論文 / 著書情報
Article / Book Information

Title	Design Method for 2-Channel Signal Word Decomposed Filters with Minimum Output Error and Their Effective VLSI Implementation
Authors	Kouhei Hosokawa, Mitsuhiko Yagyu, Akinori Nishihara
Citation	IEICE Trans. Fundamentals., Vol. E88-A, No. 8, pp. 2044-2054
Pub. date	2005, 8
URL	http://search.ieice.org/
Copyright	(c) 2005 Institute of Electronics, Information and Communication Engineers

PAPER Special Section on Papers Selected from the 19th Symposium on Signal Processing

Design Method for 2-Channel Signal Word Decomposed Filters with Minimum Output Error and Their Effective VLSI Implementation

Kouhei HOSOKAWA[†a)], *Member*, **Mitsuhiko YAGYU**[††], *Nonmember*, *and* **Akinori NISHIHARA**[†††], *Fellow*

SUMMARY This paper proposes hardware-efficient VLSI architectures for 2-channel signal word decomposed filters (2-ch SWDFs) and their design method in consideration of the implemented circuit size. By consideration of the circuit size in design method, 2-ch SWDFs with a minimum output error among SWDFs whose size is equal to or smaller than a specification can be designed. Canonical Signed Digit expressions are used to effectively represent the filter coefficients of the SWDFs in order to make its circuit size small. Through precise analysis of the internal structures, circuit size can be accurately estimated. Some design examples show that the proposed method can design filters whose output error is about −12 dB lower than that of the linear FIR filters. Compared to an exhaustive search method, our method is much faster and can design filters whose output errors are only about 2 dB more.

key words: *signal word decomposed filters, VLSI architecture, transistor count*

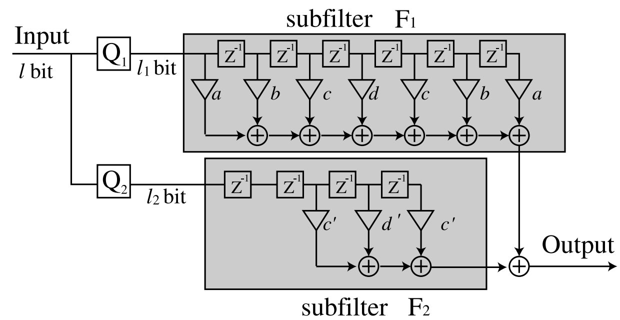


Fig. 1 Parallel structure of 2-ch SWDF with symmetric coefficients.

1. Introduction

1.1 Background

In digital signal processing, output signals of FIR filters are often rounded off, when, for example, the wordlength of the output signal is too long for a permissible noise given by specifications. We have proposed filters to aggressively incorporate this round-off operation into the FIR filters [1], [2]. These filters are called signal word decomposed filters (SWDFs) (see Fig. 1). Although an SWDF has nonlinear operations, the upper bound of its output error spectra for any input signal can be theoretically calculated. Then the boundary spectrum can be minimized and shaped by optimization in the frequency domain, and its peak can be far less than the peak output error of conventional FIR filters which carry out exact convolution, with the same circuit size. Besides, by that shaping, some priority can be given to minimizing the boundary spectrum in specified frequency bands.

References [1] and [2] showed that a hardware cost of SWDFs implemented as wired-logic can be less than that of the conventional FIR filters with the same output error. However, in the case where SWDFs are implemented as processor, SWDFs need more multiplication and addition instructions in principle than the conventional FIR filters do. Therefore, in Refs. [1], [2], and this paper, we implement an SWDF as the wired-logic employing full-adders, half-adders, and flip-flops, and evaluate the transistor count as the hardware cost.

In Refs. [1], [2], a simple arithmetic model is used to evaluate circuit size of SWDFs when designing a filter. However, these design methods cannot always design the optimum SWDFs, since this simple arithmetic model sometimes does not reflect an actual circuit size. To solve this problem we propose a method which uses an expression to estimate the transistor count that will be used in SWDFs.

1.2 Paper Organization

In Sect. 2, the variables used in this paper are defined, and we give a brief description of SWDFs output error. In addition, the output error of SWDFs is discussed.

In Sect. 3, a hardware-efficient architecture of SWDFs is proposed. This makes it possible to accurately estimate the transistor count of SWDFs.

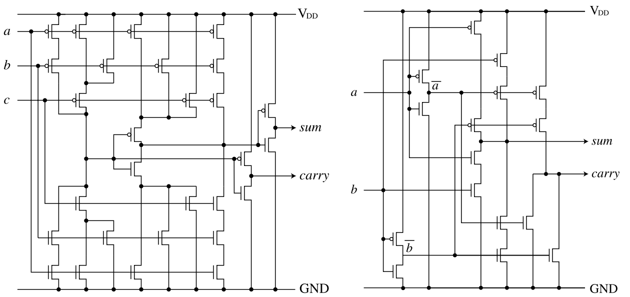
In Sect. 4, an expression that estimates the transistor count of SWDFs is derived in terms of taps T_i and the maximum coefficient wordlength c_i of each subfilter. Because, as is shown in Sect. 2, the output error can be written as a function of T_i and c_i.

In Sect. 5, a design method to minimize the output error of SWDFs under the transistor count given by specifications is proposed. Finally, in Sect. 6, filters designed by the proposed method are compared with those designed by the Remez exchange algorithm [3].

Manuscript received December 3, 2004.
Manuscript revised March 14, 2005.
Final manuscript received May 2, 2005.

[†]The author is with the NEC Media & Information Laboratories, Kawasaki-shi, 216-8666 Japan.
[††](For contact, correspond with the first author.)
[†††]The author is with the Center for Research and Development of Educational Technology, Tokyo Institute of Technology, Tokyo, 152-8552 Japan.
a) E-mail: k-hosokawa@da.jp.nec.com

DOI: 10.1093/ietfec/e88–a.8.2044



(a) Full-adder (28 Tr). (b) Half-adder (16 Tr).

Fig. 2 Conventional complementary CMOS circuits of full-adder and half-adder.

1.3 Assumptions

In this paper, we deal with 2-channel SWDFs, and the two subfilters are assumed to have the same constant group delay. The input signal is expressed by two's complement fixed number, and all filters will be designed not to have overflows. It is further assumed that $C_F \leq 2 \times C_H$ where C_F is the transistor count of one full-adder and C_H is that of one half-adder. The full-adder and half-adder circuits shown in Fig. 2 are used in evaluating the transistor count [4].

The output error is defined as a peak output error in the passband and stopband for any input signal. Such SWDFs can be applied to a prefilter of an IFIR filter [5] and a frequency response masking filter [6].

From Ref. [1], the number of taps T_i $(i = 1, 2)$ satisfies $T_1 \geq T_2$ and the coefficient wordlength c_i satisfies $c_1 \geq c_2$.

2. Output Error of SWDFs

Figure 1 shows a block diagram of a 2-ch SWDF which consists of two linear-phase FIR filters F_1 and F_2.

The ℓ-bit input signal is decomposed into two signals by the two quantizers Q_1 and Q_2. Q_1 generates high ℓ_1 bits with a sign bit from the ℓ-bit input signal, likewise Q_2 generates the remaining low $\ell_2 = \ell - \ell_1$ bits. These generated signals are processed by subfilters F_i, and the output signals of the subfilters are added together to produce the overall output.

Next, the zero-phase frequency response of the subfilter F_i is defined as $H_i(\omega) = D(\omega) + R_i(\omega),\ i = 1, 2$, where $D(\omega)$ is the ideal zero-phase frequency response given by specifications and $R_i(\omega)$ is its error response. From [1], the peak of the maximum output error $R_{\text{worst}}(\omega)$ is

$$R_{\text{worst}}(\omega) = r_1|R_1(\omega)| + r_2|R_2(\omega)|, \tag{1}$$

where r_i is the maximum amplitude of input signal of subfilter F_i.

From the assumption in Sect. 1.3, the output error of SWDFs is defined as a peak of $R_{\text{worst}}(\omega)$, that is,

$$e_p \overset{\text{def}}{=} \max_{\omega \in \substack{\text{passband and} \\ \text{stopband}}} R_{\text{worst}}(\omega). \tag{2}$$

Here, the upper bound of Eq. (2) will be considered. The coefficients of SWDFs are rounded off when the SWDFs are implemented in a real chip. In consideration of this fact, Eq. (2) is

$$e_p \leq \sum_{i=1,2} r_i \max_{\omega \in \Phi_S} |H_i(\omega) - D(\omega)| \tag{3}$$

$$\leq \sum_{i=1,2} r_i \max_{\omega \in \Phi_S} |H_i(\omega) - H_i^*(\omega)| + \sum_{i=1,2} r_i \max_{\omega \in \Phi_S} |H_i^*(\omega) - D(\omega)|, \tag{4}$$

where $H_i^*(\omega)$ is the zero-phase frequency response of the subfilter F_i with unrounded coefficients, and Φ_S is a set of frequency points in the specified passband and stopband.

Furthermore, from [7], the first term in the inequality (4) can be written as

$$\max_{\omega \in \Phi_S} \left| H_i(\omega) - H_i^*(\omega) \right| \lessapprox 2^{1-c_i} \sqrt{\frac{2T_i - 1}{3}}. \tag{5}$$

According to Ref. [1], suboptimal filter coefficients of the subfilter can be obtained by application of the Remez exchange algorithm [3] to the given T_i and c_i. From this fact and Ref. [8], the second term in the inequality (4) can be approximated to

$$\max_{\omega \in \Phi_S} \left| H_i^*(\omega) - D(\omega) \right| \approx 10^{\frac{\alpha T_i + \beta}{20}}, \tag{6}$$

where α and β are constants that depend on the cut-off frequency.

By substitution of the inequality (5) and Eq. (6) into the inequality (4), e_p is

$$e_p \lessapprox \widetilde{e}_p = \sum_{i=1}^{2} r_i \left\{ 10^{\frac{\alpha T_i + \beta}{20}} + 2^{1-c_i} \sqrt{\frac{2T_i - 1}{3}} \right\}. \tag{7}$$

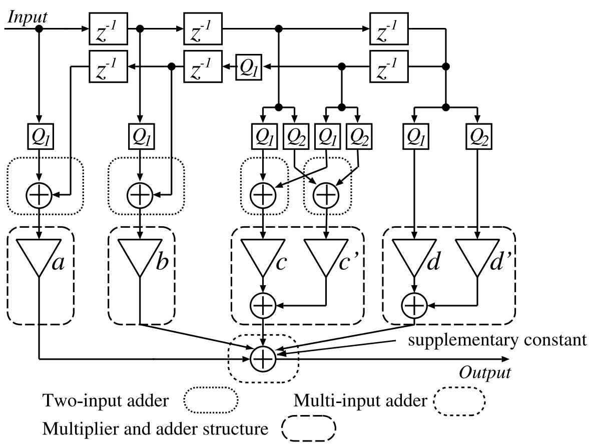
We use $\widetilde{e}_p$ in the inequality (7) as the estimation expression for the output error of SWDFs. However, The inequality (7) is only used at the designing stage of SWDFs, and Eq. (2) is used when the output error of SWDFs is evaluated.

3. Hardware-Efficient Architecture for 2-Ch SWDFs

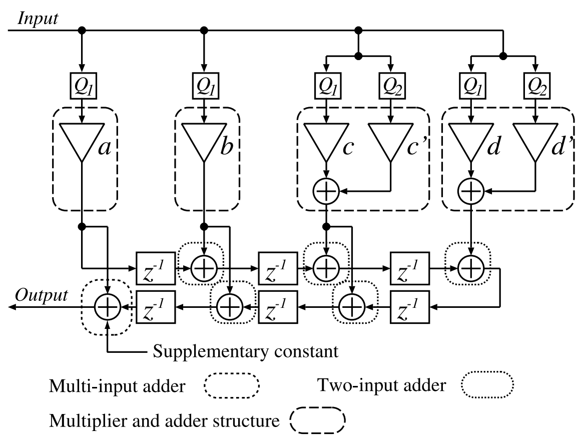
In this section, a hardware-efficient architecture for 2-ch SWDFs is proposed. This architecture is used throughout this paper for the design of SWDFs and the derivation of expressions that estimate transistor count. Figures 3(a) and (b) show hardware-efficient design examples of the direct and transposed forms of SWDFs in Fig. 1.

From these figures, it is seen that both the direct and transposed forms of SWDFs consist of 4 components—multiplier and adder structures (MASs), two-input adders, multi-input adders, and delays. The multipliers in Fig. 3 can be regarded as MASs whose one coefficient is 0.

The techniques based on sharing adders among MASs

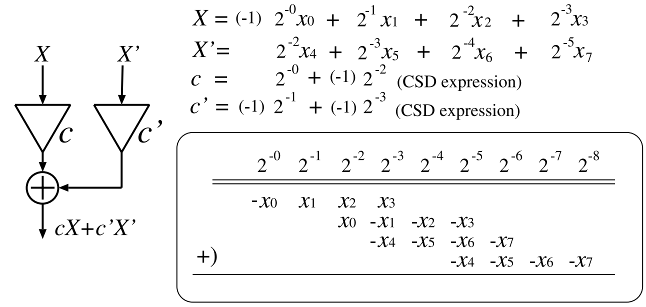


(a) Direct form. ($T_1 = 7,\ T_2 = 3$)

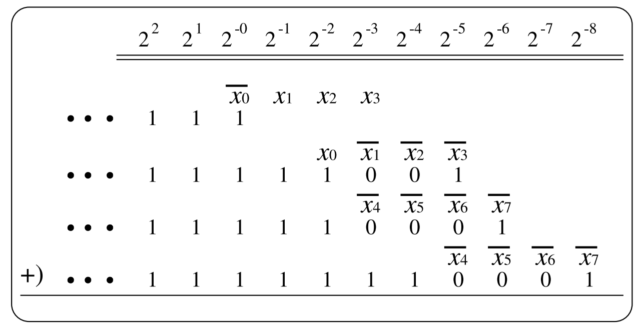


(b) Transposed form. ($T_1 = 7,\ T_2 = 3$)

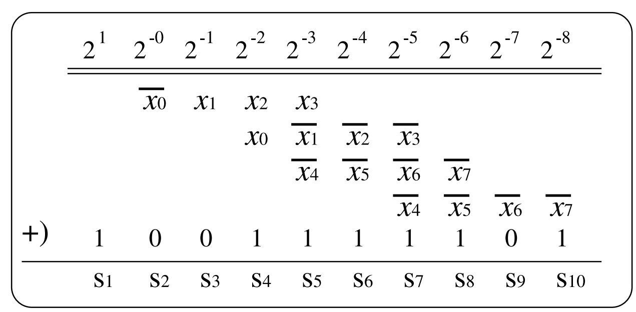
Fig. 3 Direct and transposed forms of 2-ch SWDF in Fig. 1.



(a) Example of multipliers and multiplicands.



(b) Transform subtractions into additions.



(c) Reduce additions.

Fig. 4 Design example for MAS using by CSD coefficients.

(e.g. [9], [10]) can reduce the total number of adders employed in a digital filter and may be applicable also to this SWDF architecture. In order to synthesize a 2ch-SWDF architecture by using those techniques, however, we need other algorithm for the synthesis since currently our SWDF architecture is precisely optimized at gate level. Those techniques do not optimize the gate counts, but simply optimize just the number of adders in the signal flow graph. In the future, we will consider the application of those techniques to the SWDF. In addition, those techniques can reduce the number of adders but do not uniquely determine its gate level structure. Therefore the conventional filter architecture which is optimized by the techniques [9], [10] cannot be simply compared with our optimized 2ch-SWDF architecture. We will also demonstrate this comparison in the future.

Here, we will describe a design method of these 4 components to minimize their transistor count by using full-adders, half-adders, and D-flip-flops. The design method of MASs is described at first, and that of the multi-input adder (including two-input adders) is described next. An N-bit delay can be constructed by N D-flip-flops.

3.1 MAS Structure

The canonical signed digit (CSD) expression [4] is applied to all coefficients of each multiplier in order to reduce the transistor count for MASs. The algorithm [12] can reduce the size of MASs more than CSD algorithm. However, for the same reason as [9], [10], the application of this algorithm to MASs is an issue in the future. Figure 4 is an example of synthesizing the CSD-based MAS. In this figure, the MAS multiplies the input signals X and X' by the coefficients c and c', respectively, followed by the addition of cX and $c'X'$.

The CSD expression of the coefficient (Fig. 4(a)) includes several negative bits. This example is effectively implemented by only using full-adders and half-adders as follows.

First, the bits with negative weight in Fig. 4(a) are transformed to Fig. 4(b) by introduction of bit-inversions and an addition of a constant "1." This utilizes the fact that the negative number $\cdots 0000(-x)_{(2)},\ x \in \{0, 1\}$ can be

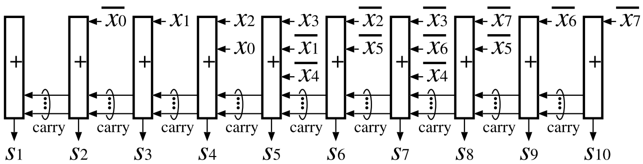


Fig. 5 Multi-input adder that realizes Fig. 4(c).

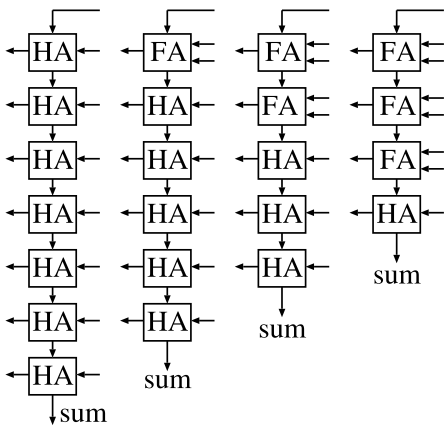


Fig. 6 Synthesis examples of a partial adder with 8 inputs.

transformed to $\cdots 0000(-x)_{(2)} = \cdots 1111\overline{x}_{(2)} + \cdots 00001_{(2)}$, where $\overline{x}$ represents inversion of x [11].

Next, Fig. 4(c) is obtained by summing up the constant numbers in Fig. 4(b). The constant numbers in all MASs of SWDFs can be calculated in advance, since the constant in Fig. 4(c) is independent of its input signal. Therefore in our proposed architecture, the sum of constant numbers in all MASs is added once at the output of SWDFs. In Fig. 3, this constant number is given as "supplementary constant."

Next, the operation without the constant numbers in Fig. 4(c) can be regarded as a multi-input adder. Therefore, a MAS in Fig. 4(c) can be effectively implemented by application of the design method for multi-input adders described in the next section.

3.2 Multi-Input Adder Structure

From the previous section, MASs with CSD coefficients can be regarded as multi-input adders. The two-input adders are naturally a kind of multi-input adders. Therefore, all three components except delays—MASs, multi-input adders, and two-input adders—are effectively implemented by using the design method for multi-input adders described below.

By using an example, the effective implementation of multi-input adders will be described below. Figure 5 shows a multi-input adder to realize Fig. 4(c) without the constant number. The multi-input adder can be implemented by cascade connections of partial adders shown in Fig. 6. A partial adder calculates the summation of all input signals to produce a 1-bit sum and several carry signals. There are many implementations for a partial adder, since a full-adder can be implemented by two half-adders as shown in Fig. 6. The multi-input adder with minimum size among many implementations can be found as follows.

First, a partial adder with $X + Y$ inputs that constructs one digit is considered, where X is the number of input signals except carry signals and Y is the number of carry signals from one digit less than that. Then, relating to a partial adder, we obtain the following three equations

$$\begin{aligned} &\text{(transistor count for a partial adder)} \\ &\quad = C_F N_F + C_H N_H, \qquad (8) \\ &\text{(the number of carries)} \\ &\quad = N_F + N_H, \qquad (9) \\ &\quad \text{and } 2N_F + N_H = X + Y - 1, \qquad (10) \end{aligned}$$

where the number of full-adders is N_F, the number of half-adders is N_H, and the transistor counts for a full-adder and a half-adder are C_F and C_H, respectively.

Next, an M digit multi-input adder is considered. The number of full-adders at each partial adder is defined as $N_F(p)$, $p = 1, \cdots, M$, and that of half-adders is defined as $N_H(p)$. In addition, $X(p)$ is defined as the number of input signals at the p-th partial adder which does not include the carry signals from the low-order digit. For example, $X(7) = 3$, $X(8) = 2$ in Fig. 5. As mentioned above, the optimization problem to minimize the transistor count for a multi-input adder can be written as Table 1.

Table 1 Optimization problem to obtain the multi-input adder with minimum transistor count.

$$\begin{aligned} &\text{Minimize} \quad C_F \sum_{p=1}^{M} N_F(p) + C_H \sum_{p=1}^{M} N_H(p), \\ &\text{Subject to} \quad 2N_F(M) + N_H(M) = X(M) - 1, \qquad (11) \\ &\qquad 2N_F(p) + N_H(p) \\ &\qquad\quad = X(p) + N_F(p+1) + N_H(p+1) - 1 \\ &\qquad\qquad \text{for } p = 1, \cdots, M-1, \\ &\qquad \text{and} \\ &\qquad \text{all the variables should be integers.} \end{aligned}$$

A multi-input adder with the least transistor count can be obtained by solving the optimization problem given by Table 1. It is found that a multi-input adder with the maximum number of full-adders is the best structure with the least transistor count since $C_F \leq 2 \times C_H$ is assumed in this paper.

4. Transistor Count Estimation of SWDFs

In this section, expressions that estimate the transistor count for each component of SWDFs are derived. Estimation expressions are classified according to whether these expressions depend on the filter coefficients or not. The expressions that do not depend on the filter coefficients—the delays and two-input adders in direct form—can be derived theoretically. Contrariwise, it is difficult to theoretically derive the other expressions—MASs, a multi-input adder in direct form, delays and two-input adders in transposed form—, since these expressions become clearly very complex and non-linear functions. Therefore, we introduce an approximation method.

In our approximation method, at first, several SWDFs are designed by the Remez exchange algorithm by the appropriate sets of c_i and T_i. Next, this approach approximates the transistor count obtained at the first step by the function of c_i and T_i. At this time, the least-square method is used.

Hereafter, the estimation expressions are explained separately for the direct and transposed forms.

4.1 Direct Form

In this subsection, the equations that estimate the transistor count for each component in the direct form SWDFs are derived separately.

4.1.1 Transistor Count of Delays

It is clear that the transistor count for delays in direct form is independent of the SWDFs filter coefficients. Therefore, the transistor count $C_{\text{delay-line}}$ is

$$C_{\text{delay-line}} = \begin{cases} C_{\text{DFF}}\left(\ell\,(T_1-1) - \ell_2\dfrac{T_1-T_2}{2}\right), & T_2>0 \\ C_{\text{DFF}}\ \ell_1(T_1-1), & T_2=0 \end{cases} \quad (12)$$

where C_{DFF} is the transistor count for a D-flip-flop, and ℓ_i is the wordlength of input signal of subfilter F_i.

4.1.2 Transistor Count of Two-Input Adders

It is clear that the input wordlength of two-input adder is ℓ_i which is independent of the SWDFs filter coefficients. The transistor count for two-input adders is

$$\sum_{i=1}^{2}(C_F(\ell_i-1)+C_H)\left\lfloor\frac{T_i}{2}\right\rfloor. \quad (13)$$

However, it is difficult to treat the floor function in Eq. (13) correctly in the optimization problem because of its discontinuities. Therefore, we simply approximate the above by eliminating the floor function as follows;

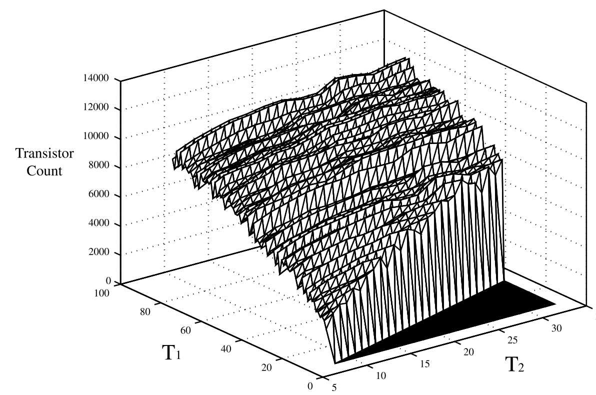
$$C_{\text{adder}} = \sum_{i=1}^{2}(C_F(\ell_i-1)+C_H)\frac{T_i}{2}. \quad (14)$$

4.1.3 Transistor Count of MASs

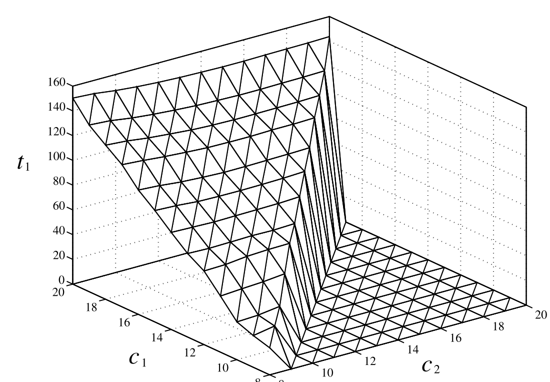
As we mentioned earlier, it is difficult to estimate the correct transistor count for MASs since MASs depend on the filter coefficients. The estimation expression for the transistor count for MASs is obtained by using a number of design samples by the algorithm in Table 2. Then, the number of design samples to obtain the expression can be written as

Table 2 Approximation method for MAS.

```
for(c1 = c1s; c1 < c1e; c1+=c1step){
  for(c2 = c2s; c2 < c2e; c2+=c2step){
    for(t1 = t1s; t1 < t1e; t1+=t1step){
      for(t2 = t2s; t2 < t2e; t2+=t2step){
        MAS[t1][t2] = Calculate_MAS_size(c1,c2,t1,t2);
      }
    }
  temp[c1][c2] = first_order_
            approximation(MAS); ⇐ approximation 1
  }
}
M = first_order_approximation(temp);⇐ approximation 2
```



(a) Transistor count for the SWDF's MASs that are designed with various T_1, T_2. The $c_1 = 14$, $c_2 = 8$, passband is 0–0.2, and stopband is 0.25–0.5.



(b) An example of how $t_1(c_1, c_2)$ changes.

Fig. 7 Approximation example of MASs transistor count.

$$\prod_{i=1}^{2}\frac{c_{ie}-c_{is}}{c_{istep}} \times \prod_{i=1}^{2}\frac{t_{ie}-t_{is}}{t_{istep}}. \quad (15)$$

The reason the expression is obtained by the two steps in Table 2 is that we can easily confirm the approximation accuracy in each step.

At first, Fig. 7(a) shows how the transistor count for MASs varies when T_1 and T_2 are varied at $(c_1, c_2) = (14, 8)$. As shown in Fig. 7(a), the transistor count for MASs can be approximated by a first order plane

$$C_{\text{MAS}} = t_1(c_1,c_2)T_1 + t_2(c_1,c_2)T_2 + t_3(c_1,c_2). \quad (16)$$

This figure shows the validity of "approximation 1" in Table 2. We define Eq. (16) as the estimation expression for MASs transistor counts when c_1 and c_2 are fixed.

Next, Fig. 7(b) shows how $t_1(c_1, c_2)$ varies when c_1 and c_2 are varied. As shown in the figure, $t_1(c_1, c_2)$ can be approximated by a first order plane

$$t_1(c_1, c_2) = M_{11}c_1 + M_{12}c_2 + M_{13}. \tag{17}$$

This figure shows the validity of "approximation 2" in Table 2. The t_2 and t_3 can also be approximated by first order planes.

From the facts mentioned above, the following equation is obtained by substituting Eq. (17) into Eq. (16)

$$C_{\text{MAS}} = (\mathbf{MC})^T\mathbf{T}, \tag{18}$$

where each element of $\mathbf{M}$ is defined as M_{ij} in Eq. (17), $\mathbf{C}$ is $[c_1\ c_2\ 1]^T$, and $\mathbf{T}$ is $[T_1\ T_2\ 1]^T$. We define Eq. (18) as the estimation equation for MASs transistor count when $T_2 > 0$.

However, the most suitable SWDFs sometimes do not have a subfilter F_2, that is, this filter is only composed of the subfilter F_1. Then the transistor count for MASs can be approximated accurately by the second order function of c_1 and T_1. We estimate the transistor count for MASs by another function $(M'_{11}c_1 + M'_{13})T_1 + (M'_{31}c_1 + M'_{33})$ when $T_2 = 0$.

As mentioned above, the estimation equation for MASs transistor counts can be written as

$$C_{\text{MAS}} = \begin{cases} (\mathbf{MC})^T\mathbf{T}, & T_2 > 0 \\ (\mathbf{M'C})^T\mathbf{T}, & T_2 = 0, \end{cases} \tag{19}$$

where each element of $\mathbf{M'}$ is M'_{ij}, and M'_{12}, M'_{32}, and M'_{2j}, $j = 1, 2, 3$ is defined as 0.

4.1.4 Transistor Count of a Multi-Input Adder

As we mentioned earlier, similarly to MASs, the transistor count for the multi-input adder is expressed by the approximation function since its transistor count depends on the filter coefficients. However, unlike MASs, the problem to obtain the approximation function of the estimation expression for SWDF's multi-input adder can be reduced to the problem of the linear FIR filter which is not decomposed so that the number of design samples to obtain the approximate function is reduced. Hereafter, the term "linear FIR filter" is used to mean a linear phase FIR filter which is not decomposed. Therefore, compared with that of MASs, the number of design samples to obtain the estimation expression can be dramatically decreased to

$$\frac{c_e - c_s}{c_{step}} \times \frac{t_e - t_s}{t_{step}}. \tag{20}$$

The transistor count for the multi-input adder depends on the wordlength and the number of input signals. The wordlength of multi-input adder's input signal, that is, the wordlength of MAS's output signal will be considered below. The number of input signals is T_1 from Sect. 1.3.

Figure 8 describes the wordlength of MAS's output

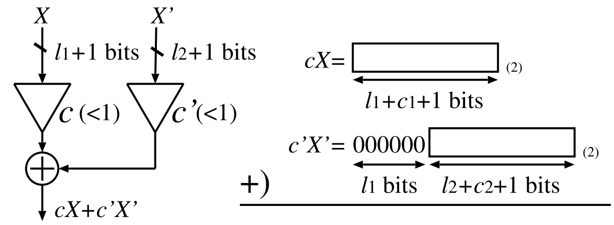


Fig. 8 Output wordlength of an MAS in direct form.

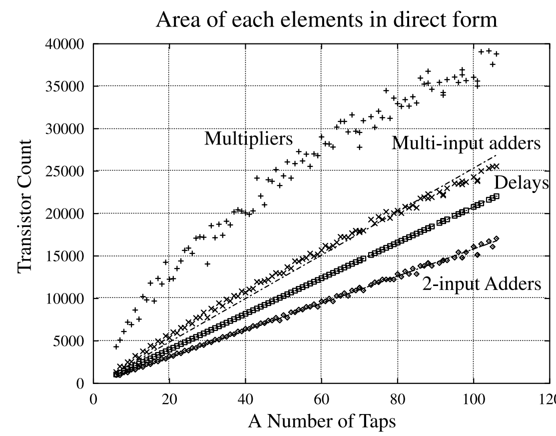


Fig. 9 Transistor count of each element in a direct form linear FIR filter with $c = 14$, $l = 16$, passband is 0–0.2, and stopband is 0.25–0.5.

signal in direct form. The two input signals X and X' are the same as the output of two-input adders. So, their wordlengths are $\ell_1 + 1$ and $\ell_2 + 1$ bits, respectively. If $T_2 = 0$ or $c_1 \geq \ell_2 + c_2$, that is, $\ell_1 + c_1 + 1 \geq \ell_1 + \ell_2 + c_2 + 1$, the wordlength of MAS's output signals is $\ell_1 + c_1 + 1$ bits from Fig. 8. Contrariwise, if $T_2 > 0$ and $c_1 < \ell_2 + c_2$, that is, $\ell_1 + c_1 + 1 < \ell_1 + \ell_2 + c_2 + 1$, the wordlength of MAS's output signals is $\ell_1 + \ell_2 + c_2 + 1$ bits.

(1) In the case of $T_2 = 0$ or $c_1 \geq \ell_2 + c_2$

If the multi-input adder of SWDFs meets this condition, the wordlength of multi-input adder's input signal is $\ell_1 + c_1 + 1$ bits and the number of input signals is T_1. It can be inferred that the transistor count for the multi-input adder of SWDFs is the same as the multi-input adder in a T_1-tap linear FIR filter whose input signal and coefficient wordlength are ℓ_1 and c_1 bits, respectively. This is because the wordlength and number of input signals in the multi-input adder of this linear FIR filter are $\ell_1 + c_1 + 1$ bits and T_1, respectively.

From Fig. 9, the transistor count for the multi-input adder in this linear FIR filter can be approximated by $a_1(T_1 - 1)$. In addition, a_1 can be approximated by $a_1 = u_1c_1 + u_2$ since a_1 varies as shown in Fig. 10 when c_1 varies. Therefore, the transistor count for the multi-input adder of this linear FIR filter, that is, the transistor count for the multi-input adder of SWDFs can be written as

$$(u_1c_1 + u_2)(T_1 - 1). \tag{21}$$

In the case of SWDFs, there is a possibility that $c'X'$ in

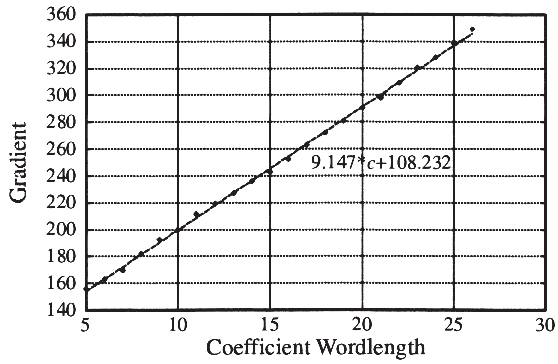


Fig. 10 The variation of a_1 in a linear FIR filter.

Fig. 8 becomes 0. However, the output wordlength of MAS is not changed in this case since the output wordlength of MASs is only decided by cX. Therefore, it is not necessary to consider the effect of $c'X' = 0$ in this case.

(2) In the case of $T_2 > 0$ and $c_1 < \ell_2 + c_2$

Contrariwise, from Fig. 8, the wordlength of the output signal of MASs is $\ell_1 + \ell_2 + c_2 + 1$ bits in the case of $T_2 > 0$ and $c_1 < \ell_2 + c_2$. From the same explanation mentioned above, it can be inferred that the transistor count for the multi-input adder in SWDFs is the same as the multi-input adder in a T_1-tap linear FIR filter whose input signal and coefficient wordlength are $\ell_1 + \ell_2$ and c_2 bits, respectively. Therefore, the transistor count for multi-input adder can also be written as

$$(u_1' c_2 + u_2')(T_1 - 1). \tag{22}$$

However, unlike the previous case, the $(\ell_1 + \ell_2 + c_2 + 1) - (\ell_1 + c_1 + 1) = \ell_2 + c_2 - c_1$ bits among the output signal of MASs with $c'X' = 0$ are fixed to 0. And, the number of MASs with $c'X' = 0$ is $(T_1 - T_2)/2$. Therefore, the transistor count for the multi-input adder in SWDFs can be written as

$$(u_1' c_2 + u_2')(T_1 - 1) - \Delta, \tag{23}$$

where Δ is the effect of the MASs with $c'X' = 0$. It will be explained in Appendix that Δ can be written as a second order function of c_1, c_2, T_1 and T_2.

(3) Summary

As mentioned above, the estimation expression of transistor count for the multi-input adder $C_{\text{adder-line}}$ can be written as

$$C_{\text{adder-line}} = \begin{cases} (u_1' c_2 + u_2')(T_1 - 1) - \Delta & \text{for } T_2 > 0 \text{ and} \\ & \quad c_1 < \ell_2 + c_2 \\ (u_1 c_1 + u_2)(T_1 - 1) & \text{otherwise} \end{cases} \tag{24}$$

4.1.5 Summary of Direct Form

From Eq. (12), (14), (19) and (24), the estimation expression for direct form can be written as

$$C_{\text{DFF}} + C_{\text{adder}} + C_{\text{MAS}} + C_{\text{adder-line}}. \tag{25}$$

Equation (25) is the second order function of c_1, c_2, T_1 and T_2.

4.2 Transposed Form

In this section, the estimation expression for transistor count of the transposed form of SWDFs is derived. In the case of the transposed form, the analysis method used in direct form can be used. The different points in the transposed form are the estimation expressions for delays and two-input adders. In the case of MASs, the method used in direct form can be used without modification.

4.2.1 Transistor Count of Delays and Two-Input Adders

These equation can be obtained by using the similar method for the multi-input adder in direct form.

(1) In the Case of $T_2 = 0$ or $c_1 \geq \ell_2 + c_2$

In this case, the wordlength of MASs output signal is $\ell_1 + c_1$ bits. Therefore, it can be inferred that the transistor counts for delays and two-input adders in SWDFs are the same as these delays and two-input adders in a T_1-tap linear FIR filter whose input signal and coefficient wordlength are ℓ_1 and c_1 bits, respectively.

As shown in Fig. 11(a), the transistor count for each element in a linear FIR filter can be approximated accurately by a first order function. In addition, Fig. 11(b) shows how the approximate coefficient varies when its coefficient wordlength varies. From this figure, the approximate coefficients can also be approximated by a first order function.

In the case of $c'X' = 0$, similarly to direct form, the transistor count of delays and two-input adders are not changed since the wordlength of MAS's output signal is not changed.

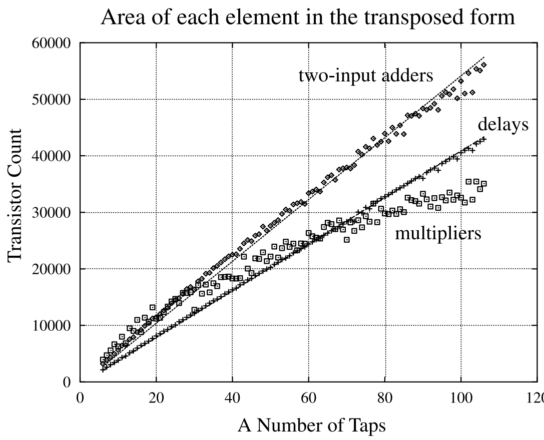
(2) In the Case of $T_2 > 0$ and $c_1 < \ell_2 + c_2$

Contrariwise, the wordlength of the output signal of MASs is $\ell_1 + \ell_2 + c_2$ bits in the case of $T_2 > 0$ and $c_1 < \ell_2 + c_2$. Therefore, it can be inferred that the transistor counts for the delays and two-input adders in SWDFs are the same as the delays and two-input adders in a T_1-tap linear FIR filter whose input signal and coefficient wordlengths are $\ell_1 + \ell_2$ and c_2 bits, respectively.

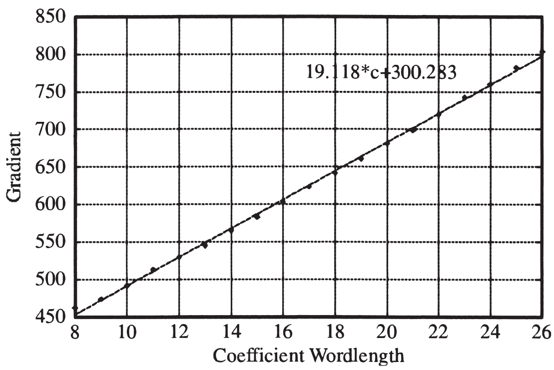
However, if a MAS with $c'X' = 0$, the lower $(\ell_1 + \ell_2 + c_2) - (\ell_1 + c_1) = \ell_2 + c_2 - c_1$ bits of a MAS's output signal are fixed to 0. Therefore, it can be inferred that the transistor count for the delays is decreased from the linear FIR filter by $C_{\text{DFF}}(\ell_2 + c_2 - c_1)(T_1 - T_2)/2$ since the number of MASs with $c'X' = 0$ is $(T_1 - T_2)/2$. Likewise, it can be inferred that the transistor count for the two-input adders is decreased by $C_F(\ell_2 + c_2 - c_1)(T_1 - T_2 - 1)$ transistors.

(3) Summary

As mentioned above, the estimation expression for delays is



(a) Each element's transistor count in the transposed form of linear FIR filter.



(b) Variation of the approximate coefficient of 2-input adder in the transposed form of linear FIR filter.

Fig. 11 Figures are related to the transistor count in the transposed form of linear FIR filter.

$$C_{\text{delay-line}} = \begin{cases} (d_1' c_2 + d_2')(T_1 - 1) - \Delta_d & \text{for } T_2 > 0 \text{ and } c_1 < \ell_2 + c_2 \\ (d_1 c_1 + d_2)(T_1 - 1) & \text{otherwise} \end{cases} \quad (26)$$

where Δ_d is equal to $C_{\text{DFF}}(\ell_2 + c_2 - c_1)(T_1 - T_2)/2$.

Likewise, the estimation equation for two-input adders is

$$C_{\text{adder}} = \begin{cases} (a_1' c_2 + a_2')(T_1 - 1) - \Delta_a & \text{for } T_2 > 0 \text{ and } c_1 < \ell_2 + c_2 \\ (a_1 c_1 + a_2)(T_1 - 1) & \text{otherwise} \end{cases} \quad (27)$$

where Δ_a is equal to $C_F(\ell_2 + c_2 - c_1)(T_1 - T_2 - 1)$.

5. Formulation of Optimization Problem

As mentioned above, the optimization problem to obtain the SWDF that has a minimum output error among SWDFs whose transistor count is below A_c, which is given by specification, can be formalized as

$$\begin{aligned} &\text{Minimize} && \sum_{i=1}^{2} r_i \left\{ 10^{\frac{\alpha T_i + \beta}{20}} + 2^{1-c_i} \sqrt{\frac{2T_i - 1}{3}} \right\} \\ &\text{Subject to} && c_1 \geq c_2, T_1 \geq T_2 \qquad (28) \\ & && C_{\text{delay-line}} + C_{\text{adder}} + C_{\text{adder-line}} \\ & && \qquad + C_{\text{MAS}} \leq A_c \end{aligned}$$

where $C_{\text{adder-line}}$ is 0 in the transposed form.

However, Eq. (28) is decomposed into three optimization problems since $C_{\text{delay-line}}$, C_{adder}, $C_{\text{adder-line}}$ and C_{MAS} have discontinuity. Therefore, the solution of Eq. (28) is the solution with a minimum output error among the three optimization problem's solutions.

6. Design Example

In this section, the transistor count and output error of SWDFs designed by two methods—the proposed method and the exhaustive search algorithm—and those of the linear FIR filter are compared. In the proposed method, we use the MATLAB optimization toolbox [15] to solve the non-linear optimization problem in Eq. (28).

In the exhaustive search algorithm, we first design all SWDFs by using all conceivable taps and coefficient wordlengths. Next, the SWDF that has a minimum transistor count is selected and plotted among the SWDFs that meet the output error of the specification. Through this algorithm, the specification of the output error is updated in increments of 1 dB in each design.

6.1 Design Specifications

The design specifications are that ℓ_1, ℓ_2 are 8 bits, passband is 0.0–0.2, and stopband is 0.25–0.5. In addition, the transistor count for a full-adder, half-adder, D-flip-flop, and an inverter are assumed to be 28, 16, 16, and 2, respectively.

The designed filters, including SWDFs and linear FIR filters, are estimated by using EX-OR logic to calculate the sum at the MSB in all adders, since full-adders and half-adders waste the transistor count here. Then, the transistor count for 2-input EX-OR and 3-input EX-OR are assumed to be 12 and 22, respectively.

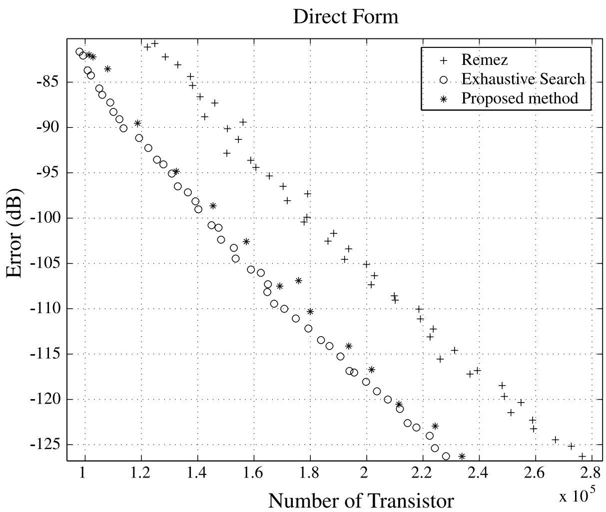
6.2 Proposed Method

The proposed method uses 527 design samples to obtain the approximate expressions. That is, c_{1step}, c_{2step}, t_{1step}, and t_{2step} in Table 2 are 5. In the proposed method, SWDFs are designed by A_c in increments of 10000 in Eq. (28).

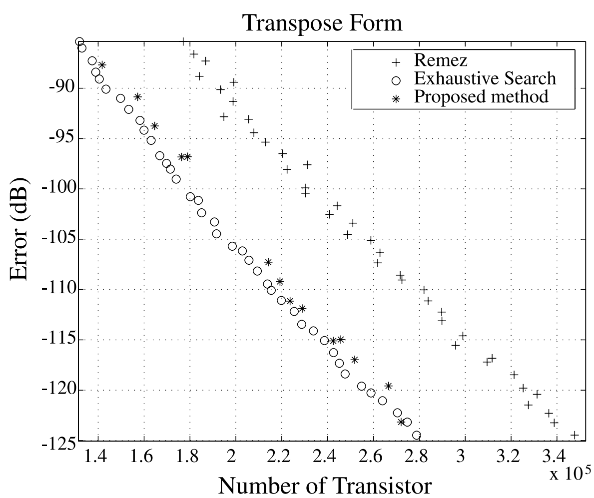
6.3 Results

Figure 12 shows the SWDFs and linear FIR filters whose output error is in the range from about −80 dB to −125 dB.

From these figures, we found that the proposed method



(a) Error e_p vs. transistor count in direct form.



(b) Error e_p vs. transistor count in transposed form.

Fig. 12 Comparison of filters designed by the proposed methods and the linear FIR filters.

can design an SWDF whose output error is about −12 dB lower than that of the linear FIR filters with the same transistor count. In other words, the transistor count of the SWDFs designed with the proposed method are about 18% less than that of the linear FIR filters with the same output error.

In addition, the proposed method can design SWDFs that are close to the optimum filters (exhaustive-search algorithm). The average output error difference between the proposed and optimum SWDFs is only about 2 dB.

Tables 3 and 4 show the computing time for each design when the Intel Xeon 2.4 GHz processor is used. Compared to the Remez exchange algorithm, the proposed method consumes larger computing time. However, the proposed method can design SWDFs in practical computing time, and about 4350 times faster than the exhaustive search algorithm.

Table 3 Comparison of computing time among all algorithms in direct form.

Method	Time (sec)
Remez	< 1
Proposed	7.15
Exhaustive Search	30987

Table 4 Comparison of computing time among all algorithms in transposed form.

Method	Time (sec)
Remez	< 1
Proposed	6.98
Exhaustive Search	30375

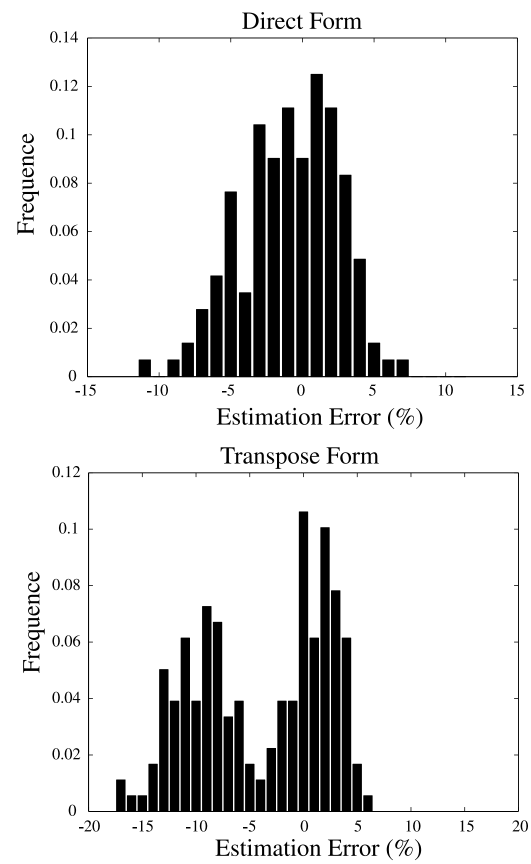


Fig. 13 Estimation error distribution of Eq. (28).

6.4 Estimation Error of Transistor Count

In this section, the difference between the transistor count estimated by Eq. (28) and actually evaluated transistor count is considered. Figure 13 shows the estimation errors in the direct and transposed forms in 144 and 179 design examples, respectively.

The average of the estimation error is 2.75% in direct form, and 5.61% in transposed form. This result shows that the estimation error of transposed form is worse than that of direct form. We conjecture the reason that the transposed form has more components which depend on the filter coefficient than the direct form does.

7. Conclusion

In the conventional SWDF design method, a simple arithmetic model is used to evaluate circuit size. However, these design methods cannot always design most suitable SWDFs, since such simple arithmetic models sometimes are based on inaccurate assumptions and so sometimes seriously degrade the estimation of actual circuit size.

In the proposed method, this problem is solved by using a correct size model. First, the hardware-efficient architecture was introduced to implement an SWDF, and a method to accurately estimate its circuit size was proposed. Next, the circuit size for each sub-component of the SWDF was considered, and then an expression to estimate its subcomponent was derived from the coefficient wordlength and taps of the SWDF. Through this proposed estimation, the proposed method can design SWDFs whose output error is about −12 dB lower than that of conventional linear FIR filters. In addition, the average output error difference between the proposed and optimum SWDFs is only about 2 dB.

At the optimization problem in this paper, the coefficient wordlength and taps are treated as real number. Therefore, the solutions in this optimization problem are rounded off. In the future, the optimization problem will be solved as an integer optimization problem. In addition, we will research that coefficients of each subfilter will be optimized directly over the discrete space by fully utilizing the design freedom.

References

[1] M. Yagyu, A. Nishihara, and N. Fujii, "Minimization of output errors of FIR digital filters by multiple decompositions of signal word," IEICE Trans. Fundamentals, vol.E81-A, no.3, pp.407–419, March 1998.
[2] M. Yagyu and A. Nishihara, "Design of signal word decomposed filters with optimal filter length and coefficient wordlength assignment to each subfilter," European Conference on Circuit Theory and Design, pp.269–272, Stresa, Sept. 1999.
[3] T.W. Parks and J.H. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," IEEE Trans. Circuit Theory, vol.CT-19, no.2, pp.189–194, March 1972.
[4] P. Pirsch, Architectures for Digital Signal Processing, Wiley, 1996.
[5] Y. Neuvo, C.Y. Dong, and S.K. Mitra, "Interpolated finite impulse response filters," IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-32, no.3, pp.563–570, June 1984.
[6] Y.C. Lim, "Frequency-response masking approach for the synthesis of sharp linear phase digital filters," IEEE Trans. Circuits Syst., vol.CAS-33, no.4, pp.357–364, April 1986.
[7] D.S.K. Chan and L.R. Rabiner, "Analysis of quantization errors in the direct form for finite impulse response digital filters," IEEE Trans. Audio Electroacoust., vol.AU-21, no.4, pp.354–366, Aug. 1973.
[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] D.R. Bull and D.H. Horrocks, "Primitive operator digital filters," Proc. Inst. Elec. Eng. G: Circuits, Devices, Syst., vol.138, no.3, pp.401–412, June 1991.
[10] A.G. Dempster and M.D. Macleod, "Use of minimum-adder multiplier blocks in FIR digital filters," IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process., vol.42, no.9, pp.569–577, Sept. 1995.
[11] S. Kim, J. Chung, and K.K. Parhi, "Low error fixed-width CSD multiplier with efficient sign extension," IEEE Trans. Circuits Syst. II, Analog. Digit. Signal Process., vol.50, no.12, pp.984–993, 2003.
[12] A.G. Dempster and M.D. Macleod, "Constant integer multiplication using minimum adders," Proc. Inst. Elec. Eng. Circuits, Devices, Syst., vol.141, no.5, pp.407–413, Oct. 1994.
[13] M. Yagyu, A. Nishihara, and N. Fujii, "Analytical guess of minimum output error of signal word decomposed filters," IEICE Technical Report, CAS97-90, Jan. 1998.
[14] 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, no.8, pp.1391–1402, Aug. 1997.
[15] The MathWorks, Inc. http://www.mathworks.com/

Appendix: Estimation of Decreased Transistor Count of an Multi-Input Adder

In this section, Δ in Eq. (23) is considered. This Δ results from decreasing the output wordlength of MAS from $\ell_1 + \ell_2 + c_2 + 1$ to $\ell_1 + c_1 + 1$ bits, since its lower bits are fixed to zero.

A.1 Decreased Transistor Count by Fixing One Bit to Zero in a Multi-Input Adder

At first, the decreased transistor count is considered when one bit at d-th digit from MSB is fixed to 0 among the input signals of the multi-input adder.

At first, we consider the d-th partial adder. In this partial adder, the number of half-adders is one and the rest are full-adders if the number of input signals including carry signals is even. Contrariwise, if the number of input signals is odd, this partial adder consists of only full-adders, that is, there is no half-adder.

If the probability that the number of input signals is even is assumed to 1/2, the transistor count decreases by $(C_F - C_H)/2 + C_H/2 = C_F/2$ because one bit is fixed to zero at d-th partial adder. In addition, the carry signals to the $(d-1)$-th partial adder decrease by 1 with a 1/2 probability. That is, the number of input signals at the $(d-1)$-th partial adder decrease by 1 with a 1/2 probability. By recursively iterating this process, the transistor count decreases by

$$\frac{C_F}{2}\sum_{n=0}^{d-1}\frac{1}{2^n}, \qquad \text{(A· 1)}$$

in the case that one input signal at the d-th partial adder is fixed to 0.

From Fig. 8, the bit place of d meets

$$d \geq l_1 + c_1 + 2. \qquad \text{(A· 2)}$$

That is, d is large enough. Therefore, Eq. (A· 1) can be written approximately as

$$\lim_{d\to\infty}\frac{C_F}{2}\sum_{n=0}^{d-1}\frac{1}{2^n} = C_F. \qquad \text{(A· 3)}$$

The decreased transistor count is C_F if one input signal at multi-input adder is fixed to 0.

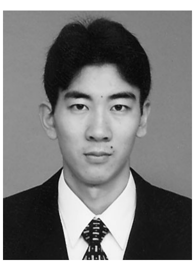
A.2 Value of Δ

In direct form, the number of input signals fixed to 0 at multi-input adders is

$$\frac{T_1 - T_2}{2} \times (\ell_2 + c_2 - c_1).$$

Therefore, the Δ in Eq. (23) can be written as

$$\Delta = C_F \times (l_2 + c_2 - c_1)\frac{T_1 - T_2}{2}. \tag{A$\cdot$4}$$



Kouhei Hosokawa received the B.E. and M.E. degrees in electronics from Tokyo Institute of Technology, Tokyo, Japan, in 1999 and 2001, respectively. He joined Media & Information Research Labs., NEC Corporation in 2001 and has been working in the area of VLSI design technology since then.

Mitsuhiko Yagyu His biography and photograph are not available.

Akinori Nishihara 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 Professor of the Center for Research and Development of Educational Technology. His main research interests are in signal processing, and its application to educational technology. He served as an Associate Editor of many international journals, including the IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences from 1990 to 1994. He served in IEEE Region 10 Executive Committee, as Student Activities Committee Chair (1995–1996), Treasurer (1999–2000), and Educational Activities Committee Chair (2001–2004). He was also an Executive Committee Member of IEEE Tokyo Section and IEEE Japan Council for the last 10 years. He is serving as a member of the Board of Governors, IEEE Circuits and Systems Society since 2004. He was Chair of the IEICE Technical Group on Circuits and Systems from 1997 to 1998, and since 1998 he has been serving as an Advisor of that Technical Group. He received Best Paper Award of the IEICE in 1999, and IEEE Third Millennium Medal in 2000. He also received 4th LSI IP Design Award in 2002. Prof. Nishihara is a Fellow of IEEE, and a member of EURASIP, European Circuits Society and Japan Society for Educational Technology.