An Improved Systolic Architecture for 2-D Digital Filters

Naresh R. Shanbhag, *Student Member, IEEE*

***Abstract*—An improved systolic architecture for two-dimensional infinite-impulse-response (IIR) and finite-impulse-response (FIR) digital filters is presented. Comparisons with recently published work [2], [3] are made. When compared with the architecture in [2], a substantial reduction in the number of delay elements is observed. This reduction is of the order of 10^2 for a 2-D IIR filter and equals $N + 1$ for an Nth order 2-D FIR filter. The clock period has been made independent of the order of the filter. The speed-up factor is the maximum achievable and is independent of the filter order. Comparison with [3] shows an improvement in the latency of the systolic array, which has been reduced from 1 to 0. A reduction of $N + 1$ delay elements has been achieved for the FIR filter. An error analysis for the new architecture is made, while error expressions for the architectures in [2], [3] are also presented.**

I. Introduction

SYSTOLIC architectures, first developed by Kung [1], are characterized by a high degree of modularity, regularity, localized data communication, global clocking and increased speed of computation. Implementation of such architectures, through VLSI techniques, is facilitated by the repetitive nature of the processing elements (PE's).

Recently, systolic architectures for 2-D filtering have been proposed [2]–[4]. While the architecture in [2] is derived from the filter transfer function, the one in [4] is based on the local state space model. Though the realization in [2] has certain advantages over the one in [4], which include employment of simpler PE's and data input to a raster scan format, it must be mentioned that a large number of shift registers are required to implement the architecture. In most image processing applications, this feature may prevent a monolithic implementation of the filter. Therefore, it is of considerable interest to develop architectures which require fewer delay elements and at the same time have equal if not better performance. This objective has been achieved up to a certain extent by the architectures in [3]. We extend this line of work by presenting another improved systolic architecture for 2-D filtering. This architecture has been derived from the signal flow graph (SFG) representation of the filter, by the application of a systolizing procedure given in [5].

Manuscript received June 16, 1990; revised October 18, 1990.

The author is with the Department of Electrical Engineering, University of Minnesota, Minneapolis, MN 55455.

IEEE Log Number 9042726.

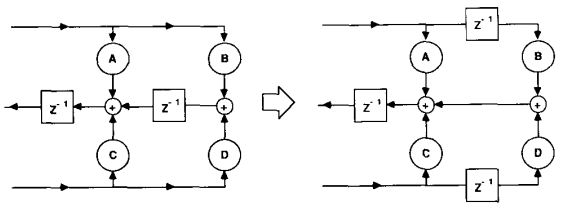


Fig. 1. The systolic transformation.

The paper is organized as follows. The new architecture for 2-D IIR and FIR filters is presented in Section II. In Section III, a comparison in terms of the number of adders, multipliers, registers and the clock period, with the realization in [2] and [3], is made. Error analysis of the new architecture is carried out and final error expressions for the architecture in [2] and [3] are presented in Section IV. The paper concludes with Section V.

II. Improved Systolic Architecture

It has been rigorously proved in [5] that any computable SFG can be systolized by rescaling the delays. A relevant case of this procedure, which shall henceforth be referred to as the systolic transformation (ST), is shown in Fig. 1. The systolic architecture, which was presented in [7] for a 1-D IIR filter, can also be derived through the ST. If the SFG which has to be systolized is canonical in the number of delays, then this approach would yield a very different realization in terms of the number of registers. We therefore first derive an SFG, which is canonical in the number of delays (canonical SFG) and then, employing the ST (Fig. 1), develop the systolic architecture.

A. 2-D IIR Filter

Unlike in [7], where the 1-D transfer function was taken to be strictly proper, we assume a more general transfer function. An Nth-order 2-D IIR filter transfer function is defined as

$$H(z_1, z_2) = \frac{\sum_{i=0}^{N}\sum_{j=0}^{N} a_{i,j} z_1^{-i} z_2^{-j}}{1 - \sum_{i=0}^{N}\sum_{j=0}^{N} b_{i,j} z_1^{-i} z_2^{-j}} \qquad (2.1)$$

1053-587X/91/0500-1195$01.00 © 1991 IEEE

where $b_{0,0} = 0$ and shall remain so for the rest of the treatment.

If $Y(z_1, z_2)$ and $X(z_1, z_2)$ represent the output and input data in the Z-domain, respectively, then

$$\begin{aligned} Y(z_1, z_2) &= X(z_1, z_2)\left[\sum_{i=0}^{N}\sum_{j=0}^{N} a_{i,j}z_1^{-i}z_2^{-j}\right] \\ &\quad + Y(z_1, z_2)\left[\sum_{i=0}^{N}\sum_{j=0}^{N} b_{i,j}z_1^{-i}z_2^{-j}\right] \\ &= \sum_{j=0}^{N} z_2^{-j}\sum_{i=0}^{N} a_{i,j}z_1^{-i}X(z_1, z_2) \\ &\quad + \sum_{j=0}^{N} z_2^{-j}\sum_{i=0}^{N} b_{i,j}z_1^{-i}Y(z_1, z_2) \\ &= \sum_{j=0}^{N} z_2^{-j}\left[\sum_{i=0}^{N} (a_{i,j}z_1^{-i}X(z_1, z_2) \right. \\ &\quad \left. + b_{i,j}z_1^{-i}Y(z_1, z_2))\right] \\ &= \sum_{j=0}^{N} z_2^{-j}\,[X(z_1, z_2)\,(a_{0,j} \\ &\quad + z_1^{-1}(a_{1,j} + z_1^{-1}(\cdots + z_1^{-1}(a_{N,j})\cdots))) \\ &\quad + Y(z_1, z_2)(b_{0,j} + z_1^{-1}(b_{1,j} \\ &\quad + z_1^{-1}(\cdots + z_1^{-1}(b_{N,j})\cdots)))]. \end{aligned} \tag{2.2}$$

From (2.2), the canonical SFG, for any 2-D IIR filter of order N, can be derived. In Fig. 2(a), we show the canonical SFG for $N = 2$. Next, we systolize the SFG in Fig. 2(a) by employing the ST (Fig. 1), to derive the SFG for a systolic architecture (Fig. 2(b)). It is clear that the ST introduces two delays for every other z_1^{-1} delay in the canonical SFG.

In order to map the systolic SFG (Fig. 2(b)) to an architecture, we assume that the sequence of input data is in raster scan format. In other words, the input data sequence is $x(0, 0)$, $x(0, 1)$, $\cdots$, $x(0, M-1)$, $x(1, 0)$, $x(1, 1)$, $\cdots$, etc. Therefore, the length of a row of input is M. The architecture for a second-order IIR filter is presented in Fig. 2(c), where the z_2^{-1} delays are replaced by shift registers of length M and the z_1^{-1} delays are single registers. It can be seen that two types of PE's (PE1 and PE2) are needed. To generate the architecture for higher order filters, all that is needed is to cascade PE1's in each subblock and add more subblocks in parallel. It can be easily confirmed that, if N' represents the number of PE1's in each subblock, then

$$N' = [N/2]. \tag{2.3}$$

B. 2-D FIR Filter

In a fashion similar to the IIR case, we can repeat the analysis for FIR filters. In fact, all that is required is to set all b_{ij} equal to zero and repeat the analysis presented in Section II(A). For the sake of brevity, we present the final systolic architecture for $N = 2$ (Fig. 3). Again, two types of PE's are needed (PE1 and PE2) and the architecture for higher order filters can be generated by cascading the PE1's in each subblock and adding the requisite number of subblocks in parallel.

III. Comparison with Existing Architectures

In this section we compare our architecture with those proposed in [2] and [3]. The architecture in [6] is an application of the systolic array, which was presented in [2], to complex digital filters. In [8], a hybrid of systolic and parallel architectures is presented, where it is assumed that the whole data array is available for processing. Notice that this is different from the raster scan input format assumed for our architecture. At present, the architecture in [8] is applicable to first and second-order filters with orthogonal symmetry and separable denominators. Therefore, it suffices to make comparisons with the work in [2] and [3].

The parameters that are compared are the number of adders, multipliers, registers, the clock period, the latency and the speed-up factor (SUF) [5]. The SUF measure, which is used to compare the speed efficiencies of systolic arrays, is defined below

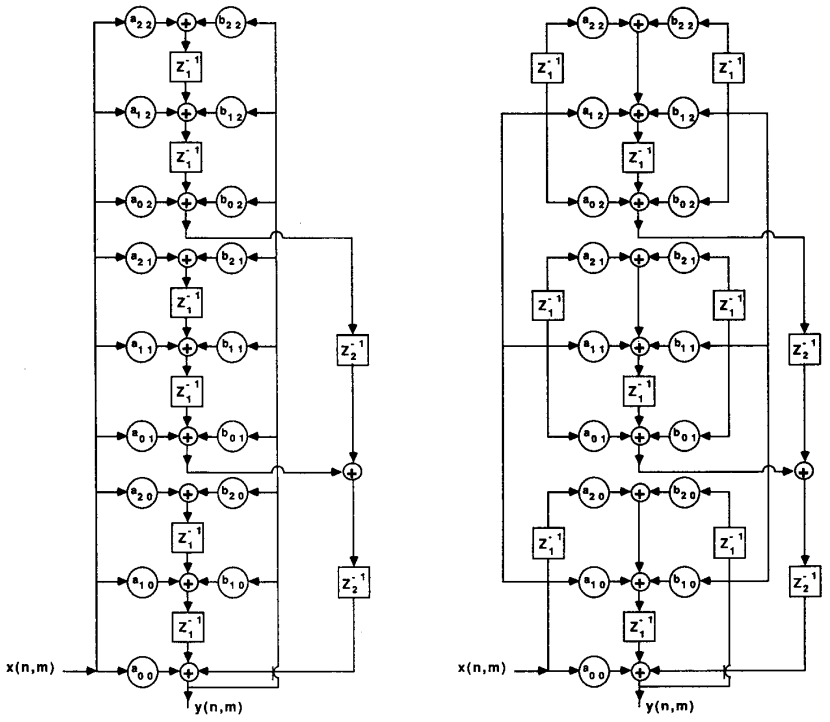
$$\text{SUF} = \frac{\text{Processing Time in a Single Processor}}{\text{Processing Time in the Array Processor}}. \tag{3.1}$$

For the purpose of comparison, we assume that all adders and multipliers are 2-operand and T_m and T_a are the times required to complete one real addition and multiplication, respectively.

A. 2-D IIR Filter Comparison

In [3], three different systolic architectures (Fig. 4) for 2-D IIR filtering are presented. These three architectures, which shall henceforth be referred to as SCH1 (Fig. 4(b)), SCH2 (Fig. 4(c)) and SCH3 (Fig. 4(d)) respectively, are all based on the same PE as [2] (Fig. 4(a)). It must be mentioned that SCH2 is identical to the architecture proposed in [2]. Comparison of our 2-D IIR architecture with SCH1, SCH2, and SCH3 is tabulated in Table I. It is clear that we have achieved a substantial reduction (of the order of MN) in the number of delay elements as compared to SCH2. This is due to the fact that M is usually of the order of 10^2. As compared to SCH3, the reduction in the number of delay elements equals N. On the other hand, SCH1 requires $N^2/2 - N/2 - 1$ fewer latches than ones. For most practical applications, the reduction achieved by SCH1 is of the order of 10. In fact, for $N = 2$ our architecture requires the same number of latches as SCH1. The reduction in the delay elements achieved by SCH1 is at the cost of increased latency. Along with SCH2 and SCH3, SCH1 has a latency of one, while our architecture has the minimum achievable latency of zero. This fact can be checked easily by observing that the first output ($y(0, 0)$) in our architecture, is available in the same clock cycle

pages. Payment of voluntary page charges (first eight only) for this IEEE TRANSACTIONS, like those journals of other professional societies, is not a necessary prerequisite for publication. However, payment of excess page charges is a prerequisite. The author will receive 100 free reprints (without covers) only if the voluntary page charge is honored. Detailed instructions will accompany the proofs.



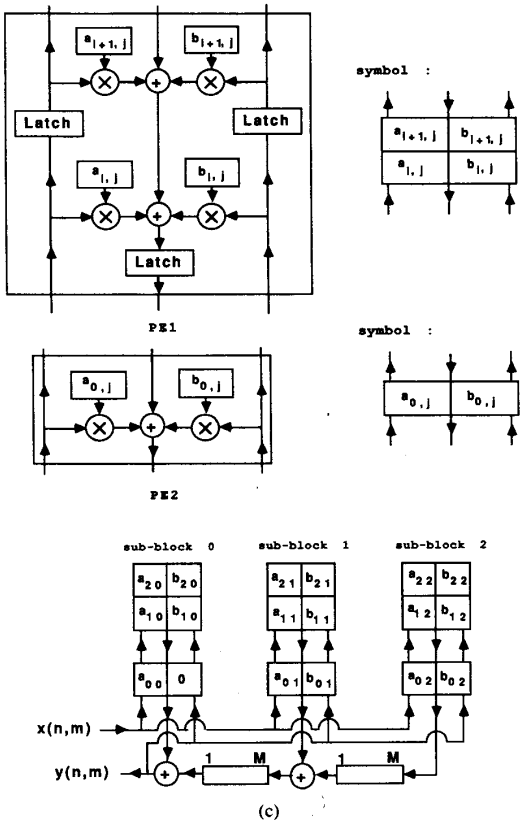


Fig. 2. Two-dimensional IIR filter: (a) canonical SFG, (b) systolized SFG, and (c) the systolic architecture.

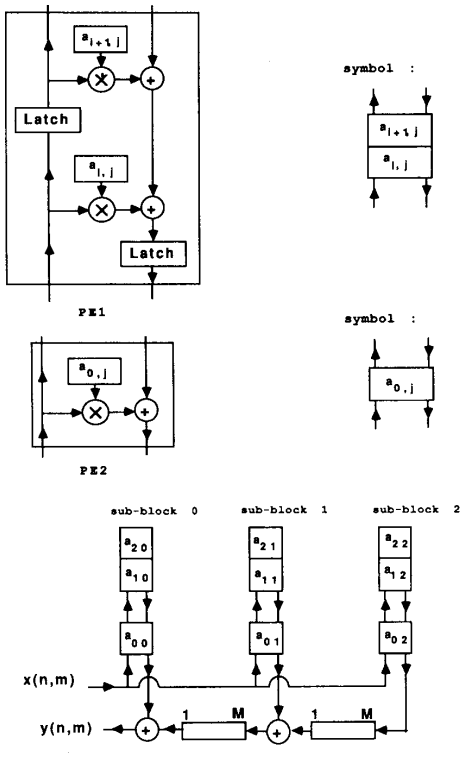


Fig. 3. The systolic architecture for 2-D FIR filter.

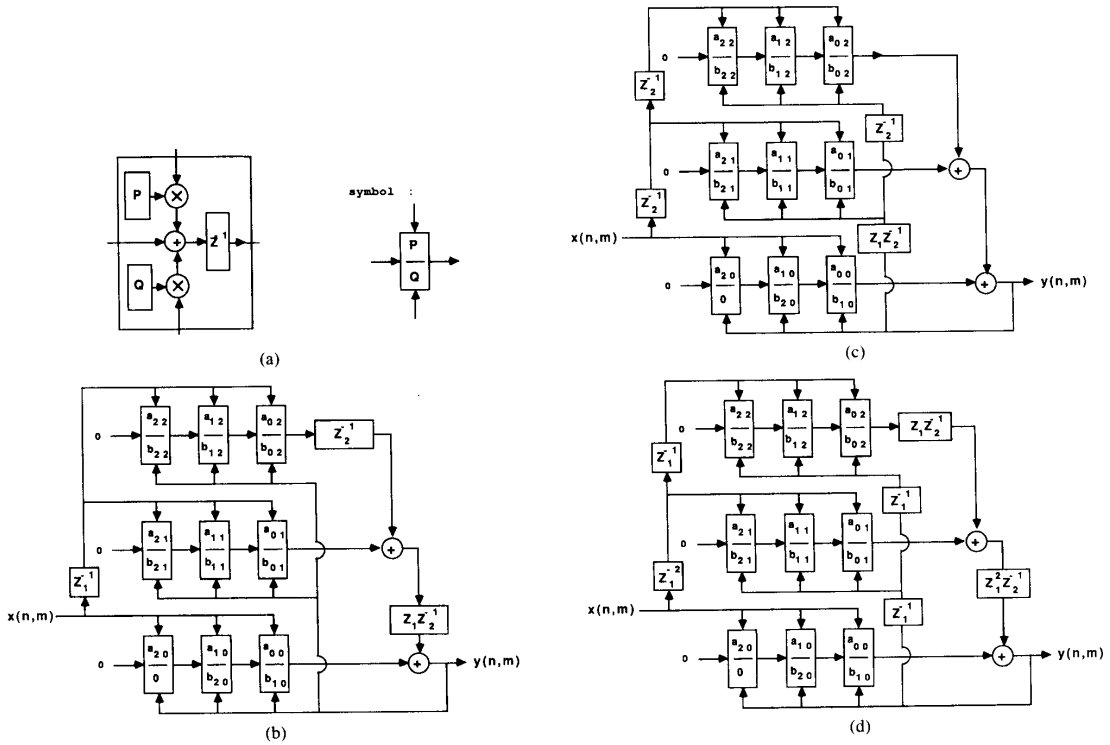


Fig. 4. Existing systolic architectures: (a) basic PE, (b) SCH1, (c) SCH2, and (d) SCH3.

pages. Payment of voluntary page charges (first eight only) for this IEEE TRANSACTIONS, like those journals of other professional societies, is not a necessary prerequisite for publication. However, payment of excess page charges is a prerequisite. The author will receive 100 free reprints (without covers) only if the voluntary page charge is honored. Detailed instructions will accompany the proofs.

TABLE I
COMPARISON WITH THE IIR ARCHITECTURE IN [2], [3]

Parameters	New	SCH1	SCH2	SCH3
No. of latches	$\frac{3N}{2}(N+1)+MN$	$(N+1)^2+MN$	$(N+1)^2+2MN$	$\left\lfloor\frac{3N}{2}+1\right\rfloor(N+1)+MN$
Cycle period	T_m+3T_a	T_m+2T_a	$\max\{(T_m+2T_a), T_a\lceil\log_2(N+1)\rceil\}$	T_m+2T_a
SUF	1.0	1.0	$\frac{T_m+2T_a}{\max\{(T_m+2T_a), T_a\lceil\log_2(N+1)\rceil\}}$	1.0
No. of adders	$2(N+1)^2-2$	$2(N+1)^2-2$	$2(N+1)^2-2$	$2(N+1)^2-2$
No of multipliers	$2(N+1)^2-1$	$2(N+1)^2-1$	$2(N+1)^2-1$	$2(N+1)^2-1$
Latency	zero	one	one	one

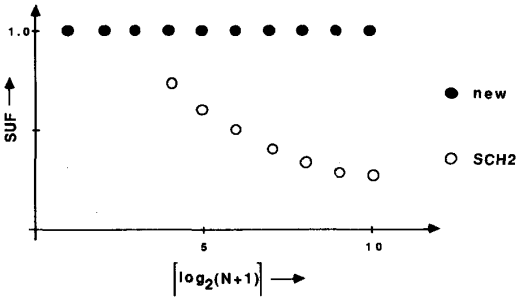


Fig. 5. Variation of the SUF measure for SCH2.

in which the first input ($x(0, 0)$) is made available to the circuit.

The clock period for our architecture is marginally longer (by T_a) than that of SCH1 and SCH3, while it is clearly shorter than that of SCH2. This disadvantage is more than made up for by the improvement in the latency. The rest of the comparison parameters, i.e., the SUF measure, number of adders and multipliers, are identical for all the architectures under consideration except SCH2. Unlike SCH1, SCH3, and the new architecture, where the SUF measure is equal to 1, the SUF measure for SCH2 deteriorates for increasing filter orders. If we assume that $T_m = T_a$, then the cycle period for SCH2 is an increasing function of N for $N > 7$. The SUF measure (Fig. 5) for SCH2 keeps decreasing for filter orders higher than 7.

B. 2-D FIR Filter Comparison

Though 2-D FIR architectures are not presented in [3], they were derived, from the corresponding IIR filter architectures, by equating the b_{ij} coefficients to zero. Let SCH1′, SCH2′, and SCH3′ be the FIR architectures derived from SCH1, SCH2, and SCH3, respectively. Again, SCH2′ is identical to the 2-D FIR architecture presented in [2].

Comparison of our architecture with SCH1′, SCH2′, and SCH3′ was done, the results of which are tabulated in Table II. It can be seen that the new architecture requires the least number of delay elements. Specifically, it requires $N + 1$ fewer registers than SCH1′ and SCH2′. Compared with SCH3′, our architecture requires $(N^2/2 + 3N/2 - 2)$ fewer registers. The rest of the factors compare in a fashion similar to the IIR case. In Fig. 6, we show the variation of the SUF measure with the filter order N, under the assumption of $T_m = T_a$. This time the SUF measure for the SCH2′ deteriorates for $N > 3$. Similar to the IIR case, the latency of our architecture is the minimum achievable.

IV. Error Analysis

It is well known that finite-precision arithmetic results in quantization errors. Therefore, it is essential to have an estimate of the errors involved. We present, in this section, a detailed error analysis of our architecture for the 2-D IIR case. Final error expressions for SCH1, SCH2, and SCH3 are also presented. Though the error expressions for FIR filters are not calculated, it is clear that these can easily be derived from the corresponding expressions for IIR filters. It must be mentioned that this analysis is a direct extension of the error analysis done in [6] for 1-D IIR filters.

For the purpose of error analysis, it would be convenient to consider the aggregation of subblocks in Fig. 2(c) as a two dimensional array of PE's, with PE's in each subblock forming a column. Let $PE_{i,j}$ denote the processing element in the jth column and the ith row, where $i = 0$ is the row containing PE2's and $0 \le i \le N'$. If X represents the true value of a variable then its quantized value would be represented by $\overline{X}$. Let $\alpha_{i,j}$ and $\beta_{i,j}$ represent the coefficient quantization errors in the representation of $a_{i,j}$ and $b_{i,j}$, respectively. Also, let $e_X(n, m)$ and $e_Y(n, m)$ de-

TABLE II
COMPARISON WITH THE FIR ARCHITECTURE IN [2], [3]

Parameters	New	SCH1′	SCH2′	SCH3′
No. of latches	$N(N+1)+MN$	$(N+1)^2+MN$	$(N+1)^2+MN$	$\left\lfloor \frac{3N}{2}+1 \right\rfloor (N+1)+MN-3$
Cycle period	T_m+2T_a	T_m+T_a	$\max\{(T_m+T_a), T_a\lceil\log_2(N+1)\rceil\}$	T_m+T_a
SUF	1.0	1.0	$\frac{T_m+T_a}{\max\{(T_m+T_a), T_a\lceil\log_2(N+1)\rceil\}}$	1.0
No. of adders	$N(N+2)$	$N(N+2)$	$N(N+2)$	$N(N+2)$
No of multipliers	N^2	N^2	N^2	N^2
Latency	zero	one	one	one

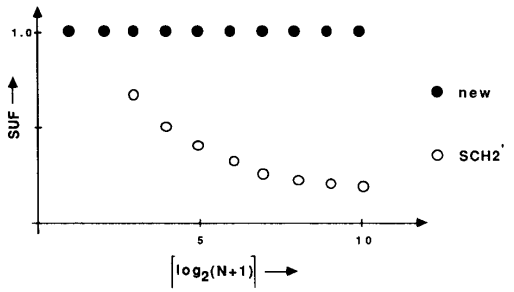


Fig. 6. Variation of SUF measure for SCH2′.

note the errors in the representation of input $x(n, m)$ and the final output $y(n, m)$, respectively.

We first derive the error expression at the output of $PE_{i,j}$ for $1 \le i \le N'$. In other words, we first consider PE's for type PE1. Let $y_{i,j}(n, m)$ denote the true output of $PE_{i,j}$ and let $i' = 2i - 1$ for $1 \le i \le N'$, then

$$\begin{aligned} y_{i,j}(n, m) &= a_{i'+1,j}x(n - i' - 1, m) + a_{i',j}x(n - i', m) \\ &\quad + b_{i'+1,j}y(n - i' - 1, m) \\ &\quad + b_{i',j}y(n - i', m) + y_{i+1,j}(n - 1, m). \end{aligned} \tag{4.1}$$

Therefore, the quantized value of $y_{i,j}(n, m)$ is given by

$$\begin{aligned} \bar{y}_{i,j}(n, m) &= [\bar{a}_{i'+1,j}\bar{x}(n - i' - 1, m)]_q \\ &\quad + [\bar{a}_{i',j}\bar{x}(n - i', m)]_q \\ &\quad + [\bar{b}_{i'+1,j}\bar{y}(n - i' - 1, m)]_q \\ &\quad + [\bar{b}_{i',j}\bar{y}(n - i', m)]_q \\ &\quad + \bar{y}_{i+1,j}(n - 1, m) - s_{i,j} \end{aligned} \tag{4.2}$$

where $s_{i,j}$, representing the storage error, is defined as the error caused by storing the output of an adder in a latch. In [6], it has been discussed that for fixed point data representation with a dynamic range between $1/2$ and 1, rounding with $(t+1)$ bit registers results in $s_{i,j} \le 2^{-t-1}$. Representing the error at the output of $PE_{i,j}$ by $f_{i,j}(n, m)$ and neglecting the second-order terms of the type $\alpha_{i',j}e_X(n - i', m)$, we get

$$\begin{aligned} f_{i,j}(n, m) &= y_{i,j}(n, m) - \bar{y}_{i,j}(n, m) \\ &= P_{i'+1,j} + P_{i',j} + C_{i'+1,j} + C_{i',j} \\ &\quad + f_{i+1,j}(n - 1, m) \\ &\quad + s_{i,j} + a_{i'+1,j}e_X(n - i' - 1, m) \\ &\quad + a_{i',j}e_X(n - i', m) \\ &\quad + b_{i'+1,j}e_Y(n - i' - 1, m) \\ &\quad + b_{i',j}e_Y(n - i', m) \\ &\quad + \alpha_{i'+1,j}x(n - i' - 1, m) \\ &\quad + \alpha_{i',j}x(n - i', m) \\ &\quad + \beta_{i'+1,j}y(n - i' - 1, m) \\ &\quad + \beta_{i',j}y(n - i', m) \end{aligned} \tag{4.3}$$

where $P_{i,j}$ and $C_{i,j}$ are the multiplication roundoff errors and are defined as

$$\begin{aligned} P_{i,j} &= \bar{a}_{i,j}\bar{x} - [\bar{a}_{i,j}\bar{x}]_q \\ C_{i,j} &= \bar{b}_{i,j}\bar{y} - [\bar{b}_{i,j}y]_q. \end{aligned} \tag{4.4}$$

Recall that (4.3) is applicable for $1 \le i \le N'$, where N' is defined in (2.3). For $i = 0$, i.e., for PE's of type PE2, the combined error at the output of $PE_{0,j}$ can be derived in a similar fashion and is stated below

$$\begin{aligned} f_{0,j}(n, m) &= P_{0,j} + C_{0,j} + f_{1,j}(n - 1, m) \\ &\quad + \alpha_{0,j}x(n, m) \\ &\quad + \beta_{0,j}y(n - i, m) \\ &\quad + \alpha_{0,j}e_X(n, m) + b_{0,j}e_Y(n - 1, m). \end{aligned} \tag{4.5}$$

Solving recursively for $f_{0,j}(n, m)$, we get

$$\begin{aligned} f_{0,j}(n, m) &= \sum_{i=0}^{N} (P_{i,j} + c_{i,j} + \alpha_{i,j}x(n - i, m) \\ &\quad + \beta_{i,j}y(n - i, m) + a_{i,j}e_X(n - i, m) \\ &\quad + b_{i,j}e_Y(n - i, m)) + \sum_{i=0}^{N'} s_{i,j}. \end{aligned} \tag{4.6}$$

pages. Payment of voluntary page charges (first eight only) for this IEEE TRANSACTIONS, like those journals of other professional societies, is not a necessary prerequisite for publication. However, payment of excess page charges is a prerequisite. The author will receive 100 free reprints (without covers) only if the voluntary page charge is honored. Detailed instructions will accompany the proofs.

Summing (4.6) over all j, we get the combined error at the final output as follows:

$$f(n, m) = \sum_{j=0}^{N} f_{0,j}(n, m - j)$$
$$= m_E + c_E + I_E + s_E \tag{4.7}$$

where m_E (multiplication roundoff errors), c_E (coefficient quantization error), I_E (input quantization error) and s_E (storage error) are defined as

$$m_E = \sum_{j=0}^{N} \sum_{i=0}^{N} (P_{i,j} + C_{i,j})$$
$$c_E = \sum_{j=0}^{N} \sum_{i=0}^{N} (\alpha_{i,j} x(n - i, m - j) + \beta_{i,j} y(n - i, m - j))$$
$$I_E = \sum_{j=0}^{N} \sum_{i=0}^{N} (a_{i,j} e_X(n - i, m - j) + b_{i,j} e_Y(n - i, m - j))$$
$$s_E = \sum_{j=0}^{N} \sum_{i=0}^{N'} s_{i,j} + \sum_{i=1}^{N} s_i \tag{4.8}$$

where $s_1, s_2, \cdots, s_N$, which occur in the expression for s_E, are the additional storage errors due to the adders which add the outputs of the $PE_{0,j}$'s. Therefore, (4.7) and (4.8) represent the final expressions for the combined error at the filter output.

Error analysis, similar to the one presented above, was done on SCH1, SCH2, and SCH3. The final error expressions for SCH1 (4.9), SCH2 (4.10), and SCH3 (4.11) are presented below:

$$f(n, m) = \sum_{j=0}^{N} \sum_{i=0}^{N} (P_{i,j} + C_{i,j}) + \sum_{j=0}^{N} \sum_{i=0}^{N} (\alpha_{i,j} x(n - 1, m - j) + \beta_{i,j} y(n - 1, m - j)) + \sum_{i=0}^{N} (\alpha_{i,0} x(n - 1, m) + \beta_{i+1,0} y(n - 1, m)) + \sum_{j=0}^{N} \sum_{i=0}^{N} (a_{i,j} e_X(n - 1, m - j) + b_{i,j} e_Y(n - 1, m - j)) + \sum_{i=0}^{N} (a_{i,0} e_X(n - 1, m) + b_{i+1,0} e_Y(n - 1, m)) + \sum_{j=0}^{N} \sum_{i=0}^{N} s_{i,j} + \sum_{i=1}^{N-1} \tag{4.9}$$

$$f(n, m) = \sum_{j=0}^{N} \sum_{i=0}^{N} (P_{i,j} + C_{i,j}) + \sum_{j=1}^{N} \sum_{i=0}^{N} (\alpha_{i,j} x(n - 1, m - j) + \beta_{i,j} y(n, m - j)) + \sum_{i=0}^{N} (\alpha_{i,0} x(n - 1, m) + \beta_{i+1,0} y(n - 1, m)) + \sum_{j=0}^{N} \sum_{i=0}^{N} (a_{i,j} e_X(n - 1, m - j) + b_{i,j} e_Y(n, m - j)) + \sum_{i=0}^{N} (a_{i,0} e_X(n - 1, m) + b_{i+1,0} e_Y(n - 1, m)) \sum_{j=0}^{N} \sum_{i=0}^{N} s_{i,j} + s_1 \tag{4.10}$$

and finally

$$f(n, m) = \sum_{j=0}^{N} \sum_{i=0}^{N} (P_{i,j} + C_{i,j}) + \sum_{j=0}^{N} \sum_{i=0}^{N} (\alpha_{i,j} x(n - 1, m - j) + \beta_{i,j} y(n - j - 1, m - j)) + \sum_{i=0}^{N} (\alpha_{i,0} x(n - 1, m) + \beta_{i+1,0} y(n - 1, m)) + \sum_{j=1}^{N} \sum_{i=0}^{N} (a_{i,j} (a_{i,j} e_X(n - 1, m - j) + b_{i,j} e_Y(n - j - 1, m - j)) + \sum_{i=0}^{N} (a_{i,0} e_X(n - 1, m) + b_{i+1,0} e_Y(n - 1, m)) + \sum_{j=0}^{N} \sum_{i=0}^{N} s_{i,j} + \sum_{i=1}^{N} s_1 \tag{4.11}$$

where $\beta_{N+1,0}$ and $b_{N+1,0}$ in (4.9)–(4.11) are equal to zero and s_i's are the storage errors due to the adders not belonging to any of the PE's.

If it is assumed that $s_{i,j}$, for all the architectures under consideration, are all of a similar nature then it is apparent that our architecture has the lowest storage error. This is due to the fact that our architecture has fewer PE's and thus fewer adder results are stored. By equating all $b_{i,j}$'s, $\beta_{i,j}$'s, and $C_{i,j}$'s to zero, error expressions for the corresponding FIR architectures can also be derived.

V. Conclusions

An improved systolic architecture for 2-D IIR and FIR filters is presented. These, have been derived by applying a systolizing procedure [5] to the canonical SFG's of the

corresponding filters. Comparisons with most recent existing architectures [2], [3], have been made. As compared to [2], we have achieved a reduction in the number of registers of the order of 10^2, for IIR filters. For FIR filters, this reduction equals $N + 1$. The SUF measure and the clock period have been made independent of the filter order, for both IIR and FIR filters. The SUF measure, for our architecture, is also the maximum achievable. In [3], three different systolic architectures (SCH1, SCH2, and SCH3) have been presented, for IIR filtering. For comparison purposes, FIR architectures (SCH1′, SCH2′, and SCH3′) have been derived from the corresponding IIR architectures. All three IIR architectures have the same PE's as [2] and have a latency of 1. The latency of our architecture is zero and thus is an improvement over [3]. For IIR filters SCH1, which is the most efficient of the three, requires the fewest registers. On the other hand, for FIR filters, our architecture has $N + 1$ fewer registers than SCH1′. A detailed error analysis, done along the same lines as in [6], is presented for our architecture, while the final error expressions for SCH1, SCH2, and SCH3 are stated. The error expressions include the multiplication roundoff error (m_E), the coefficient quantization error (c_E), the input quantization error (I_E), and the storage error (s_E). Under the assumption that the storage error per PE is the same for all the architectures, our architecture has the least storage error.

It would interesting to develop filter structures which are a hybrid of systolic and parallel architectures. This has been done to some extent in [8], where a systolic-cum-parallel architecture for 2-D IIR filters with transfer functions, which have orthogonal symmetry and separable denominators, has been presented. Extension to generalized filter transfer functions and for orders higher than two is a topic of future research.

ACKNOWLEDGMENT

The author wishes to acknowledge the perennial encouragement offered by Prof. D. Nagchoudhuri. The suggestions of the reviewers are greatly appreciated.

REFERENCES

[1] H. T. Kung, "Why systolic architectures," *Comput. Mag.*, vol. 15, pp. 37-46, Jan. 1982.
[2] M. A. Sid-Ahmed, "A systolic realization for 2-D digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, pp. 560-565, Apr. 1989.
[3] S. Sunder, F. El-Guibaly, and A. Antoniou, "Systolic implementations of two-dimensional recursive digital filters," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 1990, pp. 1034-1037.
[4] G. A. Lampropoulos and M. M. Fahmy, "A new realization for 2-D digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 533-542, Apr. 1987.
[5] S-Y. Kung, "On supercomputing with systolic/wavefront array processors," *Proc. IEEE*, vol. 72, pp. 867-884, July 1984.
[6] D. Raghuramireddy and R. Unbehauen, "A systolic structure for complex digital filters," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 1990, pp. 1252-1255.
[7] Jayadeva, "A new systolic design for digital IIR filters," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 653-654, May 1990.
[8] H. K. Kwan, "Systolic and parallel realizations of 2D IIR filters," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 1990, pp. 2345-2348.

Naresh R. Shanbhag (S'86) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, New Delhi, India, in 1988, and the M.S. degree in electrical engineering from Wright State University, Dayton, OH, in 1990. Currently he is working towards the Ph.D. degree in electrical engineering at the University of Minnesota at Minneapolis. His research interests include VLSI architectures and algorithms for signal processing, multiple-valued logic, and robust control.

pages. Payment of voluntary page charges (first eight only) for this IEEE TRANSACTIONS, like those journals of other professional societies, is not a necessary prerequisite for publication. However, payment of excess page charges is a prerequisite. The author will receive 100 free reprints (without covers) only if the voluntary page charge is honored. Detailed instructions will accompany the proofs.