OPTIMAL DESIGN METHOD FOR FIR FILTER WITH DISCRETE COEFFICIENTS BASED ON INTEGER SEMI-INFINITE LINEAR PROGRAMS

Rika ITO, Kenji SUYAMA and Ryuichi HIRABAYASHI

Faculty of Engineering, Science University of Tokyo
rika@ms.kagu.sut.ac.jp, suyama@ms.kagu.sut.ac.jp, hira@ms.kagu.sut.ac.jp

ABSTRACT

The purpose of the paper is to propose a new design method of FIR filters with discrete coefficients considering optimality. In the proposed method, the design problem of FIR filters is formulated as a Mixed Integer Semi-Infinite Linear Programming problem (MISILP), which can be solved by a branch and bound technique. Then, it is possible to obtain the optimal discrete coefficients, and the optimality of the obtained solution can be guaranteed. It was confirmed that optimal coefficients of linear phase FIR filter with discrete coefficients could be designed in reasonable computational time with sufficient precision based on the results of computational experiments.

1. INTRODUCTION

In hardware implementation of FIR filters, the filter coefficients corresponding to multiplier coefficients are presented as the finite word length numbers. When the coefficients are simply rounded to the nearest discrete number, performance of filters are degraded from the one with the optimal real coefficients. Therefore, design methods of FIR filters with discrete coefficients have been widely researched[1]$\sim$[4]. Among various methods, some methods used the branch and bound (B & B) technique based on LP or Remez algorithm. For example, Cho and Lee [3] proposed the B & B technique based on LP focusing only on active constraints to decrease the computational time. In [8], the design algorithm based on LP, which has only finite constraints, was proposed. However, optimality of the solution obtained by the algorithm cannot be assured because of the finite constraints.

In this paper, we propose a new design method of linear phase FIR filters with discrete coefficients, which guarantees the optimality of the solution obtained. In the proposed method, the design problem is formulated as a Mixed Integer Semi-Infinite Linear Programming problem (MISILP) which has infinite constraints and solved by B & B technique. In the B & B technique, a branching tree is generated and, on each node, it is necessary to solve Semi-Infinite Linear Programming problem (SILP)[5]. Here, SILP is a linear programming problem that has infinitely many constraints. SILP is solved by using 3 Phase method [6].

It is shown that the results of some computational experiments can be certified the performance of the proposed method.

2. PROBLEM FORMULATION

The transfer function of an FIR filter with length $N+1$ is denoted as

$$H(z) = \sum_{k=0}^{N} h_k z^{-k}. \tag{1}$$

When h_k, $k = 0, 1, \ldots, N$ is the even symmetric impulse response and, N is an even number, the linear phase characteristic with $N/2$ delay is achieved. Then, the magnitude response $H(\omega)$ can be expressed as

$$H(\omega) = \sum_{n=0}^{N} a_n \cos n\omega. \tag{2}$$

Suppose, a desired response $D(\omega)$ is given as follows

$$D(\omega) = \begin{cases} K, & 0 \le \omega \le \omega_p, \\ 0, & \omega_s \le \omega \le \pi. \end{cases} \tag{3}$$

Where K is a scaling factor, ω_p is the passband cutoff frequency, and ω_s is the stopband cutoff frequency, respectively. Then, the optimization problem to approximate $H(\omega)$ to $D(\omega)$ in a min-max sense can be written as

$$\min_{a_0,\ldots,a_N} \max_{\omega\in\Omega} |D(\omega) - H(\omega)|. \tag{4}$$

where Ω is the approximation band,

$$\Omega = [0, \omega_p] \cup [\omega_s, \pi]. \tag{5}$$

If we introduce a new variable δ' that corresponds to the L_∞-approximation error, it is easy to convert the above min-max problem to the following minimization problem:

$$\begin{array}{lrcll} \min & \delta' & & & \\ \text{sub.to} & H(\omega) + \delta' & \ge & D(\omega), & \omega \in \Omega, \\ & -H(\omega) + \delta' & \ge & -D(\omega), & \omega \in \Omega. \end{array} \tag{6}$$

Now, we assume that coefficients a_n $(i = 0 \dots N)$ are limited to discrete coefficients of p bit. Then, (6) can be re-formulated as a MILSIP as follows:

$$\begin{array}{lrl} \min & \delta & \\ \text{sub.to} & \tilde{H}(\omega) + \delta & \geq 2^p D(\omega), \omega \in \Omega, \\ & -\tilde{H}(\omega) + \delta & \geq -2^p D(\omega), \omega \in \Omega \\ & x_0, \dots, x_N & \geq -2^p, \\ & -x_0, \dots, -x_N & \geq -(2^p - 1), \\ & x_i \in \mathbf{Z}, & i = 0, \dots, N, \end{array} \quad (7)$$

where,

$$\begin{aligned} \tilde{H}(\omega) &= 2^p H(\omega) && (8) \\ &= \sum_{n=0}^{N} 2^p a_n \cos n\omega && (9) \\ &= \sum_{n=0}^{N} x_n \cos n\omega, && (10) \\ x_i &= 2^p a_i, \ \ i = 0, \dots, N, && (11) \\ \delta &= 2^p \delta'. && (12) \end{aligned}$$

While the ω in (7) is treated as a continuous variable, that of MILP is discretized. Therefore, the optimality of the obtained solution using MILP is not always assured. It is impossible to solve (7) directly. Hence, we adopt the B & B technique in the next section. This technique is constructed with the branching process and the bounding process. At first, we ignore the integer constraints, and we get an SILP. To solve the SILP, we need the dual problem of the SILP [5]:

$$\begin{array}{ll} \max & \sum_{\omega \in (\Omega)} 2^p D(\omega) y_1(\omega) - \sum_{\omega \in (\Omega)} 2^p D(\omega) y_2(\omega) \\ & -\sum_{i=0}^{N} 2^p y_{i+3} - \sum_{i=0}^{N} (2^p - 1) y_{i+N+4} \\ \text{sub.to} & \sum_{\omega \in (\Omega)} \boldsymbol{a}_1(\omega) y_1(\omega) + \sum_{\omega \in (\Omega)} \boldsymbol{a}_2(\omega) y_2(\omega) \\ & +\sum_{i=0}^{N} \boldsymbol{e}_{i+1} y_{i+3} - \sum_{i=0}^{N} \boldsymbol{e}_{i+1} y_{i+N+4} \\ & = \boldsymbol{e}_{N+2}, \\ & y_1(\omega), y_2(\omega) \geq 0, \ \forall \omega \in \Omega, \\ & y_3, \dots, y_{2N+4} \geq 0. \end{array} \quad (13)$$

Where, $\sum_{\omega \in (\Omega)}$ means that the summation is taken over only finite numbers of dual variables $y_k(\omega) > 0$, $k = 1, 2$, and all the remaining $y_k(\omega)$, $k = 1, 2$ are zero. Also, $\boldsymbol{a}_1 = (1, \cos\omega, \dots, \cos N\omega, 1)^T$, $\boldsymbol{a}_2 = (-1, -\cos\omega, \dots, -\cos N\omega, 1)^T \in I\!R^{N+2}$ and $\boldsymbol{e}_{i+1}, i = 0, \dots, N$ are the $(i+1)$-th unit vectors in $I\!R^{N+2}$ and $\boldsymbol{e}_{N+2}$ is also the $(N+2)$-th unit vector in $I\!R^{N+2}$.

3. 3 PHASE METHOD FOR SOLVING SILP

In this section, we describe briefly how to solve SILP by means of 3 Phase method since it is very tedious to describe the method exactly.

By use of the Carathéodory's theorem [5], we can show that there is an optimal solution for (13) that has at most $N + 2$ positive dual variables. We assume that $y_1(\omega_1), \dots, y_1(\omega_{k_1}), y_2(\omega_{k_1+1}), \dots, y_2(\omega_{N+2})$ are positive in the optimal solution. In order to simplify, we set $y'_i = y_1(\omega_i), i = 1, \dots, k_1$ and $y'_i = y_2(\omega_i), i = k_1 + 1, \dots, N + 2$. Then, we have the following equations.

$$\sum_{i=1}^{k_1} \boldsymbol{a}_1(\omega_i) y'_i + \sum_{i=k_1+1}^{N+2} \boldsymbol{a}_2(\omega_i) y'_i = \boldsymbol{e}_{N+2}. \quad (14)$$

When we notice the complementary slackness theorem of SILP [5], we have the following complementarity for the primal optimality solution $x_0, \dots, x_N, \delta$:

$$\begin{aligned} y'_i \left[\tilde{H}(\omega_i) + \delta - 2^p D(\omega_i)\right] &= 0, i = 1, \dots, k_1, \\ y'_i \left[-\tilde{H}(\omega_i) + \delta + 2^p D(\omega_i)\right] &= 0, i = k_1 + 1, \dots, N + 2. \end{aligned}$$

Noticing $y_i > 0$, then we have the following equations:

$$\begin{aligned} \tilde{H}(\omega_i) + \delta &= 2^p D(\omega_i), \ \ i = 1, \dots, k_1, && (15) \\ -\tilde{H}(\omega_i) + \delta &= -2^p D(\omega_i), i = k_1 + 1, \dots, N + 2. && (16) \end{aligned}$$

For each $\omega_i, i = 1, \dots, k_1$, (15) holds, and in a neighbourhood of ω_i, the inequality $\tilde{H}(\omega_i) + \delta > 2^p D(\omega_i)$ holds. Hence, each ω_i, $i = 1, \dots, k_1$ is a local minimum of the function $\tilde{H}(\omega_i) + \delta$. Therefore, if ω_i is in the interior of Ω, we have the following equations:

$$\frac{\partial}{\partial \omega} \left[\tilde{H}(\omega_i) + \delta\right]_{\omega = \omega_i} = 0, \ i = 1, \dots, k_1. \quad (17)$$

If ω_i is on the boundary of Ω, we have to describe the Karush-Kuhn-Tucker condition on ω_i[7]. For ω_i, $i = k_1 + 1, \dots, N + 2$, by a similar discussion, we have :

$$\frac{\partial}{\partial \omega} \left[-\tilde{H}(\omega_i) + \delta\right]_{\omega = \omega_i} = 0, i = k_1 + 1, \dots, N + 2. \quad (18)$$

Now, we have $3(N + 2)$ variables $x_0, \dots, x_N, \delta, y'_1, \dots, y'_{N+2}, \omega_1, \dots, \omega_{N+2}$ and $3(N + 2)$ equations (14), (16), (17), (18). Hence, we can solve the primal SILP (7) and its dual (13) simultaneously if we can solve equations (14), (16), (17), (18). In this paper, the Newton method is used to solve the nonlinear equations. It is widely known that when we use Newton method, we need a good initial solution that approximates an optimal solution well. For obtaining a good initial solution, we discretize the primal and dual probrems:

$$\begin{array}{lrl} \min & & \delta \\ \text{sub.to} & \tilde{H}(\omega_i)+\delta & \geq 2^p D(\omega_i), \\ & & i=1,\dots,q_1, \\ & -\tilde{H}(\omega_i)+\delta & \geq -2^p D(\omega_i), \\ & & i=q_1+1,\dots,q_2, \\ & x_0,\dots,x_N & \geq -2^p, \\ & -x_0,\dots,-x_N & \geq -(2^p-1), \end{array} \tag{19}$$

and

$$\begin{array}{ll} \max & \sum_{i=1}^{q_1} 2^p D(\omega_i) y_1(\omega_i) - \sum_{i=q_1+1}^{q_2} 2^p D(\omega_i) y_2(\omega_i) \\ & -\sum_{i=0}^{N} 2^p y_{i+3} - \sum_{i=0}^{N} (2^p-1) y_{i+N+4} \\ \text{sub.to} & \sum_{i=1}^{q_1} \boldsymbol{a}_1(\omega_i) y_1(\omega_i) + \sum_{i=1}^{q_2} \boldsymbol{a}_2(\omega_i) y_1(\omega_i) \\ & +\sum_{i=0}^{N} \boldsymbol{e}_{i+1} y_{i+3} - \sum_{i=0}^{N} \boldsymbol{e}_{i+1} y_{i+N+4} \\ & = \boldsymbol{e}_{N+2}, \\ & y_1(\omega_i) \geq 0,\ i=1,\dots,q_1, \\ & y_2(\omega_i) \geq 0,\ i=q_1+1,\dots,q_2, \\ & y_3,\dots,y_{2N+4} \geq 0. \end{array} \tag{20}$$

Problem (19), (20) are linear programming problems and can be solved by the simplex method. It is much easier to solve the problem (20) than (19). When we get a dual optimal solution $\bar{\boldsymbol{y}}$ of problem (20), it is an easy exercise to obtain a primal optimal solution $\bar{\boldsymbol{x}}, \bar{\delta}$ by using the duality theory of linear programming. Now we can use $(\bar{x}, \bar{\delta}, \bar{y}, \omega_1, \dots, \omega_{q_2})$ as the initial solution for the Newton method. This phase is called Phase 1.

For solving MISILP by using B & B technique, we have to solve many SILP subproblems with additional constraints. In that case, it might be happened that some x_i is zero in a primal optimal solution. Then, the numbers of optimal dual variables that have positive value becomes less than $N+2$. However, the optimal solution for the discretized dual problem has $N+2$ basic dual variables. Hence, it is necessary to reduce the number of dual variables. For that, at first, we delete all the dual basic variables that are zero. Next, for all pair (i,j) that ω_i and ω_j are very close and $\bar{y}'_i$, $\bar{y}'_j > 0$ holds, we set

$$y'_i := y'_i + y'_j, y_j := 0, \omega_i = (\omega_i + \omega_j)/2. \tag{21}$$

This phase is called Phase 2.

Now, we have the initial solution for the Newton method. In 3 Phase, we solve equations (14), (16), (17), (18) and obtain an optimal solution for (7). 3 Phase is also called as local reduction method.

In the following, 3 Phase method is described. Here SILP is formulated as (22) that has some additional constraints which corresponds to subproblems in the B & B technique described in the next section.

$$\begin{array}{lrl} \min & & \delta \\ \text{sub.to} & \tilde{H}(\omega)+\delta & \geq 2^p D(\omega), \omega \in \Omega, \\ & -\tilde{H}(\omega)+\delta & \geq -2^p D(\omega), \omega \in \Omega, \\ & x_0,\dots,x_N & \geq -2^p, \\ & -x_0,\dots,-x_N & \geq -(2^p-1), \\ & x_{j_i} \geq d_i, & i=1,\dots,\ell, \\ & -x_{k_i} \geq -f_{ij}, & i=1,\dots,m, \end{array} \tag{22}$$

where $x_{j_i}, x_{k_i} \in \{x_0,\dots,x_N\}$. The daul problem of the above (22) is:

$$\begin{array}{ll} \max & \sum_{\omega\in(\Omega)} 2^p D(\omega) y_1(\omega) - \sum_{\omega\in(\Omega)} 2^p D(\omega) y_2(\omega) \\ & -\sum_{i=0}^{N} 2^p y_{i+3} - \sum_{i=0}^{N} (2^p-1) y_{i+N+4} \\ & -\sum_{i=1}^{\ell} d_i y_{i+2N+4} - \sum_{i=1}^{m} f_i y_{i+2N+\ell+4} \\ \text{sub.to} & \sum_{\omega\in(\Omega)} \boldsymbol{a}_1(\omega) y_1(\omega) + \sum_{\omega\in(\Omega)} \boldsymbol{a}_2(\omega) y_2(\omega) \\ & +\sum_{i=0}^{N} \boldsymbol{e}_{i+1} y_{i+3} - \sum_{i=0}^{N} \boldsymbol{e}_{i+1} y_{i+N+4} \\ & +\sum_{i=1}^{\ell} \boldsymbol{e}_{j_i+1} y_{i+2N+4} - \sum_{i=1}^{m} \boldsymbol{e}_{k_i+1} y_{i+2N+\ell+4} \\ & +\sum_{i=1}^{\ell} \boldsymbol{e}_{j_i+1} y_{i+2N+4} - \sum_{i=1}^{m} \boldsymbol{e}_{k_i+1} y_{i+2N+\ell+4} \\ & = \boldsymbol{e}_{N+2}, \\ & y_1(\omega), y_2(\omega) \geq 0,\ \forall \omega \in \Omega, \\ & y_3,\dots,y_{2N+\ell+m+4} \geq 0. \end{array} \tag{23}$$

4. A NEW DESIGN METHOD USING BRANCH AND BOUND TECHNIQUE

Our aim is to solve MISILP (7), but it is impossible to solve (7) directly. Hence, we solve SILP ignoring the integer constraints. However, since SILP is a continuous optimization problem, an optimal solution obtained is not always an integer solution. A standard technique for solving this difficulty is to exploit the B & B technique.

If there are some $\bar{x}_i$'s that are not integers, then select one non-integer variable x_j and generate two subproblems, which one has an additional constraint $-x_j \geq -\lfloor \bar{x}_j \rfloor$ and the other has an additional constraints $x_j \geq \lceil \bar{x}_j \rceil$. Notice here, that the two generated subproblems are also SILP and can be solved by 3 Phase method. We can continue this procedure and call this process as branching process.

If we continue the branching process, then after finite iterations, we can obtain an integer solution. The obtained integer solution is an optimal solution for the subproblem and a feasible solution for MISILP (7), but might not be optimal for MISLIP. However, we can use the objective function value that corresponds to the integer solution as an upper bound for MISILP (incumbent value) since we can fathom subproblems that have the optimal value greater than or equal to the upper bound. This is true, because, if we add some additional constraints, the optimal value of the subproblem becomes always bigger. The process that we fathom all subproblems which have greater optimal value than the incumbent value is called the bounding process.

5. COMPUTATIONAL EXPERIMENTS

We executed some computational experiments to certify the performance of the proposed filter design method. We set $\omega_p = 2/5\,\pi$, $\omega_s = 4/7\,\pi$. Two kinds of computational experiments were performed.

(a)The scaling factor is fixed to $K = 1$. The bit length p was set from 3 to 10 with pitch 1, and the filter order was fixed to $N = 3, 4, \ldots, 20$ for each value of p.

(b)We fixed $p = 6, 7, 8$ and $N = 9, 10, 11, 12$. Then, the scaling factor K was changed from 0.5 to 2.0 with pitch 0.1.

The result of experiment (a) for $N = 12$, $p = 3, \ldots, 10$ is shown in figure 1. In this figure, it was shown that the optimal value decreased slowly for p over 7bit, on the other hand, the computational time increased rapidly. Therefore, we can attain fast with the enough approximation with only 6bit word-length.

Figure 2 shows the result of experiment (b) for $p = 5$, $N = 12$, $K = 0.5 \ldots 2.0$. From this results it was indicated that the appropriate value of K is from 1.1 to 1.5.

Figure 3 shows the (A)magnitude response for $p = 4$, $N = 12$, $K = 1$ and the (B)magnitude response using the coefficients which are simply rounded to 4bit.

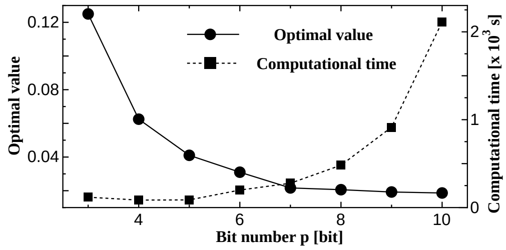

Fig. 1. Experimental result for $p = 3 \ldots 10, N = 12$

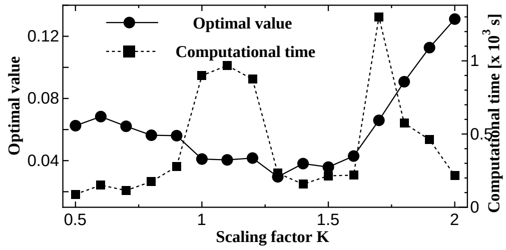

Fig. 2. Experimental result for $p = 5, N = 12, K = 0.5 \ldots 2.0$

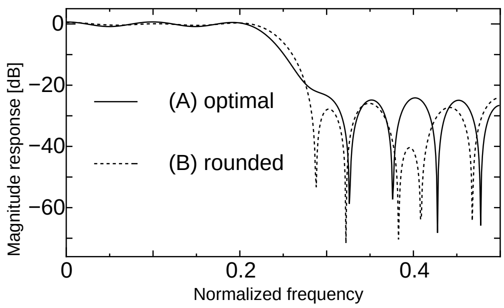

Fig. 3. Magnitude response for (A) $p = 4, N = 12, K = 1$ and (B) with coefficients rounded to $p = 4$

6. CONCLUSION

In this paper, we proposed a new design method for FIR filters with discrete coefficients using B & B technique. In this method, by formulating the FIR filter design problem to MISILP, we can guarantee the optimality of the obtained filter coefficients. The computational experiments showed that the proposed method performed the enough approximation even for only 6bit word-length with the reasonable computational time.

7. REFERENCES

[1] Young-Woo Kim, Young-Mo Yang, Joe-Tack Yoo, and Soo-Won Kim, "Low-Power Digital Filtering Using Approximate Processing with Variable Canonic Signed Digit Coefficients," Proc. ISCAS 2000, pp.337-340, 2000.

[2] Y. C. Lim and S. R. Parker, "FIR Filter Design Over a Discrete Powers-of-Two Coefficient Space," IEEE Trans. ASSP., Vol.31, No.3, pp.583–591, 1983.

[3] Nam IK Cho and Sang Uk Lee, "Optimal Design of Finite Precision FIR Filters Using Linear Programming with Reduced Constraints," IEEE Trans. SP., Vol.46, No.1, pp.195-199, 1998.

[4] Pierre Siohan and Christian Roche, "Optimal Design Of 1-D And 2-D FIR Multiplierless Filters," Proc. ICASSP91, pp.2877-2880, 1991.

[5] Goberna, M. A. and M. A. Lopéz, "Linear Semi-infinite Optimization," Wiley Series in Mathematical Methods in Practice 2, John Wiley & Sons, Chichester, 1998.

[6] Polak, E., "Optimization, -Algorithms and Consistent Approximations," Applied Mathematical Sciences 124, Springer-Verlag, New York, 1997.

[7] Glashoff, K and A-Å. Gustafson, "Linear Optimization and Approximation," Applied Mathematical Sciences 45, Springer-Verlag, New York, 1983.

[8] Y.C.Lim, "Design of Discrete-Coefficient-Value Linear Phase Filters with Optimum normalized Peak Ripple Magnitude," IEEE Trans. CAS., Vol.37, pp.1480-1486, 1990.