] R. J. Hartnett, L. B. Jackson, and G. F. Boudreaux-Bartels, "IIR filters with reduced mutipliers using cyclotomic polynomial numerators," *Proc. IEEE Int. Conf. Acoust. Speech Signal Processing.*, Vol. IV, pp. 321-324, 1992.

[11] *Using the CPLEX callable library and CPLEX mixed integer library*, CPLEX optimization, Inc. Incline Village, NV, 1994.

Table I. Comparison and Summary of the Examples

Examples	Lowpass filter		Bandpass filter	
	[9]	Proposed	[10]	Proposed
No. of adder	12	10	16	33
No. of multiplier	15	-	9	-
Passband ripple(dB)	0.09	0.082	0.2	0.18
Stopband attenuation(dB)	33.58	38.7	42.1	40.5

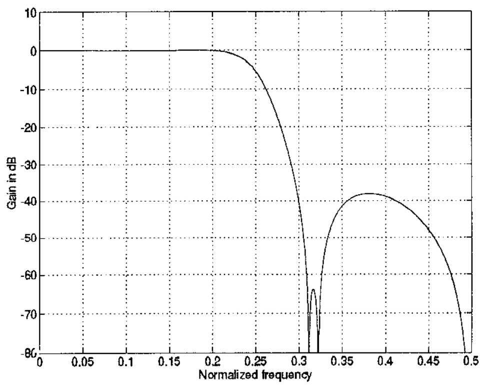

Figure 1. Magnitude response of
the designed filter of example 1.

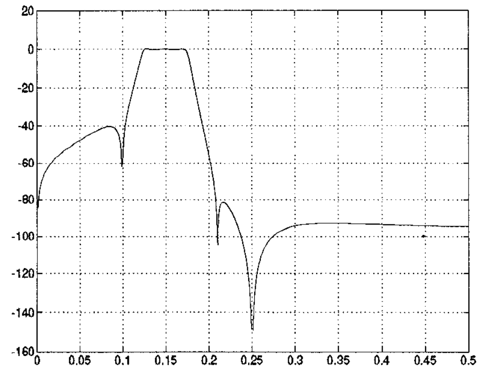

Figure 2. Magnitude response of
the designed filter of example 2.