Design for IIR Digital Filters with Discrete Coefficients Using Weighted Modified Least-Square Criterion

Masayoshi Nakamoto, Yuuki Hayakawa, Yutaka Kurumaji and Takao Hinamoto
Graduate School of Engineering, Hiroshima University
1-4-1 Higashi-Hiroshima 739-8527, Japan
E-mail : masayoshi@ieee.org

Abstract: In this paper, we formulate the filter design problem based on modified least square criterion in the frequency domain. Here the problem is expressed as quadratic form and we optimize the filter coefficients by using branch and bound method in discrete space. In the branch and bound method, we calculate the relaxation solution based on the Lagrange multiplier method and search in the domain in which good possible solutions exist. In this method, the stable solutions whose poles are all inside the unit circle are just acceptable, so the stability of the filter can be guaranteed. Finally, numerical example is given to illustrate the utility of the proposed method.

1. Introduction

When designing the digital filters, it is general to assume that the filter coefficients are expressed as continuous values (not discrete values). However, when implementing the digital filter, the filter coefficients will be changed and the frequency characteristics may be degraded. Furthermore, the filter may be unstable.

It is well known that the synthesis of low-sensitivity digital filter structures based on the coordinate transformation is an effective method to reduce the degradation. However this method is only available when the filter is described by the state-space model. Thus when the filter is described by the rational transfer function, the coordinate transformation can not be used. Design of the filter coefficients in discrete space is an alternative. However, if we search the filter coefficients in discrete space, it will be time consuming. Hence, the design of digital filter with discrete coefficients is really hard. The work for the design of digital filter coefficients with finite word-length has been done [1]-[3]. In [3], the modified least squared (MLS) error criterion [4] is used in the time domain.

In this paper, we propose a design technique for approximating the given frequency characteristic with the finite word-length coefficients. Here the weighted modified least-square (WMLS) criterion is derived in the frequency domain. Using WMLS criterion, the design problem can be expressed as a quadratic form and the finite word-length coefficients can be designed based on the branch and bound method. The relaxation solution can be calculated by the Lagrange multiplier method. Also, in this method, the coefficients whose poles are all inside the unit circle are acceptable. Hence the stability of the filter is guaranteed. Using the branch and bound method, the search time (calculation cost) can be reduced.

Finally, we design the 10-th order low-pass digital filters by the proposed method and show the magnitude response and the phase characteristic of digital filter.

2. Problem Formulation

2.1 Weighted Modified Least-Square Criterion

Let the transfer function be

$$H(\omega) = \frac{B(\omega)}{A(\omega)} \tag{1}$$

where

$$A(\omega) = \sum_{l=0}^{n} a_l e^{-jl\omega},\ a_0 = 1 \tag{2}$$

$$B(\omega) = \sum_{k=0}^{m} b_k e^{-jk\omega}. \tag{3}$$

Here define a weighted estimation function as

$$E = \int_0^{\pi} W(\omega)\big|B(\omega) - H_d(\omega)A(\omega)\big|^2 d\omega \tag{4}$$

where $W(\omega)$ is the weighting function and $H_d(\omega)$ is the desired response. The purpose of this paper is to design the lowpass filter. So the desired response $H_d(\omega)$ is given as

$$H_d(\omega) = \begin{cases} e^{-j\tau\omega} & 0 \le \omega \le \omega_p \\ 0 & \omega_s \le \omega \le \pi \end{cases} \tag{5}$$

where ω_p and ω_s are the passband frequency and stopband frequency, respectively.

Substituting (2),(3) and (5) into (4), we have

$$\begin{aligned} E &= \sum_{k=0}^{m}\sum_{l=0}^{m} a_k a_l P_{k,l} + 2\sum_{k=0}^{m}\sum_{l=0}^{n} a_k b_l Q_{k,l} \\ &+ \sum_{k=0}^{n}\sum_{l=0}^{n} b_k b_l R_{k,l} \\ &= \boldsymbol{a}^T \boldsymbol{P} \boldsymbol{a} + 2\boldsymbol{a}^T \boldsymbol{Q} \boldsymbol{b} + \boldsymbol{b}^T \boldsymbol{R} \boldsymbol{b} \\ &= \begin{bmatrix} \boldsymbol{a}^T & \boldsymbol{b}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{P} & \boldsymbol{Q} \\ \boldsymbol{Q}^T & \boldsymbol{R} \end{bmatrix} \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix} \\ &= \boldsymbol{x}^T \boldsymbol{K} \boldsymbol{x} \end{aligned} \tag{6}$$

where

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix} \qquad \boldsymbol{K} = \begin{bmatrix} \boldsymbol{P} & \boldsymbol{Q} \\ \boldsymbol{Q}^T & \boldsymbol{R} \end{bmatrix}.$$

Here (9)-(11) are expressed as

$$\boldsymbol{a} = \begin{bmatrix} a_0 & a_1 & \cdots & a_m \end{bmatrix}^t, \; a_0 = 1 \tag{7}$$

$$\boldsymbol{b} = \begin{bmatrix} b_0 & b_1 & \cdots & b_n \end{bmatrix}^t \tag{8}$$

$$P_{k,l} = \int_0^{\omega_p} W(\omega) \cos\Big((k-l)\omega\Big) d\omega, \quad k = 0, 1 \cdots m, \; l = 0, 1 \cdots n \tag{9}$$

$$Q_{k,l} = -\int_0^{\omega_p} W(\omega) \cos\Big((k-l+\tau)\omega\Big) d\omega, \quad k = 0, 1 \cdots m, \; l = 0, 1 \cdots n \tag{10}$$

$$R_{k,l} = \int_0^{\omega_p} W(\omega) \cos\Big((k-l)\omega\Big) d\omega + \int_{\omega_s}^{\pi} W(\omega) \cos\Big((k-l)\omega\Big) d\omega, \quad k = 0, 1 \cdots m, \; l = 0, 1 \cdots n. \tag{11}$$

The problem is to find the a_k and b_l which minimizes (4). Note that (4) does not equal to minimize the mean square error(MSE). Hence if we minimize (4), the optimality in the MSE sense is not guaranteed. However the MSE can be reduced effectively by minimizing (4). In this paper, we use the matrix $\boldsymbol{K}$ for the design of IIR digital filter with finite wordlength coefficients.

2.2 Weighting Function

By adjusting the weighting function, the desired specification in terms of passband and stopband can be achieved.

Consider the case when the weighting function in terms of passband is given as follows

$$W(\omega) = \begin{cases} W_1 & \omega_0 = 0 < \omega \le \omega_1 \\ W_2 & \omega_1 < \omega \le \omega_2 \\ \vdots & \\ W_N & \omega_{N-1} < \omega \le \omega_N = \omega_p \end{cases} \tag{12}$$

where $W_1, W_2, \cdots W_N$ are all constants.

Let

$$\alpha_{k,l}^{(i)} = W_i \int_{\omega_{i-1}}^{\omega_i} \cos\Big((k-l)\omega\Big) d\omega. \tag{13}$$

If $k - l \neq 0$, we have

$$\alpha_{k,l}^{(i)} = \frac{W_i}{k-l}\Big(\sin(k-l)\omega_i - \sin(k-l)\omega_{i-1}\Big). \tag{14}$$

Also if $k - l = 0$, we have

$$\alpha_{k,l}^{(i)} = (\omega_i - \omega_{i-1})W_i. \tag{15}$$

Hence

$$P_{k,l} = \begin{cases} \displaystyle\sum_{i=1}^{N} \frac{W_i}{k-l}\Big(\sin(k-l)\omega_i - \sin(k-l)\omega_{i-1}\Big) & \\ & k-l \neq 0 \\ \displaystyle\sum_{i=1}^{N} (\omega_i - \omega_{i-1})W_i & k-l = 0 \end{cases} \tag{16}$$

is obtained.

Next, let $k - l + \tau \neq 0$ and

$$\beta_{k,l}^{(i)} = W_i \int_{\omega_{i-1}}^{\omega_i} \cos\Big((k-l+\tau)\omega\Big) d\omega, \tag{17}$$

we have

$$\beta_{k,l}^{(i)} = \frac{W_i}{k-l+\tau}\Big(\sin(k-l+\tau)\omega_i - \sin(k-l+\tau)\omega_{i-1}\Big). \tag{18}$$

Also, if $k - l + \tau = 0$, we have

$$\beta_{k,l}^{(i)} = (\omega_i - \omega_{i-1})W_i. \tag{19}$$

Hence

$$Q_{k,l} = \begin{cases} \displaystyle -\sum_{i=1}^{N} \frac{W_i}{k-l+\tau}\Big(\sin(k-l+\tau)\omega_i & \\ -\sin(k-l+\tau)\omega_{i-1}\Big) & k-l+\tau \neq 0 \\ \displaystyle -\sum_{i=1}^{N} (\omega_i - \omega_{i-1})W_i & k-l+\tau = 0 \end{cases} \tag{20}$$

is obtained.

The weighting function in terms of stopband is given as follows

$$W(\omega) = \begin{cases} W_{N+1} & \omega_s = \omega_N < \omega \le \omega_{N+1} \\ W_{N+2} & \omega_{N+1} < \omega \le \omega_{N+2} \\ \vdots & \\ W_M & \omega_{M-1} < \omega \le \omega_M = \pi \end{cases} \tag{21}$$

where $W_{N+1}, W_{N+2}, \cdots W_M$ are all constants.

Then $R_{k,l}$ is calculated as

$$R_{k,l} = \begin{cases} \displaystyle\sum_{i=1}^{M} \frac{W_i}{k-l}\Big(\sin(k-l)\omega_i - \sin(k-l)\omega_{i-1}\Big) & \\ & k-l \neq 0 \\ \displaystyle\sum_{i=1}^{M} (\omega_i - \omega_{i-1})W_i & k-l = 0. \end{cases} \tag{22}$$

3. Finite-Wordlength Design

3.1 Lower Bound Estimation

If we try to search the combination of all coefficients values, it is very difficult to search everything based on the tree model as Fig. 1. So we use the Lower bound estimation.

Assume that the temporary solution $\boldsymbol{x}_P$ is obtained. Let the discrete coefficients vector be $\boldsymbol{x}_D$ and divide it into two sets as

$$\boldsymbol{x}_D = \{\boldsymbol{x}_1, \; \boldsymbol{x}_2\} \tag{23}$$

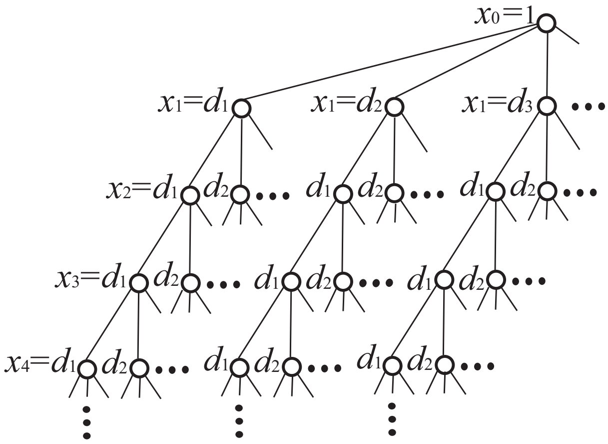

Figure 1. Tree structure of filter coefficients where $d_1, d_2 \cdots$ are candidate value for each coefficient.

where

$$\boldsymbol{x}_1 = (x_1,\ x_2 \cdots x_i) \tag{24}$$
$$\boldsymbol{x}_2 = (x_{i+1},\ x_{i+2} \cdots x_K). \tag{25}$$

Hense $\boldsymbol{x}_i$ indicates a subtree with depth i. So there are a lot of subtrees under $\boldsymbol{x}_1$.

Here the vector in which $\boldsymbol{x}_2$ is relaxed is written as

$$\boldsymbol{x}_D^* = \{\boldsymbol{x}_1,\ \boldsymbol{x}_2^*\}. \tag{26}$$

As mentioned before, $\boldsymbol{x}_1$ indicates a subtree and $x_1, x_2 \cdots x_i$ are fixed as discrete values. The vector which minimizes E under the constraints is $\boldsymbol{x}_D^*$.

If

$$E(\boldsymbol{x}_P) \leq E(\boldsymbol{x}_D^*), \tag{27}$$

a subtree $\boldsymbol{x}_2$ in $\boldsymbol{x}_D^*$ which is better than the temporary solution $\boldsymbol{x}_P$ is not exited. Hence we do not have to search a subtree which is located under $\boldsymbol{x}_1$. Thus we can bound the search space based on the branch and bound method. In this method, to find the good temporary solution quickly is very important to search effectively.

3.2 Lagrange Multiplier Method

We show how to calculate the relaxation solution of (26). Define

$$\boldsymbol{x}^t \boldsymbol{S} = \boldsymbol{p}^t \tag{28}$$

where

$$\boldsymbol{S} = (\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_i)$$
$$\boldsymbol{s}_1 = (1, 0, 0, \ldots, 0)^t$$
$$\boldsymbol{p} = (p_1, p_2, \ldots, p_i)^t.$$

(28) indicates that some elements of $\boldsymbol{x}$ are discrete values.

The relaxation solution can be obtained by the Lagrange multiplier method as follows. Let us apply the Lagrange multiplier method to the problem minimizes (6) under the constraint (28). Let

$$J(\boldsymbol{x}, \boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{x}^t \boldsymbol{K}\boldsymbol{x} - (\boldsymbol{x}^t \boldsymbol{S} - \boldsymbol{p})\boldsymbol{\lambda}. \tag{29}$$

Here differentiate (29) by $\boldsymbol{x}$ and $\boldsymbol{\lambda}$ and equating to $\boldsymbol{0}$, we have

$$\frac{\partial J(\boldsymbol{x}, \boldsymbol{\lambda})}{\partial \boldsymbol{x}} = \boldsymbol{K}\boldsymbol{x} - \boldsymbol{S}\boldsymbol{\lambda} = \boldsymbol{0} \tag{30}$$
$$\frac{\partial J(\boldsymbol{x}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = -\boldsymbol{x}^t \boldsymbol{S} + \boldsymbol{p} = \boldsymbol{0}. \tag{31}$$

From (30) and (31), we have

$$\boldsymbol{\lambda} = (\boldsymbol{S}^t \boldsymbol{K}^{-1} \boldsymbol{S})^{-1} \boldsymbol{p}. \tag{32}$$

Substituting (32) into (30), $\boldsymbol{x}$ can be obtained as

$$\boldsymbol{x} = \boldsymbol{K}^{-1} \boldsymbol{S} (\boldsymbol{S}^t \boldsymbol{K}^{-1} \boldsymbol{S})^{-1} \boldsymbol{p}. \tag{33}$$

Hence by setting $\boldsymbol{S}$ and $\boldsymbol{p}$ with the appropriate values, the relaxation solution can be calculated.

For example, $\boldsymbol{S}$ and $\boldsymbol{x}_1$ which are corresponding to the subtree $T_1 = \{x_1 = 1,\ x_2 = 0.5,\ x_3 = 1.25\}$ are as follows.

$$\boldsymbol{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix} \tag{34}$$
$$\boldsymbol{x}_1 = (1,\ 0.5,\ 1.25)^T. \tag{35}$$

3.3 Search Algorithm

Let the depth of the tree be i and the branch position be k, respectively. Thus, the coefficients are x_i and its candidates are d_k with $i = 1, 2 \cdots K,\ k = 1, 2 \cdots L$ where L is the number of candidates.

STEP 1: Let the initial search point be $x_1 = 1$ and the initial temporary solution be $\boldsymbol{x}_P$.

STEP 2: If all trees are terminated, the search is finished.

STEP 3: If $i \neq m + 1$, go to STEP 4. If $i = m + 1$, check the stability of the filter. If it is stable, go to STEP 4. If not, go to STEP 5.

STEP 4: Calculate the relaxation solution by (33). If $E(\boldsymbol{x}_P) \leq E(\boldsymbol{x}_D^*)$, go to STEP 5. If not, let $i \to i + 1$, and go to STEP 6 (See Fig. 2(a)).

STEP 5: If $x_i \neq d_L$, terminate the subtree under $x_i = d_k$ and move to the next branch $x_i = d_{k+1}$ (See Fig. 2(b)). If $x_i = d_L$, $i \to i - 1$ and move to $x_{i-1} = d_k \to x_{i-1} = d_{k+1}$ (Fig. 2(c)) and go to STEP 2.

STEP 6: If $i \neq K$, go to STEP 3. If $i = K$, compare the temporary solution with the obtained solution. If the obtained solution is better than the temporary solution, update the temporary solution and go to STEP 7.

STEP 7: If $k \neq L$, go to STEP 6. If $k = L$, go to STEP 2.

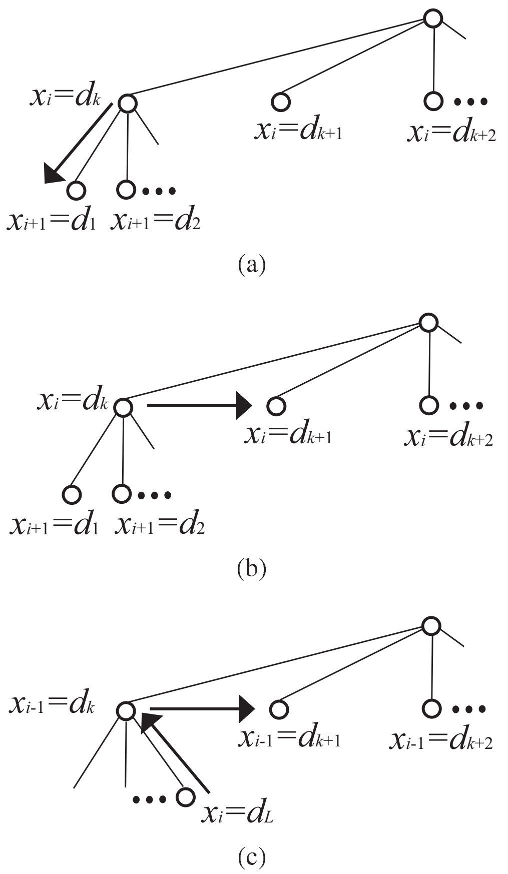

Figure 2. Movement of the search point.

4. Numerical Example

To show the effectiveness of the proposed method, we design the digital filters in which the desired frequency response:

$$H_d(\omega) = \begin{cases} e^{-j\tau\omega} & 0 \leq \omega \leq \omega_p \\ 0 & \omega_s \leq \omega \leq \pi \end{cases} . \tag{36}$$

Fig. 3 shows the magnitude response of the designed filters where the filter order is $m = n = 10$ and the wordlength of the coefficients is 10 bits. The solid line indicates the magnitude response of digital filter designed by the proposed method. For comparison, we show the result of simple roundoff case as the dotted line. Also, Fig. 4 shows the phase characteristic of the filter designed by the proposed method and simple roundoff. From Figs. 4, we can see the filter designed by the proposed method has the linear phase characteristic.

From Figs. 3 and 4, we can see the digital filter designed by the proposed method outperforms that designed by the simple roundoff.

5. Conclusion

In this paper, we have proposed the design method for IIR digital filter with discrete coefficients in the frequency domain. We have formed the filter design problem based on the WMLS criterion and search the good solution using the lower

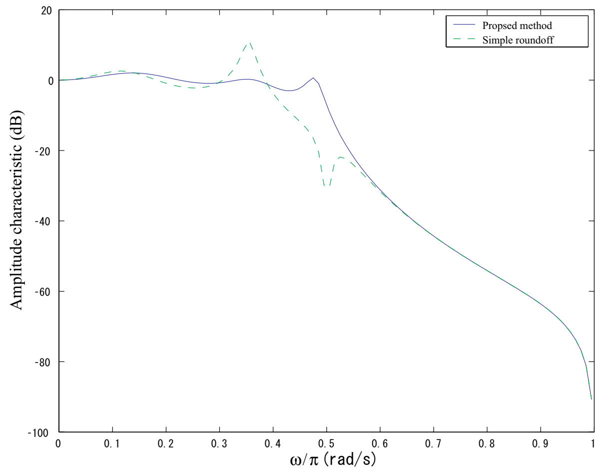

Figure 3. Magnitude response of the digital filters.

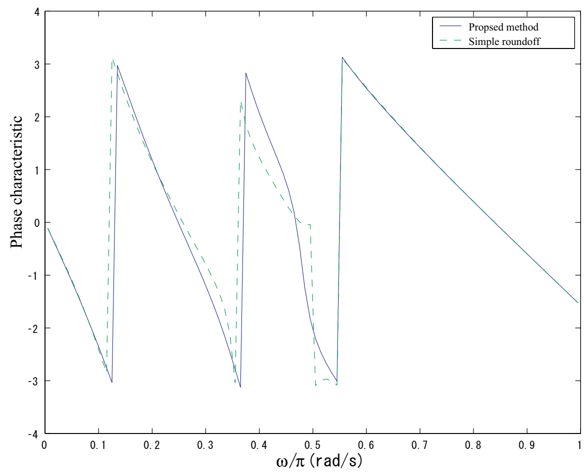

Figure 4. Phase characteristic of the digital filters.

bound estimation. Finally, we have confirmed the effectiveness of the proposed method based on a numerical example.

References

[1] J.Radecki, J.Konrad and E.Dubois,"Design of multidimensional finite-wordlength FIR and IIR filters by simulated annealing," IEEE trans. Circuits & Syst. II, vol.42, no.6, pp.424-431, June, 1995.

[2] R.Thamvichai, T.Bose and R.Haupt, "Design of 2-D multiplierless IIR filters using the genetic algorithm," IEEE Trans, Circuits & Syst. I, vol.49, no.6, pp.878-882, June 2002.

[3] M.Nakamoto and T.Hinamoto, "Design of IIR digital filters with discrete coefficients based on MLS criterion," IEICE Trans. vol.E89-A, no.4, pp.1116-1121, April 2006.

[4] C. T. Mullis and R. A. Roberts, "The use of second-order information in the approximation of discrete-time linear systems," IEEE Trans. Acoust., Speech, Signal Processing, vol. 24, no. 3, pp. 226-238, June 1976.