論文 / 著書情報
Article / Book Information

Title	Minimization of Output Errors of FIR Digital Filters by Multiple Decompositions of Signal Word
Authors	Mitsuhiko Yagyu, Akinori Nishihara, N. Fujii
Citation	IEICE Trans. Fundamentals., Vol. E81-A, No. 3, pp. 407-419
Pub. date	1998, 3
URL	http://search.ieice.org/
Copyright	(c) 1998 Institute of Electronics, Information and Communication Engineers

PAPER Special Section of Selected Papers from the 10th Karuizawa Workshop on Circuits and Systems

Minimization of Output Errors of FIR Digital Filters by Multiple Decompositions of Signal Word

Mitsuhiko YAGYU†, *Student Member*, Akinori NISHIHARA††, *and* Nobuo FUJII†, *Members*

SUMMARY FIR digital filters composed of parallel multiple subfilters are proposed. A binary expression of an input signal is decomposed into multiple shorter words, which drive the subfilters having different length. The output error is evaluated by mean squared and maximum spectra. A fast algorithm is also proposed to determine optimal filter lengths and coefficients of subfilters. Many examples confirm that the proposed filters generate smaller output errors than conventional filters under the condition of specified number of multiplications and additions in filter operations. Further, multiplier and adder structures (MAS) to perform the operations of the proposed filters are also presented. The number of gates used in the proposed MAS and its critical path are estimated. The effectiveness of the proposed MAS is confirmed.

key words: *FIR digital filters, multiple decompositions of signal word, error spectrum, the number of multiplications and additions*

1. Introduction

FIR digital filters are implemented by the finite number of multiplications and additions. Naturally, output signals of FIR filters have errors. To implement more precise arithmetic operations in filters, long wordlengths of signals and coefficients are needed. Long wordlength of the input signal does not guarantee high accuracy unless the output errors due to the finiteness of the number of taps are small enough. Thus high accuracy FIR digital filters need to have the large number of multipliers and adders with the long wordlength. This pushes up the hardware size, the operation delay and the number of executions of high-speed array multipliers [1].

Real-time implementation of direct-form N-tap FIR digital filters requires N multiplications and additions in a sampling interval. It is well known that an $l \times l$ multiplication can be performed using an $\frac{l}{2} \times l$ multiplier twice where l is the wordlength of a multiplier. In this case, the chip size is reduced at the price of longer execution time. When two $\frac{l}{2} \times l$ multipliers operate at a time to perform multiplications in filter operations, the speed of such two multipliers is faster than that of the $l \times l$ multiplier. So in this case, the execution time in filtering is reduced in comparison with that in filtering by using the $l \times l$ multiplier. In the above discussion, an $l \times l$ multiplier is decomposed into two $\frac{l}{2} \times l$ multipliers to reduce the hardware size or the execution time. Further reduction is expected in the case of multiple decompositions of an $l \times l$ multiplier. That is, several multipliers having shorter wordlengths may perform multiplications in filter operations with reduced hardware size or execution time. In a realization of filters, there are three factors for the multipliers to be specified; the number of multipliers, their wordlengths and the number of executions of the multiplications (NEM) in a sampling interval. Once the three factors are specified, it becomes important to minimize the errors of filters.

In this paper, FIR filter structures composed of parallel multiple subfilters driven by shorter signal words are proposed. We minimize the errors of the proposed structures realized with the specified multipliers. A binary expression of the input signal is decomposed into multiple shorter words, which drive the subfilters having different length. We call this type of filters as signal word decomposed filters (SWDFs). SWDFs do not have any deterministic frequency responses, because they have nonlinear operations. It is, however, possible to evaluate the output error of SWDFs by using mean squared and maximum spectra in passband, stopband and also transition band. Next, fast algorithms are also proposed to determine optimal filter length and coefficients of subfilters. In many design examples, we confirm that the SWDFs generate much smaller output errors in passband and stopband than conventional filters under the condition of specified NEM in filter operations.

Several techniques to design sharp cut-off filters have been proposed [2]–[4]. The filters designed by using these techniques have internal filters whose output signals are attenuated in transition band. Accordingly, even if SWDFs have comparatively large output errors in transition band, the output errors of SWDFs implemented as the internal filters do not become a problem. By analyzing the output errors in passband, stopband and also transition band. we confirm the superiority of the sharp cut-off filters using SWDFs as the internal filters.

Finally, multiplier and adder structures to perform the SWDF operations are presented. We analyze the number of gates used in the proposed structures and

Manuscript received July 3, 1997.

†The authors are with the Faculty of Engineering, Tokyo Institute of Technology, Tokyo, 152–8552 Japan.

††The author is with the Center for Research and Development of Educational Technology, Tokyo Institute of Technology, Tokyo, 152–8552 Japan.

their critical paths, and confirm the effectiveness of the proposed structures in comparison with conventional ones.

2. Signal Word Decomposed Filters

Figure 1 shows an M-channel SWDF structure. We assume that all filters discussed in this paper have zero phase. An ideal filter is defined as F_d, and its frequency response as $D(\omega)$, where amplitude of $D(\omega)$ is 1.0 in whole passband, and 0.0 in whole stopband. The M subfilters shown in Fig. 1 are called as F_i, $i = 1, \ldots, M$, their numbers of taps as T_i, $i = 1, \ldots, M$ and coefficients of the filter F_i as $h_i(j)$, $j = 1, \ldots, T_i$. All subfilters keep long wordlength and then only a total sum of outputs of the subfilters is rounded off so as to have the same wordlength with the input signal.

Now let the input signal expressed as l bit fixed-point binary numbers be $x(n)$, $|x(n)| \leq 1$, those quantization levels be q_i, $i = 1, \ldots, L$, $L = 2^l$, a set whose elements are q_i, $i = 1, \ldots, L$ be B_l and all filter coefficients be l_c bit binary numbers. Binary expressions of $x \in B_l$ are written as

$$x = b_1 b_2 b_3 \cdots b_{l(2)}, \tag{1}$$

where b_1 is a sign bit and $b_i \in \{0, 1\}$, $i = 1, \ldots, l$. Then x and b_i, $i = 1, \ldots, l$ satisfy

$$x = -b_1 + \sum_{i=2}^{l} 2^{-i+1} b_i. \tag{2}$$

Now input signal words are decomposed into M parts. Let the wordlength of the i-th significant signal word be l_i. Then we can write the i-th significant signal word of x shown in Eq. (1) as

$$b_{d_i+1} b_{d_i+2} b_{d_i+3} \cdots b_{d_i+l_i}, \tag{3}$$

where $d_i = l_1 + l_2 + \cdots + l_{i-1}$, $i = 2, \ldots, M$ and $d_1 = 0$. As shown in Fig. 1, input signal words are decomposed by the nonlinear function blocks $Q_i[\cdot]$, $i = 1, \ldots, M$ whose input signals are $x(n)$ having wordlength l. We

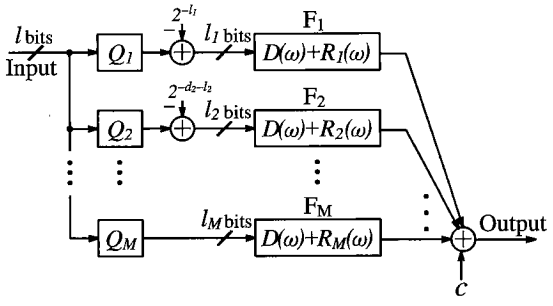


Fig. 1 An M-channel signal word decomposition filter structure.

define their outputs $Q_i[x]$, $i = 1, \ldots, M$, which correspond to an input x, as

$$Q_1[x] = -b_1 + \sum_{k=2}^{l_1} 2^{-k+1} b_k + 2^{-l_1}, \tag{4}$$

$$Q_j[x] = -2^{-d_j} \overline{b_{d_j+1}} + \sum_{k=d_j+2}^{d_j+l_j} 2^{-k+1} b_k + 2^{-d_j-l_j} \tag{5}$$

and

$$Q_M[x] = -2^{-d_M} \overline{b_{d_M+1}} + \sum_{k=d_M+2}^{d_M+l_M} 2^{-k+1} b_k, \tag{6}$$

where $j = 2, \ldots, M-1$. From Eqs. (4), (5) and (6), the signal words corresponding to the outputs $Q_i[x]$ are written as

$$Q_1[x] = b_1 b_2 \cdots b_{l_1} 1_{(2)}, \tag{7}$$

$$Q_j[x] = \overline{b_{d_j+1}} b_{d_j+2} \cdots b_{d_j+l_j} 1_{(2)} \tag{8}$$

and

$$Q_M[x] = \overline{b_{d_M+1}} b_{d_M+2} \cdots b_{d_M+l_M(2)}, \tag{9}$$

where $j = 2, \ldots, M-1$. Then $Q_i[x]$ and x also satisfy

$$x = \sum_{i=1}^{M} Q_i[x]. \tag{10}$$

All the binary expressions shown in Eqs. (7), (8) and (9) are 2's complement and thus their most significant bits are sign bits. Therefore all multiplications in the subfilters F_i, $i = 1, \ldots, M$ can be performed by using 2's complement parallel array multipliers [1]. Although the signal words given in Eqs. (7) and (8) have wordlength $l_i + 1$, respectively, the least significant bits of those signal words are equal to 1 independently of the input signal words x. So among those $l_i + 1$ bits, let the l_i bits without the least significant bit be an input to the filter F_i as shown in Fig. 1. Instead of neglecting the least significant bits,

$$c = \sum_{i=1}^{M-1} \sum_{j=1}^{T_i} 2^{-d_i-l_i} h_i(j) \tag{11}$$

is calculated beforehand. In filtering, c is added to outputs of an SWDF so that the errors caused by truncating the least significant bits are canceled.

The actual inputs of all subfilters F_i, $i = 1, \ldots, M$ have wordlength l_i so that wordlengths of all multiplications in subfilter F_i become $l_i \times l_c$ bits. Then to evaluate the amount of additions and multiplications of SWDFs, we define AAM as

$$AAM = \sum_{i=1}^{M} l_i l_c T_i. \tag{12}$$

Further, an accumulator is loaded with c before additions and multiplications not to increase AAM.

Now the $(l+1)$-th bit of the input signal $x(n)$ is regarded as 1, constantly. In that case, the functions $Q_i[x]$, $i = 1, \ldots, M$ satisfy

$$Q_i[-x] = -Q_i[x]. \tag{13}$$

3. Output Errors of SWDFs

3.1 Evaluation of Errors in Passband and Stopband

When all subfilters F_i, $i = 1, \ldots, M$ have a specified ideal frequency response, the SWDF does not generate any output errors. However, due to finiteness of AAM of SWDFs, each subfilter generally has a deviation of its frequency response from an ideal response, and the output signals of SWDFs have errors. In this section, the output errors are analyzed.

The discrete-time Fourier transforms (DTFTs) of $x(n)$ and $Q_i[x(n)]$, $i = 1, \ldots, M$ are called as $X(e^{j\omega})$ and $X_i(e^{j\omega})$, which are written as

$$X(e^{j\omega}) = \sum_n x(n)e^{-jn\omega} \tag{14}$$

and

$$X_i(e^{j\omega}) = \sum_n Q_i[x(n)]e^{-jn\omega}, \tag{15}$$

respectively. From Eq. (10), $x(n)$, $Q_i[x(n)]$ and their DTFTs satisfy

$$x(n) = \sum_{i=1}^{M} Q_i[x(n)] \tag{16}$$

and

$$X(e^{j\omega}) = \sum_{i=1}^{M} X_i(e^{j\omega}). \tag{17}$$

Let a zero-phase frequency response of the filter F_i be $H_i(\omega)$. Then the error responses $R_i(\omega)$ of $H_i(\omega)$, $i = 1, \ldots, M$ are written as

$$R_i(\omega) = H_i(\omega) - D(\omega). \tag{18}$$

From Fig. 1 and Eqs. (17) and (18), the output error $R_o(e^{j\omega})$ corresponding to the input signal $x(n)$ is obtained as

$$R_o(e^{j\omega}) = \sum_{i=1}^{M} H_i(\omega)X_i(e^{j\omega}) - D(\omega)X(e^{j\omega}) \tag{19}$$

$$= \sum_{i=1}^{M} R_i(\omega)X_i(e^{j\omega}). \tag{20}$$

By substituting Eq. (15) into Eq. (20), $R_o(e^{j\omega})$ is rewritten as

$$R_o(e^{j\omega}) = \sum_n \left\{ \sum_{i=1}^{M} R_i(\omega)Q_i[x(n)] \right\} e^{-jn\omega}. \tag{21}$$

We express the DTFT of an error response of an SWDF corresponding to an input level x as $R[x, \omega]$. Then $R[x, \omega]$ can be written as

$$R[x, \omega] = \sum_{i=1}^{M} Q_i[x]R_i(\omega). \tag{22}$$

By using Eqs. (21) and (22), the output error $R_o(e^{j\omega})$ corresponding to the input signal $x(n)$ can be rewritten as

$$R_o(e^{j\omega}) = \sum_n R[x(n), \omega]e^{-jn\omega}. \tag{23}$$

In [5], we have proposed two methods to analyze and evaluate output errors which are written as Eq. (23). Using the method in [5], we firstly analyze the mean squared output error spectrum (MSOES) of an SWDF. We assume that the stochastic input signals $x(n)$ are stationary independent process and that their probability density function is $p(q_i)$, $i = 1, \ldots, L$. Then the average and the variance of the error responses $R[q_i, \omega]$, $i = 1, \ldots, L$ are obtained as

$$R_{ave}(\omega) = \sum_{i=1}^{L} p(q_i)R[q_i, \omega] \tag{24}$$

and

$$R_{var}(\omega) = \sum_{i=1}^{L} p(q_i) \{R[q_i, \omega]\}^2 - \{R_{ave}(\omega)\}^2, \tag{25}$$

respectively. From [5], the MSOES of SWDFs is obtained as

$$R_{mse}(\omega) = R_{var}(\omega) + \frac{\{R_{ave}(\omega)\}^2}{P} \left| \sum_{n=0}^{P-1} e^{-jn\omega} \right|^2, \tag{26}$$

where P is the number of sampling points of the stochastic input signal $x(n)$. We assume that the probability density function of the input signals is an even function. From Eq. (13), $Q_j[x]$, $j = 1, \ldots, M$ are regarded as odd functions so that $Q_j[\cdot]$ and $p(\cdot)$ satisfy

$$\sum_{i=1}^{L} Q_j[q_i]p(q_i) = 0, \tag{27}$$

where $j = 1, \ldots, M$. Substituting Eqs. (22) and (27) into Eq. (24), we obtain

$$R_{ave}(\omega) = 0. \tag{28}$$

By using Eqs. (22), (25), (26) and (28), $R_{mse}(\omega)$ is rewritten as

$$R_{mse}(\omega) = \sum_{i=1}^{L} p(q_i) \left\{ \sum_{j=1}^{M} Q_j[q_i] R_j(\omega) \right\}^2. \quad (29)$$

Using the method in [5], we analyze the maximum output error spectrum (MOES) of an SWDF now. Note that the MOES $R_{\mathrm{worst}}(\omega)$ has the normalized maximum error value at frequency ω among error values of an SWDF at ω corresponding to every possible input signal [5]. Here, we define the maximum and the minimum among error responses corresponding to all input levels at a frequency ω as $R_{\max}(\omega)$ and $R_{\min}(\omega)$, respectively. Then $R_{\max}(\omega)$ and $R_{\min}(\omega)$ are written as

$$R_{\max}(\omega) = \max_{1 \leqq i \leqq L} R[q_i, \omega] \quad (30)$$

and

$$R_{\min}(\omega) = \min_{1 \leqq i \leqq L} R[q_i, \omega]. \quad (31)$$

From [5], by using Eqs. (30) and (31) $R_{\mathrm{worst}}(\omega)$ can be obtained as

$$R_{\mathrm{worst}}(\omega) = \begin{cases} \dfrac{R_{\max}(\omega) - R_{\min}(\omega)}{2}, & \omega \neq 0 \\ \max\{|R_{\max}(0)|, |R_{\min}(0)|\}, & \omega = 0 \end{cases}. \quad (32)$$

Now let the dynamic range of the internal signal $Q_i[x(n)]$ be r_i, $i = 1, \ldots, M$. Naturally, r_i satisfies

$$-r_i \leqq Q_i[q_j] \leqq r_i. \quad (33)$$

From Eqs. (4), (5) and (6), r_i, $i = 1, \ldots, M$ are given as

$$r_1 = 1 - 2^{-l_1}, \quad (34)$$

$$r_j = 2^{-d_j} - 2^{-d_j - l_j}, \quad (35)$$

where $j = 2, \ldots, M-1$, and

$$r_M = 2^{-d_M}. \quad (36)$$

From Eqs. (22), (30), (31) and (33), $R_{\max}(\omega)$ and $R_{\min}(\omega)$ become

$$R_{\max}(\omega) = \sum_{i=1}^{M} r_i |R_i(\omega)| \quad (37)$$

and

$$R_{\min}(\omega) = -\sum_{i=1}^{M} r_i |R_i(\omega)|, \quad (38)$$

respectively. By substituting Eqs. (37) and (38) into Eq. (32), $R_{\mathrm{worst}}(\omega)$ is rewritten as

$$R_{\mathrm{worst}}(\omega) = \sum_{i=1}^{M} r_i |R_i(\omega)|. \quad (39)$$

By using (29) and (39), the MSOES and the MOES of an SWDF can be evaluated in passband and stopband.

If all frequency responses of the subfilters F_i, $i = 1, \ldots, M$ are identical, the SWDF becomes a conventional linear filter which has that frequency response. Then $R_{\mathrm{worst}}(\omega)$ becomes an error response which the frequency response of the linear filter has. Further, the peak value of $R_{\mathrm{worst}}(\omega)$ becomes the Chebyshev error of the linear filter.

3.2 Evaluations in Transition Band

SWDFs do not have any deterministic frequency responses in passband, stopband and also transition band, due to their nonlinear operations $Q_i[\cdot]$, $i = 1, \ldots, M$. In design of typical frequency selective filters, an ideal response $D(\omega)$ is often unspecified in transition band. However in some applications of filters, estimations of frequency responses in transition band are needed. In this section, we analytically obtain frequency responses such that the MSOES or the MOES is minimum in transition band.

Firstly, the MSOES is minimized and then the frequency response $D_t(\omega)$ is obtained. Let an arbitrary frequency in transition band be ω_t. By using (18) and (29), the MSOES in transition band can be written as

$$R_{mse}(\omega_t) = \sum_{i=1}^{L} p(q_i) \left\{ \sum_{j=1}^{M} Q_j[q_i] \Big(H_j(\omega_t) - D_t(\omega_t) \Big) \right\}^2 \quad (40)$$

$$= \sum_{i=1}^{L} p(q_i) \left\{ {q_i}^2 D_t(\omega_t)^2 - 2q_i \sum_{j=1}^{M} Q_j[q_i] H_j(\omega_t) D_t(\omega_t) + \left(\sum_{j=1}^{M} Q_j[q_i] H_j(\omega_t) \right)^2 \right\}. \quad (41)$$

When Eq. (41) is minimum, $D_t(\omega_t)$ is obtained as

$$D_t(\omega_t) = \frac{\sum_{i=1}^{L} p(q_i) q_i \sum_{j=1}^{M} Q_j[q_i] H_j(\omega_t)}{\sum_{i=1}^{L} p(q_i) {q_i}^2}. \quad (42)$$

Next, the MOES is minimized and then the frequency response $D_t(\omega)$ is obtained. By using (18) and (39), the MOES in transition band can be written as

$$R_{\mathrm{worst}}(\omega_t) = \sum_{i=1}^{M} r_i |H_i(\omega_t) - D_t(\omega_t)|. \quad (43)$$

From Eqs. (34), (35) and (36), r_i, $i = 1, \ldots, M$ satisfy

$$r_1 > r_2 + r_3 + \cdots + r_M. \tag{44}$$

By using Eq. (44), when the MOES expressed as Eq. (43) is minimum, $D_t(\omega_t)$ is obtained as

$$D(\omega_t) = H_1(\omega_t) \tag{45}$$

as shown in Appendix A.

4. Design of SWDFs

4.1 An Algorithm to Design Optimum SWDFs

We use the peak value e_p of the MOES shown in Eq. (39) as the criterion for filter design. Then e_p is written as

$$e_p = \max_{0 \leqq \omega \leqq \pi} R_{\text{worst}}(\omega) = \max_{0 \leqq \omega \leqq \pi} \sum_{i=1}^{M} r_i |R_i(\omega)|. \tag{46}$$

$M, l_c, l_i, i = 1, \ldots, M$ and AAM are specified in design of SWDFs. Then in our algorithm, the SWDF having the minimum peak of the MOES is chosen among all SWDFs which meet the specifications. Let the numbers of taps of subfilters F_i of a designed SWDF be T_i. Then from Eq. (12), T_i, $i = 1, \ldots, M$ must satisfy

$$\sum_{i=1}^{M} l_i l_c T_i \leqq AAM. \tag{47}$$

In this paper, l_i, $i = 1, \ldots, M$ are assumed to be equal to l/M. From this assumption, Eq. (47) can be rewritten as

$$\sum_{i=1}^{M} T_i \leqq T', \tag{48}$$

where

$$T' \triangleq \frac{M \times AAM}{l\ l_c}. \tag{49}$$

T' is determined by using the specified M, l, l_i, $i = 1, \ldots, M$ and AAM.

Coefficients of a zero phase FIR filter have symmetry. So we define a vector whose elements are the independent coefficients of the filter F_i, $i = 1, \ldots, M$ as $\boldsymbol{a}_i$. Then the frequency response $H_i(\omega)$ of the filter F_i is written as

$$H_i(\omega) = \boldsymbol{A}(\omega)\boldsymbol{a}_i, \tag{50}$$

where $\boldsymbol{A}(\omega)$ is a vector whose elements are trigonometric functions [6]. From Eq. (18), the error response $R_i(\omega)$ is written as

$$R_i(\omega) = \boldsymbol{A}(\omega)\boldsymbol{a}_i - D(\omega). \tag{51}$$

To make a discussion simple, we deal with the case of $M = 2$. Then e_p shown in Eq. (46) can be written as

$$\begin{aligned} e_p = \max_{0 \leqq \omega \leqq \pi} \big[& r_1 \{\boldsymbol{A}(\omega)\boldsymbol{a}_1 - D(\omega)\} + r_2 \{\boldsymbol{A}(\omega)\boldsymbol{a}_2 - D(\omega)\}, \\ & r_1 \{\boldsymbol{A}(\omega)\boldsymbol{a}_1 - D(\omega)\} - r_2 \{\boldsymbol{A}(\omega)\boldsymbol{a}_2 - D(\omega)\}, \\ & -r_1 \{\boldsymbol{A}(\omega)\boldsymbol{a}_1 - D(\omega)\} + r_2 \{\boldsymbol{A}(\omega)\boldsymbol{a}_2 - D(\omega)\}, \\ & -r_1 \{\boldsymbol{A}(\omega)\boldsymbol{a}_1 - D(\omega)\} - r_2 \{\boldsymbol{A}(\omega)\boldsymbol{a}_2 - D(\omega)\}\big]. \end{aligned} \tag{52}$$

The minimization of Eq. (52) is a linear minimax problem and thus can be solved by using linear programming (LP). In the same way, if M is larger than 2, the minimization problem can be solved.

Further, to obtain the optimum SWDF, the numbers of taps of subfilters T_i, $i = 1, \ldots, M$ have to be determined to obtain the minimum e_p. So firstly, all combinations of T_i, $i = 1, \ldots, M$ which satisfy Eq. (48) are determined. Next e_p is minimized by using T_i, $i = 1, \ldots, M$ in each combination. Then a combination of T_i, $i = 1, \ldots, M$ corresponding to the minimum e_p is the solution. By using above algorithm, the optimum SWDF meeting the specification can be obtained. Here after, we call this algorithm as algorithm 1.

4.2 Fast Algorithms to Design SWDFs

In the algorithm 1 proposed in the above section, a linear programming is iteratively used to determine the numbers of taps of all subfilters. However in the case of large AAM and M, the algorithm 1 consumes enormous computing cost. In this section, we firstly propose an algorithm to optimize coefficients of subfilters whose numbers of taps are estimated beforehand, and then propose an algorithm to estimate the numbers of taps of subfilters which the minimal error SWDF has. We call this algorithm as algorithm 2.

To obtain the algorithm 2, we approximate e_p shown in Eq. (46) as

$$e_p = \max_{0 \leqq \omega \leqq \pi} \sum_{i=1}^{M} r_i |R_i(\omega)| \approx \sum_{i=1}^{M} \max_{0 \leqq \omega \leqq \pi} r_i |R_i(\omega)|. \tag{53}$$

So we define a criterion for fast filter design as

$$\widehat{e_p} \triangleq \sum_{i=1}^{M} \max_{0 \leqq \omega \leqq \pi} r_i |R_i(\omega)|. \tag{54}$$

In the algorithm 2, $\widehat{e_p}$, that is, the sum of the the Chebyshev errors of subfilters is minimized. Then each subfilter is independently designed by using Remez exchange algorithm [7].

Next, a subroutine in the algorithm 2 to estimate the numbers of taps of subfilters is presented. Firstly, the Chebyshev errors e_i of the subfilters F_i, $i = 1, \ldots, M$ can be approximately written as

$$20 \log_{10} e_i \approx \alpha T_i + \beta, \tag{55}$$

where α and β are negative constants and T_i is the number of taps of the designed subfilter F_i [8]. By using Eqs. (54) and (55), $\widehat{e_p}$ is written as

$$\widehat{e_p} = \sum_{i=1}^{M} r_i 10^{\frac{\alpha T_i + \beta}{20}}. \tag{56}$$

Equation (48) can be rewritten as

$$\prod_{i=1}^{M} r_i 10^{\frac{\alpha T_i + \beta}{20}} \geq 10^{\frac{\alpha T' + M\beta}{20}} \prod_{i=1}^{M} r_i. \tag{57}$$

Equation (57) is equivalent to Eq. (48), because an exponential function monotonically increases. Accordingly, Eq. (57) is a condition which T_i, $i = 1, \ldots, M$ must satisfy. Now we assume that T_i, $i = 1, \ldots, M$ are odd integers or zero. Then the problem to estimate the numbers of taps of subfilters which the minimal error SWDF has can be formulated as

$$\begin{aligned} &\textit{Minimize} \quad \sum_{i=1}^{M} r_i 10^{\frac{\alpha T_i + \beta}{20}} \\ &\textit{Subject to} \quad \prod_{i=1}^{M} r_i 10^{\frac{\alpha T_i + \beta}{20}} \geq 10^{\frac{\alpha T' + M\beta}{20}} \prod_{i=1}^{M} r_i, \\ &\quad T_i \text{ are odd integers or zero.} \end{aligned} \tag{58}$$

Since the variables T_i, $i = 1, \ldots, M$ are integers, Eq. (58) is an integer programming problem. An arbitrary continuous minimization problem derived from Eq. (58) can be solved by using the subroutine shown in Appendix B. So the integer programming problem in Eq. (58) can be solved by using the branch and bound algorithm [9].

In the algorithm 2, each subfilter having the estimated number of taps is optimized by using Remez exchange algorithm. Now we propose another fast algorithm. Two subfilters having the estimated numbers of taps are simultaneously optimized by using LP, and then by iterating such an optimization all subfilters are designed. We call this algorithm as algorithm 3 and show it as follows.

1. Let an iteration counter $k := 0$ and a provisional MOES $R_p(\omega) := 0$.
2. Solve the following minimax problem by using linear programming.

$$\begin{aligned} \textit{Minimize} \quad & \max_{0 \leq \omega \leq \pi} \{ R_p(\omega) \\ & + r_{M-k} |\boldsymbol{A}(\omega)\boldsymbol{a}_{M-k} - D(\omega)| \\ & + r_{M-k-1} |\boldsymbol{A}(\omega)\boldsymbol{a}_{M-k-1} - D(\omega)| \}. \end{aligned} \tag{59}$$

   Then the coefficient vectors $\boldsymbol{a}_{M-k}$ and $\boldsymbol{a}_{M-k-1}$ are obtained.

3. By using the obtained $\boldsymbol{a}_{M-k}$ and $\boldsymbol{a}_{M-k-1}$, $R_p(\omega)$ is updated by

$$\begin{aligned} R_p(\omega) := R_p(\omega) &+ r_{M-k} |\boldsymbol{A}(\omega)\boldsymbol{a}_{M-k} - D(\omega)| \\ &+ r_{M-k-1} |\boldsymbol{A}(\omega)\boldsymbol{a}_{M-k-1} - D(\omega)|. \end{aligned} \tag{60}$$

4. If the designs of all subfilters are finished, then stop. Otherwise let $k := k + 2$ and go to the step 2.

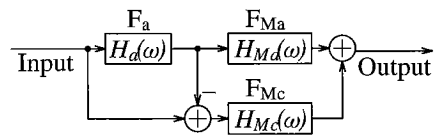
The algorithm 3 can design better SWDFs than the algorithm 2 at the price of longer computing time.

5. Application to Frequency Response Masking Technique

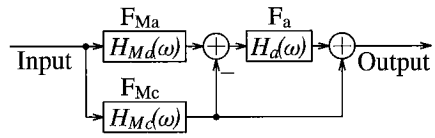
Several techniques to design sharp cut-off filters have been proposed [2]–[4]. The filters designed by using these techniques have internal filters whose output signals are attenuated in transition band. Accordingly, even if SWDFs have large output errors in transition band, the output errors of SWDFs implemented as the internal filters do not become a problem.

In this paper, we treat the frequency response masking technique (FRMT) [2] among those techniques to design sharp cut-off filters. SWDFs can be applied to other techniques [3], [4] also. The FRMT is an effective method to reduce the complexity of FIR filters with a very narrow transition band and a wide passband. Figures 2 (a) and (b) show a filter structure implemented by using FRMT, which is composed of two masking filters F_{Ma} and F_{Mc} and a prototype filter F_a, and its transposed structure, respectively. The frequency responses of three filters F_{Ma}, F_{Mc} and F_a are defined as $H_{Ma}(\omega)$, $H_{Mc}(\omega)$ and $H_a(\omega)$, respectively. Since the three internal filters are linear and time-invariant, two structures shown in Figs. 2 (a) and (b) can have same output signal corresponding to an input signal.

In FRMT, F_a have several passbands and stopbands, because F_a is an interpolated filter. Then F_{Ma} and F_{Mc} can have don't care bands where F_a has stopbands. If typical frequency selective filters are designed by using FRMT, ideal frequency responses of F_{Ma} and



(a) Original structure.



(b) Transposed structure.

Fig. 2 Filter structure implemented by using FRMT.

F_{Mc} can be unspecified in transition bands of F_{Ma} and F_{Mc} [2]. We propose to implement F_{Ma} and F_{Mc} of the structure shown in Fig. 2 (b) as SWDFs. Even if F_{Ma} and F_{Mc} implemented as SWDFs generate large output errors in their transition bands, the output errors can be attenuated by F_a.

Let the subfilters of F_{Ma} and F_{Mc} implemented as SWDFs be $F_{Ma,i}$ and $F_{Mc,i}$, $i = 1, \ldots, M$ and their frequency responses be $H_{Ma,i}(\omega)$ and $H_{Mc,i}(\omega)$, $i = 1, \ldots, M$, respectively. We redefine an ideal response of a filter implemented by using FRMT as $D(\omega)$. From Eqs. (29) and (39), the MSOES and the MOES of a filter implemented by using FRMT are obtained as

$$R_{mse}(\omega) = \sum_{i=0}^{L-1} p(q_i) \left\{ \sum_{j=1}^{M} Q_j(q_i) \left\{ H_a(\omega)\{H_{Ma,j}(\omega) - H_{Mc,j}(\omega)\} + H_{Mc,j}(\omega) - D(\omega) \right\} \right\}^2 \tag{61}$$

and

$$R_{\text{worst}}(\omega) = \sum_{i=1}^{M} r_i \left| H_a(\omega)\{H_{Ma,i}(\omega) - H_{Mc,i}(\omega)\} + H_{Mc,i}(\omega) - D(\omega) \right|, \tag{62}$$

respectively. In the same way to obtain Eqs. (42) and (45), the MSOES and the MOES can be evaluated in transition band.

[2] shows an algorithm to design filters by using FRMT. In the algorithm, several design parameters of F_{Ma} and F_{Mc} are determined by using the experience of the author of [2]. Then F_{Ma} and F_{Mc} are firstly designed. Next, the Chebyshev error of the filter implemented by using FRMT is minimized and then the optimum F_a is obtained. Next if F_{Ma} and F_{Mc} are implemented as SWDFs, F_{Ma} and F_{Mc} are firstly designed by using the algorithm 1, 2 or 3. In order to reduce the peak of the MOES of the filter implemented by FRMT, the MOES of F_{Ma} where F_{Mc} has transition band should be 50 percent smaller than the MOES of F_{Ma} in other bands from our experience. The weighted MOES is optimized by using a weighting function. In the same way, the MOES of F_{Mc} is optimized. Next, the peak of $R_{\text{worst}}(\omega)$ shown in Eq. (62) is minimized and then the optimum F_a is obtained.

6. Design Examples

In this section, SWDFs and conventional linear filters are compared in many design examples. We specify the wordlength of input signals as $l = 16$. The conventional filters are regarded as SWDFs with $M = 1$. Lowpass filters with passband $0.0-0.15$, stopband $0.22-0.5$ and $AAM/(l \times l_c)$ shown in Eq. (49) $1-79$ are designed and their transition band $0.15-0.22$ is specified as don't care band. If M multipliers with wordlengths $l/M \times l_c$ are used to implement designed filters, the multiplications of the SWDFs with M channels can be performed by using the M multipliers $AAM/(l \times l_c)$ times in a sampling interval. So the SWDFs with $M = 1$, namely, conventional filters have the same number of taps with $AAM/(l \times l_c)$. The SWDFs with $M > 1$ are designed by using the algorithm 2. By using the algorithm 1, the optimum SWDFs with $M = 2$ are also designed.

Figure 3 shows the peaks of the MOES $R_{\text{worst}}(\omega)$ of designed filters. From Fig. 3, with the growth of M, the errors of the SWDFs with M channels are decreased. Especially the errors of the SWDFs with $M = 16$ are 28 [dB] less than the errors of the SWDFs with $M = 1$, which are the conventional filters. Furthermore, we obtain the result that the errors of the SWDFs designed by using the algorithm 2 are nearly equal to those of the optimum SWDFs.

One of our goals is to design filters whose output

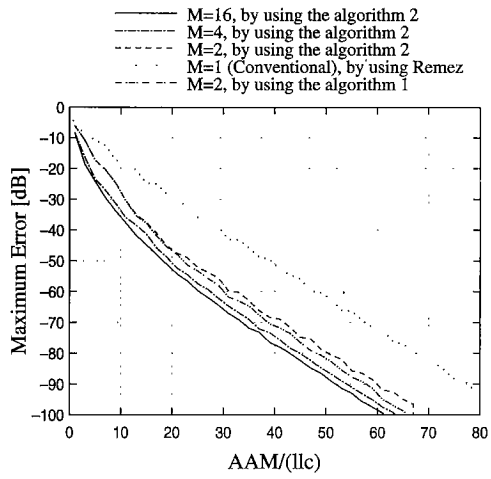


Fig. 3 The peaks of the MOES shown in Eq. (46).

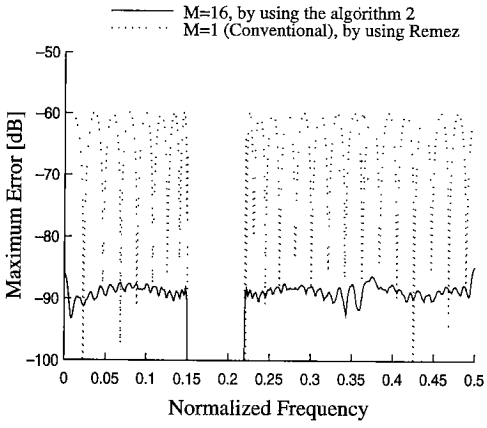
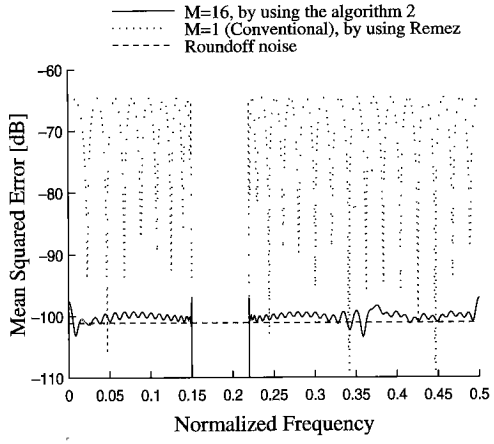
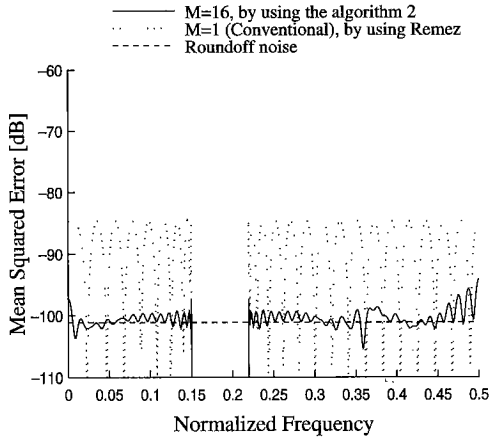


Fig. 4 The MOES of the designed filters with $AAM/(l \times l_c) = 47$.

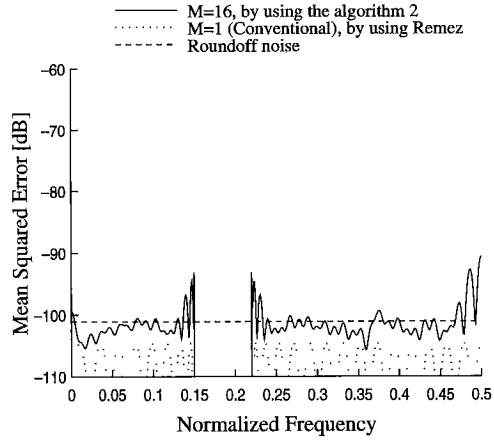
signals have small errors enough to utilize the resolution of wordlength $l = 16$. So we compare the designed SWDFs which have $AAM/(l \times l_c) = 47$ in detail. Figure 4 shows the MOES of those SWDFs with $M = 16$ and $M = 1$. The MOES of the SWDF with $M = 16$ is much smaller than that with $M = 1$ in whole passband and stopband. Next, We give three types of stochastic input signals having uniform distributions in the ranges $|x(n)| \leqq 1.0$, $|x(n)| \leqq 0.1$ and $|x(n)| \leqq 0.01$. Figures 5 (a), (b) and (c) show the MSOES corresponding to the stochastic input signals of the SWDFs with $M = 16$ and $M = 1$. The output signals of SWDFs are rounded off so as to have the wordlength $l = 16$. Accordingly, the roundoff noise are added to the output of SWDFs. The mean squared spectrum of the roundoff noise has the amplitude -101.1 [dB] and is shown in Figs. 5 (a), (b) and (c) also. From Figs. 5 (a), (b) and (c), in the case of the input signals having the large amplitude $|x(n)| \leqq 1.0$, the MSOES of the SWDF with $M = 16$ are 35 [dB] less than that with $M = 1$. In that case, the SWDF with $M = 16$ has great effectiveness. In the case of input signals having small amplitude, the SWDF with $M = 16$ does not have such great effectiveness. However, the MSOES of the SWDF with $M = 16$ is 15 [dB] less than that with $M = 1$. If the amplitude of input signals is smaller than 0.01, the MSOES of the SWDF with $M = 1$ is smaller than that with $M = 16$. The roundoff noise having the amplitude -101.1 [dB] is constantly added to the output signals of SWDFs. So we can regard that the output errors of the SWDF



(a) $|x(n)| \leqq 1.0$.

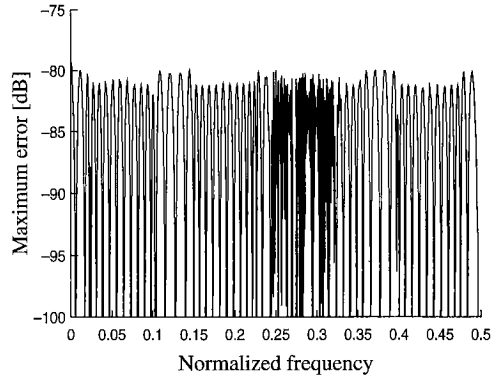


(b) $|x(n)| \leqq 0.1$.

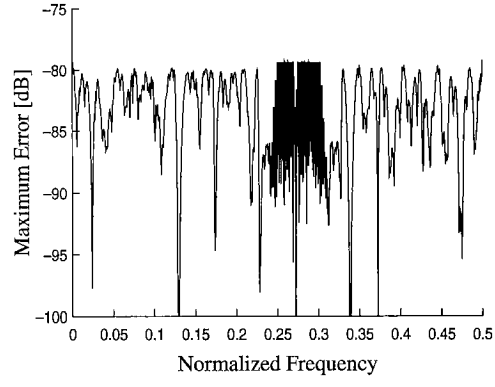


(c) $|x(n)| \leqq 0.01$.

Fig. 5 The MSOES of the designed filters with $AAM/(l \times l_c) = 47$.

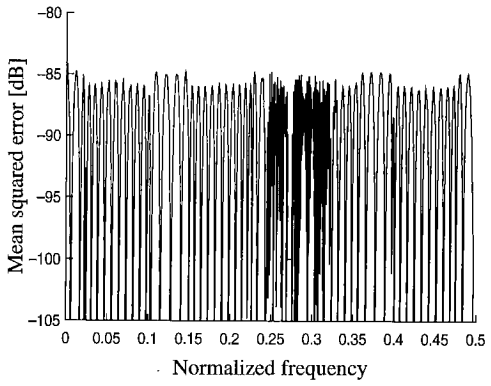


(a) Filter 1.

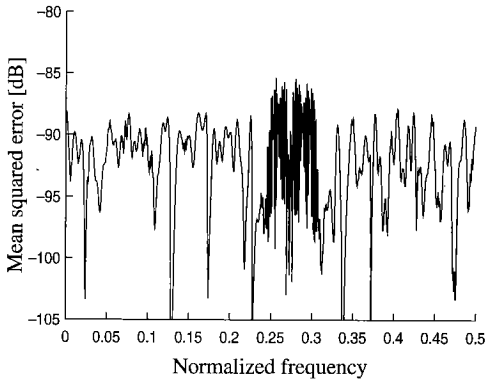


(b) Filter 3.

Fig. 6 The MOES of the filters 1 and 3.

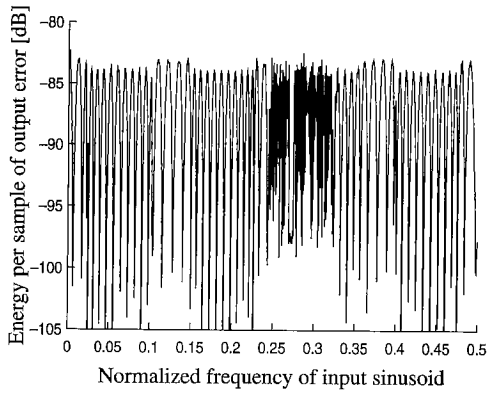


(a) Filter 1.

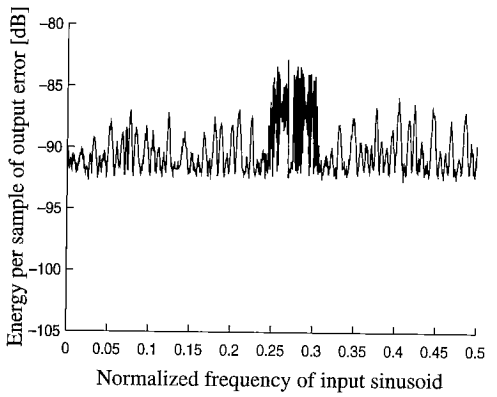


(b) Filter 3.

Fig. 7 The MSOES corresponding to the stochastic input signal $|x(n)| < 1.0$ of the filters 1 and 3.



(a) Filter 1.



(b) Filter 3.

Fig. 8 The energy per sample of the output errors corresponding to sinusoids of the filters 1 and 3.

with $M = 16$ are nearly equal to those of the SWDF with $M = 1$, even if the input signals have such small amplitude.

Next by using FRMT, the sharp cut-off filters with SWDFs are designed and then their output errors are evaluated. The specifications are; passband $0.0 - 0.27$, stopband $0.275 - 0.5$ and the peak of the MOES 78 [dB] less than the dynamic range of input signals. The masking filters are implemented as SWDFs with $M = 1$, 2 and 4 which are designed by using the algorithm 3. Then by using the algorithm in [2], their prototype filters are obtained. We call the filters designed by using FRMT which have the masking filters implemented as SWDFs with $M = 1$, 2 and 4 as the filters 1, 2 and 3, respectively. The peaks of the MOES given in Eq. (62) of the filters 1, 2 and 3 are -79.3 [dB], -78.7 [dB] and -79.2 [dB] and thus roughly equal. Figures 6 (a) and (b) show the MOES in whole band of the filters 1 and 3. We give a stochastic input signals having uniform distributions in the range $|x(n)| \leq 1.0$. Then the MSOES corresponding to the stochastic input signals of the filters 1 and 3 are shown in Figs. 7 (a) and (b). From these figures, the MSOES and the MOES of the filters 1 and 3 can be regarded roughly equal. Next, the actual output errors of the filters 1 and 3 are compared, when sinusoids $\cos(\omega_{in})n$ are applied as the input signals. We give $\omega_{in} = 0.000\pi$, 0.001π, 0.002π, $0.003\pi, \ldots, 0.999\pi$ and 1.000π and their number of sampling points as 10000. Then we calculate the energy per sample of the output error signal corresponding to each input signal in time domain. Figures 8 (a) and (b) show the energies per sample of the output error signals. From Figs. 8 (a) and (b), the output errors corresponding to the actual input signals of the filters 1 and 3 are roughly equal also. The numbers of taps of the two masking filters and the prototype filter of the filter 1 are 51, 103 and 119, respectively. However, those of the filters 1 and 2 are 95, 75, 52 and 95, 67, 53, respectively. The total number of taps of the filter 3 is 22% less than that of the filter 1 and thus the number of taps can be reduced by multiple decompositions of input signal word.

7. Multiplier and Adder Structures to Implement Operations of SWDF

We assume that l bit input signal words are decomposed into M signal words having wordlength l/M and that coefficients of SWDFs also have the wordlength

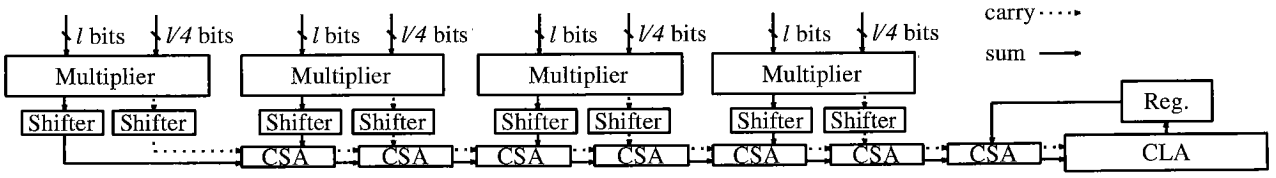


Fig. 9 A multiplier and adder structure having four $l/4 \times l$ multipliers.

l. The signals corresponding to those decomposed signal words excite their respective subfilters of an SWDF. To perform multiplications of the SWDF, several multipliers with wordlength $l/M \times l$ which operate at a time are used. In this section, M multipliers with that wordlength are assumed to be used. We call this structure as the multiplier and adder structure (MAS). The number of gates used in the MAS and its critical path are analyzed and then compared with that used in an $l \times l$ multiplier and its critical path, respectively.

Figure 9 shows an MAS having four $l/4 \times l$ multipliers to perform multiplications and additions of SWDF with $M = 4$, as an example. Those multipliers are implemented as the carry save scheme array multipliers [1] and thus have two outputs, summations and carries. All the outputs of the multipliers are shifted and summed up by using the carry save adders (CSA). The MAS shown in Fig. 9 is implemented as a chain structure. However it can be also implemented as a tree structure to obtain faster operation speed. When the MAS which has l multipliers with the wordlength $1 \times l$ is implemented as a tree structure, the MAS is called Wallace tree structure [10]. In Fig. 9, the final addition is performed by using a carry look-ahead adder (CLA).

Now let the wordlength l be 16. Then we give M as 1, 2, 4, 8 and 16. The number of gates used in the MAS and its critical path are estimated. In the case of $M = 1$, the MAS has an $l \times l$ multiplier, a CSA and a CLA and do not have any shifters. In the realization of the CLA, we use the standard scheme given in [11], which uses full internal carry look-ahead within 4-bit slices. Figure 10 shows the numbers of gates used in the MASs and the number of gates on their critical paths. From Fig. 10, with the growth of M, the total number of gates used in the MAS is increased. However, the number of gates used in the MAS having two $l/2 \times l$ multipliers is nearly equal to that used in the MAS having an $l \times l$ multiplier. The number of gates on the critical path is minimum at $M = 2$. So we can obtain the fact that the MAS which has two $l/2 \times l$ multipliers has the best performance at the lowest price. Further from Fig. 3, the effectiveness of the SWDFs having 2 channels is as large as that of the SWDFs having more than 2 channels. Therefore in the case of $l = 16$, it is most efficient that the operations of the SWDF having 2 channels are performed by using two $l/2 \times l$ multipliers. The more $l/2 \times l$ multipliers are used, the higher throughput may be obtained.

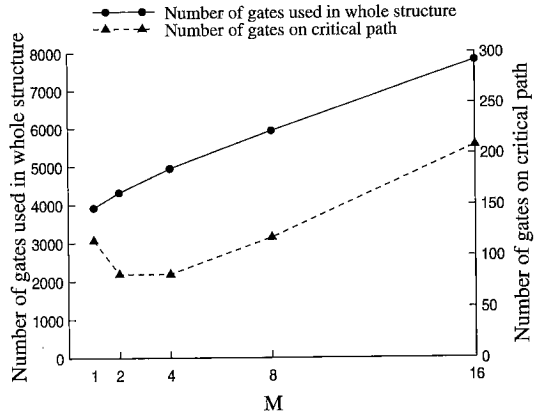


Fig. 10 The number of gates used in the MAS and the number of gates on its critical path.

8. Conclusions

This paper has presented FIR digital filters composed of parallel multiple subfilters. A binary expression of the input signal is decomposed into multiple shorter words, which drive the subfilters having different length. The proposed filters have nonlinear operations, however the output errors can be evaluated by mean squared and maximum spectra. Fast algorithms have been also proposed to determine optimal filter lengths and coefficients of subfilters. Many examples have confirmed that the maximum output error spectra of the proposed filters have been 28 [dB] less than those of the conventional filters under the condition of specified number of multiplications and additions in filter operations. Further, the actual output errors of the proposed and the conventional filters corresponding to many sinusoids are analyzed. Then we have also confirmed the superiority of the proposed filters to the conventional filters. We also have presented multiplier and adder structures (MAS) to perform the operations of the proposed filters. The number of gates used in the proposed MAS has been nearly equal to that of gates used in the conventional MAS, however the throughput of the proposed MAS has been larger than that of the conventional MAS. In future, we will analyze the chip area and the power consumption of the proposed MAS in detail.

References

[1] C.R. Baugh and B.A. Wooley, "A two's complement paral-

lel array multiplication algorithm," IEEE Trans. Comput., vol.C-22, pp. 1045–1047, Dec. 1973.
[2] Y.C. Lim, "Frequency-response masking approach for the synthesis of sharp linear phase digital filters," IEEE Trans. Circuits & Syst., vol.CAS-33, pp.357–364, April 1986.
[3] Y. Neuvo, C.Y. Dong, and S.K. Mitra, "Interpolated finite impulse response filters," IEEE Trans. Acoust. Speech, Signal Processing, vol.ASSP-32, pp.563–570, June 1984.
[4] J.W. Adams and A.N. Willson, Jr., "Some efficient digital prefilter structures," IEEE Trans. Circuits & Syst., vol.CAS-31, pp.260–266, March 1984.
[5] M. Yagyu, A. Nishihara, and N. Fujii, "Analysis and minimization of output errors of 2-D non-separable FIR digital filters with finite precision internal signals," IEICE Trans. Fundamentals, vol.E80-A, pp. 1391–1402, Aug. 1997.
[6] L.R. Rabiner and B. Gold, "Theory and application of digital signal processing," Prentice-Hall, 1975.
[7] T.W. Parks and J.H. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," IEEE Trans. Circuit Theory, vol.CT-19, pp.189–194, March 1972.
[8] J.F. Kaiser, "Nonrecursive digital filter design using the I_0-sinh window function," Proc. IEEE Int. Symp. Circuits & Syst., pp.20–23, 1974.
[9] R. Fletcher, "Practical methods of optimization," Second Edition, John Wiley & Sons, 1987.
[10] C.S. Wallace, "A Suggestion for a fast multipliers," IEEE Trans. Electronic Computers, pp.14–17, Feb. 1964.
[11] I. Flores, "The logic of computer arithmetic," Prentice-Hall, 1963.

Appendix A: Minimization of the MOES in Transition Band

The problem to minimize the MOES Eq. (43) can be formulated as

$$Minimize \quad f(x) = \sum_{i=1}^{M} r_i |x - a_i|, \tag{A·1}$$

where a variable $x \triangleq D_t(\omega_t)$ and given constant numbers $a_i \triangleq H_i(\omega_t)$. The derivative of Eq. (A·1) can be defined at $x \neq a_i$, $i = 1, \ldots, M$, and written as

$$f'(x) = \sum_{i=1}^{M} s_i(x) r_i, \tag{A·2}$$

where $s_i(x)$ satisfies

$$s_i(x) = \begin{cases} 1 & x > a_i \\ -1 & x < a_i \end{cases}. \tag{A·3}$$

Now consider $f'(x)$ where two ranges of x, $x > a_1$ and $x < a_1$, and thus by using Eq. (44), $f'(x)$ satisfies

$$f'(x) = \begin{cases} r_1 + \sum_{i=2}^{M} s_i(x) r_i > 0 & x > a_1 \\ -r_1 + \sum_{i=2}^{M} s_i(x) r_i < 0 & x < a_1 \end{cases}. \tag{A·4}$$

Equation (A·4) implies that $f(x)$ where $x > a_1$ is monotonically increasing and $f(x)$ where $x < a_1$ is monotonically decreasing. Accordingly, $f(x)$ has a minimum value at $x = a_1$.

Appendix B: A Subroutine to Solve an Arbitrary Continuous Minimization Problem Derived from Eq. (58)

In the branch and bound algorithm [9] an arbitrary continuous minimization problem derived from Eq. (58) is given as

$$\begin{aligned} &Minimize \quad \sum_{i=1}^{N} V_i \\ &Subject\ to \quad \prod_{i=1}^{N} V_i \geqq K \text{ and } 0 < a_j \leqq V_j \leqq b_j, \\ &\text{where } a_j < b_j, \quad j = 1, \ldots, N. \end{aligned} \tag{A·5}$$

The upper and lower bounds a_i, b_i, $i = 1, \ldots, N$ are given as positive numbers. Those numbers always satisfy only one condition among four conditions,

$$\text{(I)} \quad \prod_{i=1}^{N} b_i < K, \tag{A·6}$$

$$\text{(II)} \quad \prod_{i=1}^{N} b_i = K, \tag{A·7}$$

$$\text{(III)} \quad \prod_{i=1}^{N} b_i > K \quad \text{and} \quad \prod_{i=1}^{N} a_i \geqq K \tag{A·8}$$

and

$$\text{(IV)} \quad \prod_{i=1}^{N} b_i > K \quad \text{and} \quad \prod_{i=1}^{N} a_i < K. \tag{A·9}$$

Here, an arbitrary solution of the minimization problem Eq. (A·5) is defined as $\widetilde{V}_i$, $i = 1, \ldots, N$. If $a_i, b_i, i = 1, \ldots, N$ satisfy a condition (I), (II) or (III), $\widetilde{V}_i$, $i = 1, \ldots, N$ are easily obtained as

$$\widetilde{V}_i = \begin{cases} \emptyset & \text{(I)} \\ b_i & \text{(II)} \\ a_i & \text{(III)} \end{cases}, \tag{A·10}$$

respectively. If a_i, b_i, $i = 1, \ldots, N$ satisfy (IV), the solution $\widetilde{V}_i$, $i = 1, \ldots, N$ always satisfy

$$\prod_{i=1}^{N} \widetilde{V}_i = K. \tag{A·11}$$

We give the proof of Eq. (A·11) as follows.

Proof: From the condition (IV) and Eq. (A·5), one of $\widetilde{V}_i$, $i = 1, \ldots, N$ is larger than its lower bound. So we write $\widetilde{V}_1$ as $\widetilde{V}_1 = a_1 + \varepsilon$, $\varepsilon > 0$. To prove Eq. (A·11), we assume

$$(a_1+\varepsilon)\prod_{i=2}^{N}\widetilde{V}_i = K+\delta > K, \tag{A·12}$$

where $\delta > 0$. Now we define ε' as

$$\varepsilon' \triangleq \frac{K\varepsilon}{K+\delta} < \varepsilon. \tag{A·13}$$

We define $\widetilde{V}_1'$ as

$$\widetilde{V}_1' \triangleq a_1 + \frac{K\varepsilon}{K+\delta} < \widetilde{V}_1. \tag{A·14}$$

Replace $\widetilde{V}_1$ by $\widetilde{V}_1'$ as a new solution. Then $\widetilde{V}_1'$, $\widetilde{V}_i$, $i = 2,\ldots,N$ satisfy the condition given in Eq. (A·5) and

$$\widetilde{V}_1' + \sum_{i=2}^{N}\widetilde{V}_i < \sum_{i=1}^{N}\widetilde{V}_i. \tag{A·15}$$

Equation (A·15) is inconsistent with the definition of $\widetilde{V}_i$, $i = 1,\ldots,N$. □

Hence if a_i, b_i, $i = 1,\ldots,N$ satisfy (IV), the optimum solution can be obtained by solving a minimization problem given by

$$\begin{aligned} &\textit{Minimize} && \sum_{i=1}^{N} V_i \\ &\textit{Subject to} && \prod_{i=1}^{N} V_i = K \text{ and } 0 < a_j \leqq V_j \leqq b_j, \\ & && \text{where } a_i < b_i, \quad i = 1,\ldots,N. \end{aligned} \tag{A·16}$$

Let a candidate of the solutions of the minimization problem Eq. (A·16) be V_i^*, $i = 1,\ldots,N$. Then we give the following lemma.

Lemma 1: If V_p^* and V_q^*, where $1 \leqq \exists p,\ \exists q \leqq N$, satisfy $V_p^* < V_q^*$, $V_p^* < b_p$ and $V_q^* > a_q$, V_i^*, $i = 1,\ldots,N$ is not any solutions of Eq. (A·16).

Proof: We assume V_i^*, $i = 1,\ldots,N$ is a solution of Eq. (A·16) and then define ε_p and ε_q as

$$0 < \exists\varepsilon_p < \min\left\{ b_p - V_p^*,\ \frac{V_q^* - V_p^*}{2},\ \frac{(V_q^* - a_q)V_p^*}{a_q},\ \frac{(V_q^* - V_p^*)V_p^*}{V_p^* + V_q^*} \right\} \tag{A·17}$$

and

$$\varepsilon_q \triangleq \frac{\varepsilon_p V_q^*}{\varepsilon_p + V_p^*}, \tag{A·18}$$

respectively. Replace V_p^* and V_q^* by $V_p^* + \varepsilon_p$ and $V_q^* - \varepsilon_q$ as a new solution. Then the new solution satisfies the condition given in Eq. (A·16) and

$$\sum_{i=1}^{N} V_i^* > (V_p^* + \varepsilon_p) + (V_q^* - \varepsilon_q) + \sum_{\substack{1 \leqq i \leqq N \\ i \neq p,q}} V_i. \tag{A·19}$$

Equation (A·19) is inconsistent with the assumption in this proof. □

We define two sets A and B as $A = \{a_1, a_2, \ldots, a_N\}$ and $B = \{b_1, b_2, \ldots, b_N\}$, respectively. Further, three sets Φ_A, Φ_B and Φ' are defined as

$$\Phi_A = \left\{ i \mid \widetilde{V}_i \in A,\ i = 1,\ldots,N \right\}, \tag{A·20}$$

$$\Phi_B = \left\{ i \mid \widetilde{V}_i \in B,\ i = 1,\ldots,N \right\} \tag{A·21}$$

and

$$\Phi' = \left\{ i \mid \widetilde{V}_i \notin A \text{ and } B,\ i = 1,\ldots,N \right\}, \tag{A·22}$$

respectively. By applying Lemma 1, we can obtain the following theorems.

Theorem 1: $\widetilde{V}_p = \widetilde{V}_q$, where $\forall p, \forall q \in \Phi'$.

Theorem 2: If Φ' is not empty, $\max_{p\in\Phi_B} \widetilde{V}_p \leqq \widetilde{V}_q$, $\forall q \in \Phi' \leqq \min_{r\in\Phi_A} \widetilde{V}_r$.

Theorem 3: If Φ' is empty, $\max_{p\in\Phi_B} \widetilde{V}_p \leqq \min_{r\in\Phi_A} \widetilde{V}_r$.

By applying Theorems 1, 2 and 3, the minimization problem given in Eq. (A·16) can be solved. We propose a subroutine to solve the problem and then show it as follows.

1. Let a set whose elements are a_i and b_i, $i = 1,\ldots,N$ be Ψ, the i-th smallest element of the set Ψ be g_i, the number of the set Ψ be N_Ψ, an iteration counter $k := 1$ and $e^* := \infty$.
2. If $k \geqq N_\Psi$, then Stop. Otherwise go to step 3.
3. Determine three sets Φ_A, Φ_B and Φ' such that

   $$\Phi_A = \{ i \mid a_i \geqq g_{k+1},\ i = 1,\ldots,N \}, \tag{A·23}$$

   $$\Phi_B = \{ i \mid b_i \leqq g_k,\ i = 1,\ldots,N \} \tag{A·24}$$

   and

   $$\Phi' = \{ i \mid a_i \leqq g_k \text{ and } b_i \geqq g_{k+1},\ i = 1,\ldots,N \}. \tag{A·25}$$

   Then let the number of elements of the set Φ' be N'.
4. Calculate P' such that

   $$P' = \sqrt[N']{\frac{K}{P_A P_B}}, \tag{A·26}$$

   where

   $$P_A = \prod_{i\in\Phi_A} a_i \tag{A·27}$$

   and

   $$P_B = \prod_{i\in\Phi_B} b_i. \tag{A·28}$$

5. If $P' < g_k$ or $P' > g_{k+1}$, then let $k := k+1$ and go to step 2. Otherwise go to step 6.

6. Calculate e such that

$$e = N' \sqrt[N']{\frac{K}{P_A P_B}} + \sum_{i \in \Phi_A} a_i + \sum_{i \in \Phi_B} b_i. \qquad \text{(A·29)}$$

7. If $e < e^*$, then e^* and V_i^*, $i = 1, \ldots, N$ are updated by

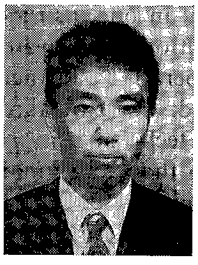
$$e^* := e \qquad \text{(A·30)}$$

and

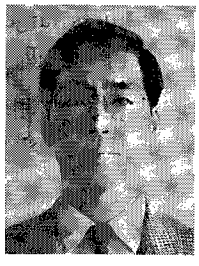
$$V_i^* := \begin{cases} a_i & i \in \Phi_A \\ b_i & i \in \Phi_B \\ P' & i \in \Phi' \end{cases}, \qquad \text{(A·31)}$$

respectively. Otherwise let $k := k+1$ and go to step 2.

Using the above subroutine, we obtain the solution V_i^*, $i = 1, \ldots, N$ of the minimization problem given by Eq. (A·16).



Mitsuhiko Yagyu was born in Osaka, Japan, on December 3, 1969. He received the B.S. and M.E. degrees from Tokyo Institute of Technology, Tokyo, Japan, in 1993 and 1995 respectively. He is currently working toward the Ph.D. degree in the Graduate School of Science and Engineering, Tokyo Institute of Technology. His main research interest is in digital signal processing.



Akinori Nishihara was born in Fukuoka, Japan, on February 26, 1951. He received the B.E., M.E. and Dr.Eng. degrees in electronics from Tokyo Institute of Technology in 1973, 1975 and 1978, respectively. Since 1978 he has been with Tokyo Institute of Technology, where he is now Professor of the Center for Research and Development of Educational Technology. His main research interests are in filter design, 1D and multi-D signal processing, and educational technology. From 1990 to 1994 he served as an Associate Editor of the IEICE Trans. Fundamentals. During 1995–1996 he was Student Activities Committee Chair, IEEE Region 10 (Asia Pacific Region). He is now serving as an Associate Editor of the IEEE Transactions on Circuits and Systems II. Dr.Nishihara is a member of IEEE, EURASIP, ECS and JET.

Nobuo Fujii received B.E. degree from Keio University, Yokohama, Japan, and M.E. and Doctor of Engineering degrees from Tokyo Institute of Technology, Tokyo, Japan, in 1966, 1968, and 1971, respectively. Since 1971, he has been with the Faculty of Engineering, Tokyo Institute of Technology where he is now a professor in the Department of Physical Electronics. From 1984 to 1985, he was a visiting scholar at the University of California, Santa Barbara. From 1990 to 1992, he served as an editor of the Transaction of the Institute of Electronics, Information and Communication Engineers, from 1995 to 1997, was the chairman of the Circuits and Systems Society of IEEE Tokyo chapter, and is now one of the chief editors of the International Journal of Analog Integrated Circuits and Signal Processing, Kluwer Academic Publishers. He is the president of the Engineering Sciences Society of IEICE and is the chairman of the technical group of electronic circuits of IEE Japan. His main interest lies in the fields of active networks, analog integrated circuits, and analog signal processing. He is the recipient of the 1983 and 1996 Best Paper Awards of the Institute of Electronics, Information Communication Engineers of Japan. He is the author of more than 10 books. Dr.Fujii is a senior member of the Institute of Electrical and Electronics Engineers, and the Institute of Electrical Engineers of Japan.