High-Performance VLSI Architecture Using Distributed Arithmetic for Higher-Order FIR Filters with Complex Coefficients

Yoshitaka Tsunekawa[1], Takeshi Nozaki[2], and Norio Tayama[3]

[1,2,3]Department of Electrical and Electronic Engineering, Faculty of Engineering,
Iwate University, Morioka, Japan
Tel. and Fax. +81-19-621-6468, E-mail: [1]tsune@iwate-u.ac.jp, [2]t5300005@iwate-u.ac.jp

Abstract: This paper proposes a high-performance VLSI architecture using distributed arithmetic for higher-order FIR filters with complex coefficients. For the purpose of realizing high sampling rate with small latency in high-order filters, we apply distributed arithmetic[1]. Moreover, in order to decrease drastically the power dissipation, the structure applying not ROM's but optimum function circuits which we have previously proposed, is utilized[2][3]. However, this structure increases in the number of adders as compared to the conventional structure applying ROM's. In order to realize a more effective method for further higher-order filter, we propose newly an implementation applying two methods which have large effects on the unit using the adders. First, we propose an implementation applying SFAs(Serial Full Adders) and SFSs(Serial Full Subtractors). Second, we propose a structure applying proposed 4-2 adders. Finally, it is shown that the proposed architecture is an effective way to realize low power dissipation and small latency while the sampling rate is kept constant for further higher-order filters with complex coefficients.

1. Introduction

In recent years, the efficient implementation of complex digital arithmetic has been increasingly important, because many signal-processing applications require processing of complex signals with complex digital filters. Examples of this can be found in baseband processing for narrow-band RF signals in homomorphic speech processing, in spectral analysis and in matched filtering for coherent radars. Also, as a number of features in complex filters have been gradually evidenced, the complex filters have attracted attention. One of the features is that the filters can obtain the same characteristic of transform function by a half order in comparison with real FIR filters.

As a general implementation of complex FIR filters, there is the way based on direct form structure using the multipliers. This can obtain high sampling rate by applying basically pipeline processing to each tap. However, the latency increases significantly for relatively high-order filters. Also, it uses four multipliers and two adders per tap, so that it requires enormous hardware complexity.

We propose a high-performance VLSI architecture for higher-order FIR filters with complex coefficients. For the purpose of realizing high sampling rate with small latency, we consider distributed arithmetic of which the processing time depends on only word length[1]. The conventional structure using ROM's is effective for low-order filters and real coefficients. However, it requires enormous power dissipation for high-order filters and complex coefficients. In order to decrease drastically the power dissipation, we utilize the new structure based on the optimum function circuits using logic gates[2][3]. However, the structure increases in the number of adders as compared with that using ROM's[3]. In order to realize a more effective method for higher-order filters, we propose newly an implementation applying two methods which have large effects on the unit using the adders.

First, we will take account of linear-phase characteristic of the complex FIR filters and propose a method applying SFAs and SFSs for real part and imaginary part, respectively. Using them, we can decrease the power dissipation drastically. Second, in order to realize smaller latency with low power dissipation, we will apply not the conventional structure based on CLA(Carry Look-Ahead Adder), but a new structure based on proposed 4-2 adders. Consequently, it will be shown that the proposed architecture is an effective way to realize low power dissipation and small latency while the sampling rate is kept constant for further higher-order filters with complex coefficients.

2. Complex FIR Filters Based on Distributed Arithmetic

When the complex coefficients and the complex input variables are $a = (a_{R1} + ja_{I1}, \ldots a_{RN} + ja_{IN})$ and $v = (v_{R1} + jv_{I1}, \ldots v_{RN} + jv_{IN})$, respectively, the input and output relationship of N-tap complex FIR filters become

$$y = \sum_{i=1}^{N} \{(a_{Ri}v_{Ri} - a_{Ii}v_{Ii}) + j(a_{Ii}v_{Ri} + a_{Ri}v_{Ii})\} \quad (1)$$

where $-1 \leq v_{Ri} < 1$ and $-1 \leq v_{Ii} < 1$. For the purpose of obtaining small latency with high sampling rate in high-order filter, we consider distributed arithmetic of which the processing time depends on only word length[1]. This arithmetic is the method in which the inner product of the constant coefficients are calculated by the table-lookup. Eq.(1) can be regarded as the inner product.

Consider the application of the distributed arithmetic. The v_{Ri} and v_{Ii} in Eq.(1) are the two's complement representation of B-bit fixed-point number. These variables are expressed as

$$v_{Ri} = -v_{Ri}^0 + \sum_{k=1}^{B-1} 2^{-k}v_{Ri}^k, v_{Ii} = -v_{Ii}^0 + \sum_{k=1}^{B-1} 2^{-k}v_{Ii}^k \quad (2)$$

where v_{Ri}^k, v_{Ii}^k are the respective k-th bit of v_{Ri} and v_{Ii}, and are either 0 or 1. Substituting Eq.(2) for Eq.(1), it follows that

$$y = y_R + jy_I, \tag{3}$$

$$y_R = -\Phi_{R1}^0 + \sum_{k=1}^{B-1} 2^{-k}\Phi_{R1}^k - \Phi_{R2}^0 + \sum_{k=1}^{B-1} 2^{-k}\Phi_{R2}^k, \tag{4}$$

$$y_I = -\Phi_{I1}^0 + \sum_{k=1}^{B-1} 2^{-k}\Phi_{I1}^k - \Phi_{I2}^0 + \sum_{k=1}^{B-1} 2^{-k}\Phi_{I2}^k \tag{5}$$

where Φ_{R1}^k, Φ_{R2}^k, Φ_{I1}^k, and Φ_{I2}^k are

$$\Phi_{R1}^k = \sum_{i=1}^N a_{Ri}v_{Ri}^k, \qquad \Phi_{R2}^k = -\sum_{i=1}^N a_{Ii}v_{Ii}^k,$$

$$\Phi_{I1}^k = \sum_{i=1}^N a_{Ii}v_{Ri}^k, \qquad \Phi_{I2}^k = \sum_{i=1}^N a_{Ri}v_{Ii}^k. \tag{6}$$

The equations from (3) to (6) show that the distributed arithmetic can be applied to the complex FIR filters. The structure based on this arithmetic is composed of three units which are called input unit, functional generation unit, and functional addition unit, respectively. In input unit, a chain of serial shift registers provides the tap delays and propagates the bits of input variables to ROM of the functional generation unit. The inner product Φ of the bits of the input variables and coefficients is stored in ROM. When the calculation is performed, this adds one data transferred from ROM and the other data which is shifted the previous accumulated value to the right. In this way, it has the advantage that the processing time depends on word length without the multipliers. However, in high-order FIR filters, it requires enormous power dissipation which is caused by a large scale of ROM.

There is the division of function Φ as a means of decreasing the scale of ROM. This method can realize a large scale of function Φ as the addition of small scale of functions. When the number of division is Q, the function Φ is expressed as

$$\Phi(v_1^k, \ldots, v_N^k) = \sum_{i=1}^{N/Q} a_i v_i^k + \cdots + \sum_{i=Q'}^N a_i v_i^k$$

$$= \Phi(v_1^k, \ldots, v_{N/Q}^k) + \cdots + \Phi(v_{Q'}^k, \ldots, v_N^k) \tag{7}$$

where Q' is

$$Q' = N(Q-1)/Q + 1. \tag{8}$$

There is the relationship of trade-off between the functional generation unit and the functional addition unit for fluctuation of the number of gates with the number of division, so that there is the number of division Q in Eq.(7) which is minimized the power dissipation[3]. The number is called optimum number of division. However, in higher-order FIR filters with complex coefficients, a large number of ROM's are used and each of them requires relatively high power dissipation. Consequently, this structure requires high power dissipation even if the division of function Φ is applied.

3. Optimum Function Circuit

In order to decrease drastically the power dissipation, we consider the structure based on the optimum function circuits using logic gate which we have previously proposed[2][3].

When the number of the address lines of ROM is m, the function Φ which is stored in a ROM, is defined as

$$W = \begin{bmatrix} \Phi(0, \ldots, 0) \\ \vdots \\ \Phi(1, \ldots, 1) \end{bmatrix} = \begin{bmatrix} w_0^{B-1} & \cdots & w_0^0 \\ \vdots & \cdots & \vdots \\ w_{2^m-1}^{B-1} & \cdots & w_{2^m-1}^0 \end{bmatrix} \tag{9}$$

where $w_i^{B-1} \cdots w_i^0$ is output of B-bit output of ROM and is stored at address i. The optimum function circuit is composed of logic gates and is based on the method which unifies the identical column and row vectors of Eq.(9). Using it, we can decrease the power dissipation drastically[2][3]. As the number of the identical column and row vectors in Eq.(9) increases, it has the advantage that the power dissipation becomes lower[2][3].

However, the optimum number of division increases in comparison with the conventional structure using ROM's[3]. Therefore, the power dissipation and latency increases in part of the functional addition unit. In order to realize a more effective method for further higher-order filters, we propose newly a implementation applying two methods which have large effects on the functional addition unit.

4. Structure using SFAs and SFSs

In order to realize lower power dissipation, we pay attention to linear-phase characteristic of complex FIR filters. In real FIR filters, it is enough to consider only real coefficients. However, the complex FIR filters have both real and imaginary coefficients. When the number of taps N is even, the coefficients of this filters have the following relationship

$$a_{R1} = a_{RN}, \ldots, a_{RN/2} = a_{RN/2+1}, \tag{10}$$

$$a_{I1} = -a_{IN}, \ldots, a_{IN/2} = -a_{IN/2+1}. \tag{11}$$

Substituting Eqs.(10) and (11) for Eq.(4), the real part of complex FIR filer can be rewritten as

$$y_R = -\Phi_{RI1}^0 + \sum_{k=1}^{B-1} 2^{-k}\Phi_{RI1}^k - \Phi_{RI2}^0 + \sum_{k=1}^{B-1} 2^{-k}\Phi_{RI2}^k \tag{12}$$
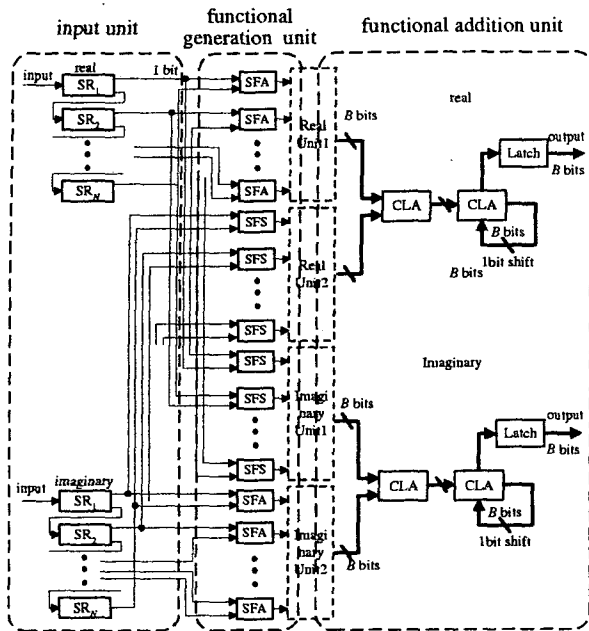
where Φ_{RI1}^k and Φ_{RI2}^k are

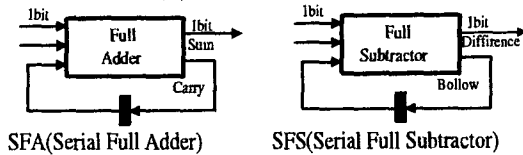$$\Phi_{RI1}^k = \sum_{i=1}^{N/2} a_{Ri}(v_{Ri}^k + v_{RN+1-i}^k), \tag{13}$$

$$\Phi_{RI2}^k = \sum_{i=1}^{N/2} a_{Ii}(v_{Ii}^k - v_{IN+1-i}^k). \tag{14}$$

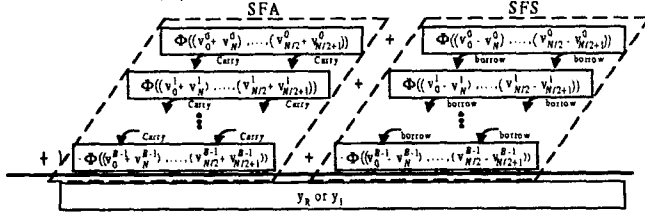Substituting Eqs.(10) and (11) for Eq.(5), the imaginary part of complex FIR filters can be rewritten as

$$y_I = -\Phi_{II1}^0 + \sum_{k=1}^{B-1} 2^{-k}\Phi_{II1}^k - \Phi_{II2}^0 + \sum_{k=1}^{B-1} 2^{-k}\Phi_{II2}^k \tag{15}$$

(a)Whole structure.

SFA(Serial Full Adder)   SFS(Serial Full Subtractor)

(b)Structure of SFA and SFS.

(c)The process of SFA and SFS.

Figure 1. Structure applying SFAs and SFSs.

where $\Phi_{\text{II}1}^k$ and $\Phi_{\text{II}2}^k$ are

$$\Phi_{\text{II}1}^k = \sum_{i=1}^{N/2} a_{\text{R}i}\left(v_{\text{R}i}^k - v_{\text{R}N+1-i}^k\right), \qquad (16)$$

$$\Phi_{\text{II}2}^k = \sum_{i=1}^{N/2} a_{\text{I}i}\left(v_{\text{I}i}^k + v_{\text{I}N+1-i}^k\right). \qquad (17)$$

If the parenthesized two input variables of equations from (12) to (17) can be calculated as 1-bit value, this can be regard as the arithmetic decreasing the number of terms by one-half. However, the addition and subtraction of these variables cause carry and borrow, respectively.

In order to overcome this problem, we propose a method applying SFAs and SFSs to the structure based on the optimum function circuits as shown in Fig.1. The carry and borrow which result from addition and subtraction are transmitted to register in SFA and in SFS as shown in Fig.1(b), respectively. Then each value is
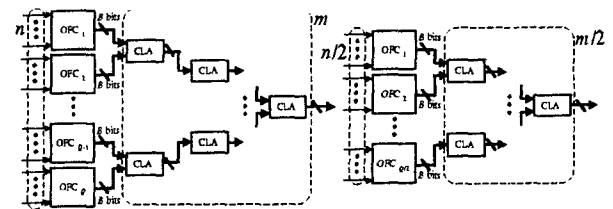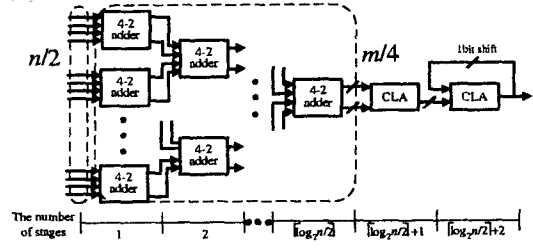

(a)Conventional structure   (b) Proposed structure
applying SFAs and SFSs

Figure 2. Structure of real and imaginary units shown in Fig. 1.(a)
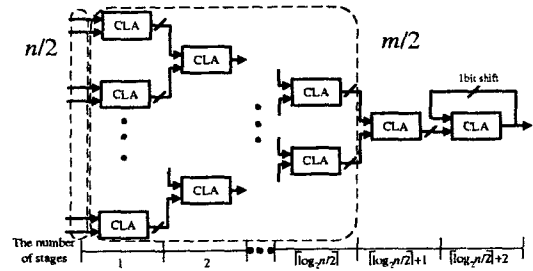
(a)Conventional Structure composed of CLAs.

(b)New structure of proposed 4-2 adders.

Figure 3. Structure of functional addition unit.

used as input variable at the time of the arithmetic of higher bit as shown in Fig.1(c). Here, the input variables must be the scaled range of $-0.5 \leq v_i < 0.5$ to prevent overflow.

Using SFAs and SFSs, we can executed the parenthesized variables as 1-bit value. In this way, the proposed structure decreases the number of input lines of the optimum function circuits by one-half as shown in Fig.2, so that the scale of function Φ can be reduced by one-half. Moreover, the number of the adders in the functional addition unit can be decreased by one-half. Therefore, it can be said that the proposed method is the way to make the most of the features of optimum function circuit, SFA, and SFS.

5. Structure using a proposed 4-2 adder

The conventional method based on the distributed arithmetic has used CLAs in functional addition unit. In order to realize not only lower power dissipation but also smaller latency, we consider a structure using 4-2 adders. The general 4-2 adder is basically the structure connecting two FAs(Full Adders) in cascade. If FA of the second stage can begin to process before that of the first stage finish processing, the speed becomes higher. In order to realize this process, we will take account

Table 1. VLSI evaluation of 60-tap complex FIR filters

	Proposed method	Conventional method using ROM's	General method using the multipliers for direct form
Power dissipation[W]	1.62	11.29	82.6
Area[mm^2]	7.06	16.9	177.2
Number of the gates	34591	82801	864480
Sampling rate[MHz]	2.65	2.65	6.71
Latency [ns]	486	567	8940

of the characteristics of two types of FAs. These FAs are defined as Type 1 FA and Type 2 FA, respectivey. Type 1 FA propagates a carry at high speed. In type 2 FA, one input variable allows a delay of the processing time of HA(Half Adder) after the other two variables transmit. Also, the number of gates is few than Type 1 FA. In consideration for the number of gates, we utilize these characteristics and propose a 4-2 adder which can reduced by the processing time of HA less from that of the general 4-2 adder.

We propose a structure applying new 4-2 adders as shown in Fig.3(b). The number of the 4-2 adders in the proposed structure can be decreased by one-half compared to that of CLAs in the conventional structure. Also, the number of the gates of the proposed 4-2 adder can be reduced as compared with the general 4-2 adder. Thus, the proposed structure becomes lower power dissipation. The proposed adder has another of advantage that the processing time is decreased significantly as compared with that of CLA. Moreover, the proposed structure is the same number of the additional stages as the conventional structure as shown in Fig.3. Therefore, the number of pipeline stages is decreased and the smaller latency can be realized.

6. VLSI Evaluation

In this section, the processors with 0.8 μm CMOS standard cell and the power dissipation voltage of 5.0V are designed and evaluated. The processors have linear-phase characteristic which is obtained by Retmez's algorithm[4]. These are low-pass filters, of which passband edge and stopband edge of normalized frequency are 0.12 and 0.15, respectively. The VLSI evaluation is shown in Table 1. We took account of the distributed arithmetic of which processing time depends on only word length. Using it, in higher-order of 60-tap for the complex FIR filters, the proposed processor can realize small latency of 486ns(about 1.29 sampling period) with high sampling rate of 2.65MHz. However, the only application of the distributed arithmetic caused problem of high power dissipation. In order to overcome this, we utilized the features of optimum function circuit, SFA, SFS, and proposed 4-2 adder. Consequently, the power dissipation of the proposed processor is only 1.62W, which implies that it has the capability of a sharp drop of 85.7% compared to the conventional processor using ROM's. Also, the latency can be decreased by 14.3%.

Moreover, the comparison with the general processor using the multipliers for direct form is made. This pro-

cessor can decrease drastically the power dissipation by 98.0%. Also, it can reduce sharply the latency by 94.6%. By taking account of the features of the proposed architecture, it is obvious that the proposed architecture is an effective way to realize low power dissipation and small latency while sampling rate is kept constant for further higher-order filters with complex coefficients.

7. Conclusions

We proposed a high-performance VLSI architecture for higher-order FIR filters with complex coefficients. Using the distributed arithmetic and optimum function circuit, we realized small latency and low power dissipation with high sampling. For the purpose of obtaining the more effective method for further higher-order filters, we proposed newly the implementation applying two methods. First, in order to realize lower power dissipation, we applied SFAs and SFSs to the structure based on the optimum function circuits. Second, in order to realize smaller latency with low power dissipation, we proposed the new structure applying proposed 4-2 adders. As a result, in higher-order of 60-tap, the proposed processor(with 0.8 μm CMOS standard cell) could realize small latency of 486ns (1.29 sampling period) with high sampling rate of 2.65MHz. The latency was much smaller than the conventional processor. Moreover, the power dissipation of the proposed processor was only 1.62W, and could be much lower.

For further higher order-filters, it was shown that the proposed architecture is an effective way to realize low power dissipation and small latency while the sampling rate is kept constant.

References

[1] C.F. Chen, "Implementing FIR Filters with Distributed Arithmetic," IEEE Trans. Acount. Speech & Signal Process., ASSP-34-4, pp.1318-1321, 1985.

[2] Y. Tsunekawa, T. Nozaki, and M. Miura, "High-Speed and Low Power Dissipation Architecture for Higher-Order FIR Filters with Very Small Latency", T.IEE Japan, Vol.118-C, No.7/8 pp.1098-1107 1998.

[3] T.Nozaki,Y.Tsunekawa, and N.Tayama, "High - Performance VLSI Architecture using Distributed Arithmetic for Higher-Order FIR Filters", ITC-CSCC2001, July 10-12, 2001.

[4] Karam,L.J., and J.H.McClellan, "Complex Chebyshev Approximation for FIR Filter Design.", IEEE Trans. on Circuits & System II.March, pp.207-216, 1995.