A New 2-D Systolic Digital Filter Architecture Without Global Broadcast

Lan-Da Van, *Member, IEEE*

***Abstract*—In this paper, we propose two-dimensional (2-D) systolic-array infinite-impulse response (IIR) and finite-impulse response (FIR) digital filter architectures without global broadcast, by the hybrid of a modified reordering scheme and a new systolic transformation. This architecture has local broadcast, lower-quantization error, and zero latency without sacrificing the number of multipliers, as well as delay elements under the satisfactory critical period. Furthermore, we extend this new architecture to a useful 2-D systolic cascade-form architecture and provide the comprehensive error analysis for the proposed architectures.**

***Index Terms*—Cascade-form architecture, local broadcast, quantization error, systolic architecture, two–dimensional (2-D) digital filter.**

I. Introduction

RECENTLY, two-dimentional (2-D) digital filters have been widely researched in a variety of digital signal processing (DSP) applications, such as image restoration [1], [2] obtained through a 2-D low-pass intraframe filter and image enhancement [2], [3] performed by a 2-D high-pass filter. Generally speaking, the category of 2-D digital filters can be divided into infinite-impulse response (IIR) and finite-impulse response (FIR) digital filters. By contrast, an IIR digital filter has the advantage of highly computational efficiency and low-hardware cost, and an FIR digital filter has the merit of stable and linear-phase properties. Although, 2-D digital filters can be simulated on a general-purpose computer, it seems unlikely to process input signals in real time due to a large amount of computation. Therefore, an application-specific integrated circuit (ASIC) design plays an important role for the realization of 2-D digital filters. One of the most attractive structures for an ASIC design is the systolic array architecture [4]–[9] since it is characterized by synchronization, a high degree of modularity and regularity, local broadcast/communication, concurrency, and extendibility. When a large number of processing elements (PEs) work together, data communication becomes more significant. In recently developing VLSI technologies, the global broadcast makes a significant impact upon speed in the circuit level [5], [6] for an ASIC design. Therefore, the local broadcast of systolic arrays is advantageous to reducing the global broadcast impact [5]–[9].

Several systolic architectures for 2-D digital filters have existed in [10]–[12]. However, we obviously observe that input and output signals globally broadcast among the existing structures. On the other hand, the research [10]–[12] lacks a generalized discussion on further reducing storage error. As a result, we are motivated to design a local broadcast and lower storage-error architecture. In this paper, we propose an improved 2-D systolic architecture [13], [14] by differently reordering delays as well as summations of the filter-output function and then by a new systolic transformation. The former scheme eliminates the ith-dimensional global broadcast and the latter cancels the jth-dimensional global broadcast and reduces storage error, where i and j are dimensional indices of 2-D digital filters. The hybrid of two schemes can achieve higher throughput than that in [10], [12]. Thus, the resulting IIR and FIR architectures have local broadcast, lower quantization error, and zero latency under the acceptable throughput rate. The structure of this paper is organized as follows. The new 2-D systolic digital filter architectures without global broadcast are proposed in Section II. The quantization error analysis of the proposed architectures is discussed in Section III. In Section IV, comparison results are tabulated in terms of local broadcast, storage error, critical period, latency, and the number of multipliers as well as delay elements. In the last section, concise statements conclude this presentation.

II. An Improved Systolic Architecture Design

The general transfer function of a 2-D IIR digital filter can be represented as

$$H(z_1, z_2) = \frac{\sum_{i=0}^{N_1}\sum_{j=0}^{N_2} a_{i,j} z_1^{-i} z_2^{-j}}{1 - \sum_{i=0}^{N_1}\sum_{j=0}^{N_2} b_{i,j} z_1^{-i} z_2^{-j}} \tag{1}$$

where $b_{0,0} = 0$, $a_{i,j}$ as well as $b_{i,j}$ and $N_1 \times N_2$ are coefficients and the order [15] of the IIR digital filter, respectively. In this paper, a square image $M \times M$ is fed to the following structures in raster-scan mode and thus the delay $z_2^{-1} = z^{-1}$ and $z_1^{-1} = z^{-M}$ where z^{-1} and M denote a unit-delay element and the width of an image, respectively.

A. 2-D Systolic Noncascade Form Digital Filter

Without loss of generality, we review and deduce the following structures under an assumption $N_1 = N_2 = N$. Equation (1) can be modified as seen in (2), at the bottom of the next page, where $X = X(z_1, z_2)$ and $Y = Y(z_1, z_2)$ are defined as input and output of the digital filter, respectively, in the z transform domain. Equation (2), referred to as SCH2 of S-G-A [11] can be mapped to Sid-Ahmed's structure [10], [11] as shown in Fig. 1, where $N = 2$. In Fig. 1, the left index $(M - 1)$ of the

Manuscript received December 1, 2000; revised October 29, 2001.

The author was with the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C. He is now with the Chip Implementation Center (CIC), National Science Council, Hsinchu 300, Taiwan, R.O.C. (e-mail: ldvan@cic.edu.tw).

Digital Object Identifier 10.1109/TVLSI.2002.800531

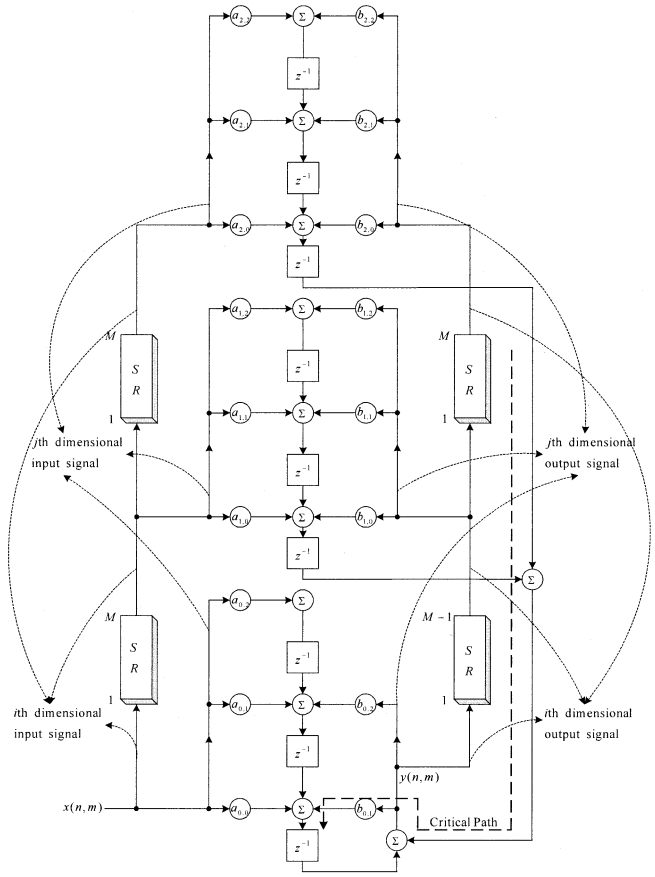

Fig. 1. An IIR digital filter proposed by [10] for $N = 2$.

shift register (SR) denotes the number of unit-delay elements and the input $x(n,m)$ as well as the output $y(n,m)$ are taken the inverse z transform of X and Y, respectively. Other 2-D IIR digital-filter structures can be realized either by reordering delays, as well as summations [11], or the systolic transformation [12]. The mapping equations for S-G-As SHC3 [11] and Shanbhag's scheme [12] are described, respectively, as seen in (3) and (4), at the bottom of the next page. The resulting structures corresponding to (3) and (4) are shown in Figs. 2 and 3, respectively, where $N = 2$. However, these structures [10]–[12] do not have local data broadcast for input and output signals.

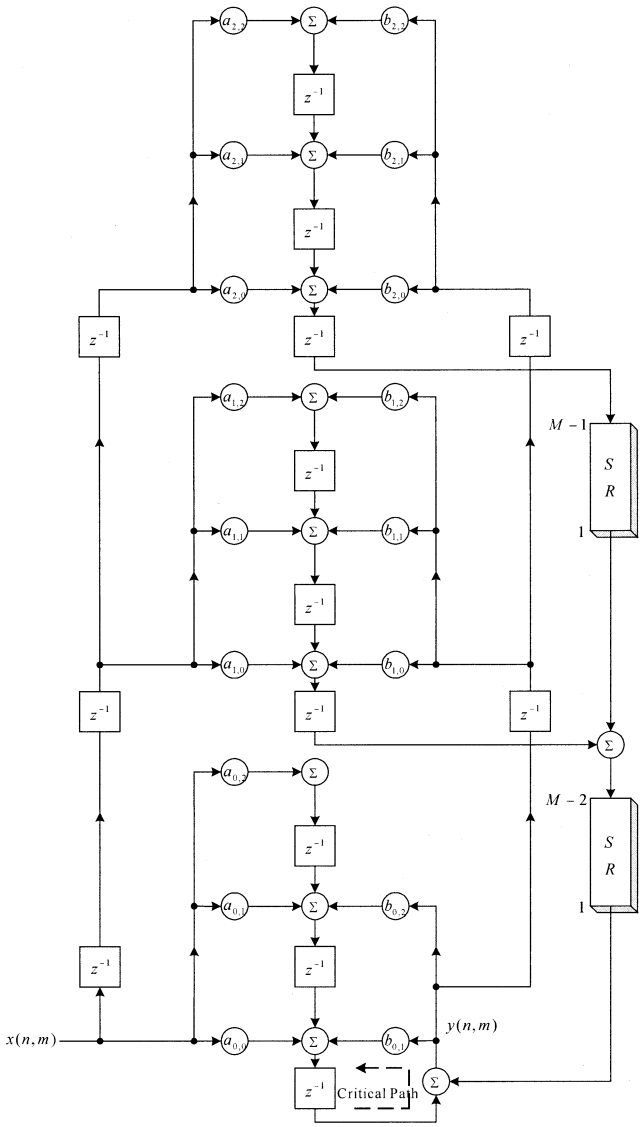

Fig. 2. An IIR digital filter proposed by [11] for $N = 2$.

For convenience of recognizing broadcast directions of input and output signal paths for 2-D digital filters, we divide them into ith-dimensional input signal, ith-dimensional output signal,

$$
\begin{aligned}
Y = &\left[\left(\sum_{j=0}^{N} a_{0,j} z_2^{-j-1}\right)(X) + \left(\sum_{j=0}^{N} b_{0,j} z_2^{-j}\right)(Y)\right] \\
&+ \left[\left(\sum_{j=0}^{N} a_{1,j} z_2^{-j-1}\right)\left(z_1^{-1} X\right) + \left(\sum_{j=0}^{N} b_{1,j} z_2^{-j-1}\right)\left(z_1^{-1} z_2 Y\right)\right] \\
&+ \left[\left(\sum_{j=0}^{N} a_{2,j} z_2^{-j-1}\right)\left(z_1^{-2} X\right) + \left(\sum_{j=0}^{N} b_{2,j} z_2^{-j-1}\right)\left(z_1^{-2} z_2 Y\right)\right] + \cdots \\
&+ \left[\left(\sum_{j=0}^{N} a_{N,j} z_2^{-j-1}\right)\left(z_1^{-N} X\right) + \left(\sum_{j=0}^{N} b_{N,j} z_2^{-j-1}\right)\left(z_1^{-N} z_2 Y\right)\right]
\end{aligned} \tag{2}
$$

jth-dimensional input signal and jth-dimensional output signal as indicated in Fig. 1. It is obvious that the above defined signal notations as shown in Fig. 1 can also be utilized in Figs. 2, 3 and the proposed architectures. Since the global broadcast leads to the low speed operation in the circuit level [6], we propose the hybrid of a modified recording scheme and a new systolic transformation to achieve local broadcast. The procedures are as follows.

First, we eliminate the ith dimensional global broadcast of input and output signals as shown in Fig. 3 by differently reordering delays and summations of the filter output function. The reordering is as shown in (5), at the bottom of the next page, where the integer variable P is restricted in the range of 1 and $M-1$ so as to maintain local broadcast in the ith-dimensional path. For simplifying the representation of (5), we define two terms as

$$F(i)=\sum_{j=0}^{N} a_{i,j} z_2^{-j}, \quad \text{for } i=0,1,\ldots,N \tag{6a}$$

$$G(i)=\sum_{j=0}^{N} b_{i,j} z_2^{-j}, \quad \text{for } i=0,1,\ldots,N. \tag{6b}$$

Substituting (6a) and (6b) into (5), we can rewrite (5) as

$$\begin{aligned} Y =&[F(0)X+G(0)Y] \\ &+\sum_{i=1}^{N} z_1^{-i} z_2^{i\times P}\left[F(i)\left(z_2^{-i\times P}X\right)+G(i)\left(z_2^{-i\times P}Y\right)\right] \\ =&[F(0)X+G(0)Y]+z_1^{-1}z_2^{P} \\ &\cdot\left(\left[F(1)\hat{X}_1+G(1)\hat{Y}_1\right]+z_1^{-1}z_2^{P}\right. \\ &\cdot\left(\left[F(2)\hat{X}_2+G(2)\hat{Y}_2\right]+\cdots\right. \\ &\left.\left.+z_1^{-1}z_2^{P}\left(\left[F(N)\hat{X}_N+G(N)\hat{Y}_N\right]\right)\cdots\right)\right) \end{aligned} \tag{7}$$

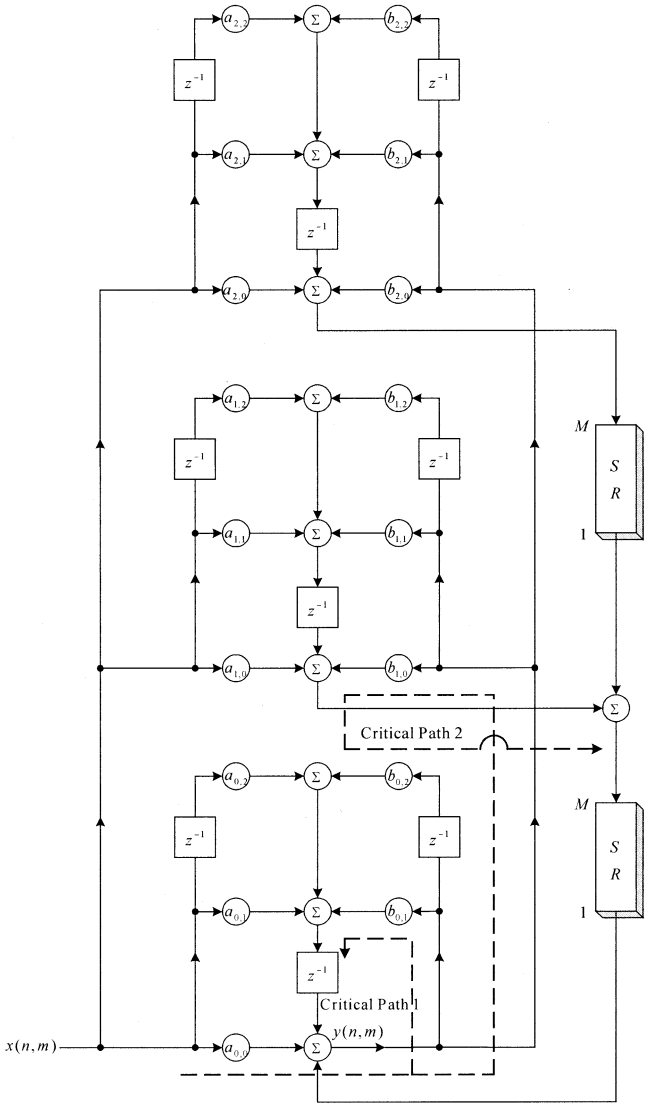

Fig. 3. An IIR digital filter proposed by [12] for $N=2$.

$$\begin{aligned} Y=&\left[\left(\sum_{j=0}^{N} a_{0,j}z_2^{-j-1}\right)(X)+\left(\sum_{j=0}^{N} b_{0,j}z_2^{-j}\right)(Y)\right] \\ &+z_1^{-1}z_2^{2}\left(\left[\left(\sum_{j=0}^{N} a_{1,j}z_2^{-j-1}\right)\left(z_2^{-2}X\right)+\left(\sum_{j=0}^{N} b_{1,j}z_2^{-j-1}\right)\left(z_2^{-1}Y\right)\right]\right. \\ &\quad+z_1^{-1}z_2^{1}\left(\left[\left(\sum_{j=0}^{N} a_{2,j}z_2^{-j-1}\right)\left(z_2^{-3}X\right)+\left(\sum_{j=0}^{N} b_{2,j}z_2^{-j-1}\right)\left(z_2^{-2}Y\right)\right]+\cdots\right. \\ &\quad\left.\left.+\ z_1^{-1}z_2^{1}\left(\left[\left(\sum_{j=0}^{N} a_{N,j}z_2^{-j-1}\right)\left(z_2^{-N-1}X\right)+\left(\sum_{j=0}^{N} b_{N,j}z_2^{-j-1}\right)\left(z_2^{-N}Y\right)\right]\right)\cdots\right)\right), \end{aligned} \tag{3}$$

and

$$\begin{aligned} Y=\sum_{i=0}^{N} z_1^{-i}&\left[X\left(a_{i,0}+z_2^{-1}\left(a_{i,1}+z_2^{-1}\left(a_{i,2}+\cdots+z_2^{-1}\left(a_{i,N}\right)\cdots\right)\right)\right)\right. \\ &\left.+Y\left(b_{i,0}+z_2^{-1}\left(b_{i,1}+z_2^{-1}\left(b_{i,2}+\cdots+z_2^{-1}\left(b_{i,N}\right)\cdots\right)\right)\right)\right]. \end{aligned} \tag{4}$$

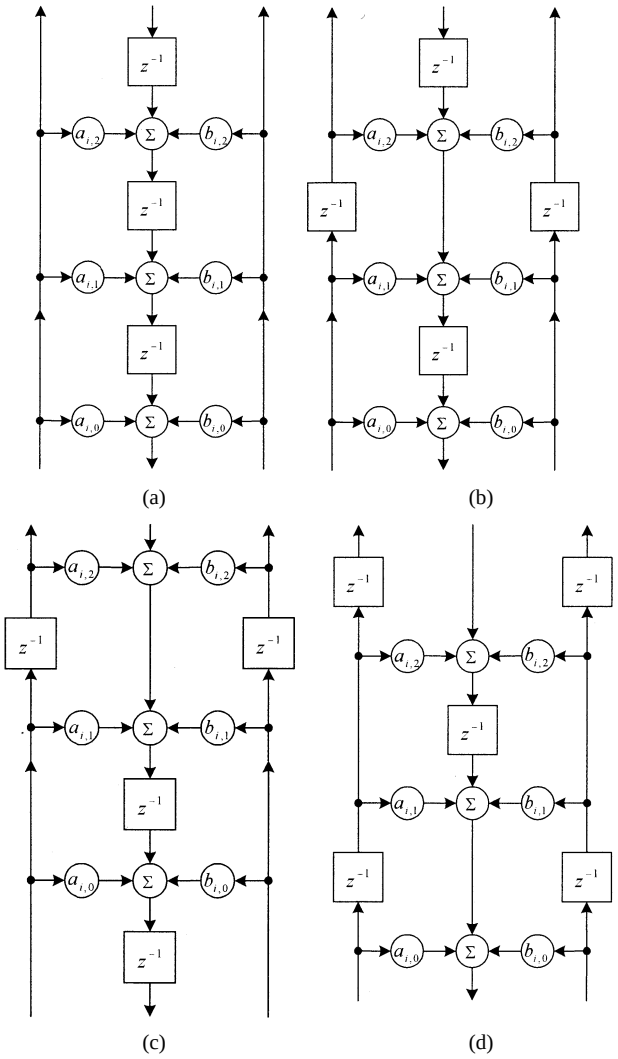

Fig. 4. Systolic transformations. (a) Original second-order relationship; (b) applied by [12]; (c) applied by [16]; and (d) applied by this paper.

where

$$\hat{X}_k = z_2^{-P}\hat{X}_{k-1}, \text{ for } k = 2, 3, \ldots, N \tag{8a}$$

$$\hat{Y}_k = z_2^{-P}\hat{Y}_{k-1}, \text{ for } k = 2, 3, \ldots, N \tag{8b}$$

and $\hat{X}_1 = z_2^{-P}X$, $\hat{Y}_1 = z_2^{-P}Y$. From (7), several 2-D local broadcast architectures in the ith-dimensional paths can be generated in the range of $1 \le P \le M-1$.

Next, we discuss how to realize the summation in the square bracket of (7) as a local broadcast architecture in the jth-dimensional paths. Shanbhag [12] utilizes the systolic transformation as shown in Fig. 4(b) instead of Fig. 4(a) to solve the jth dimensional global broadcast path. However, Shanbhag's scheme sacrifices the critical period to obtain the output, because $y(n,m)$ as shown in Fig. 3 is not immediately blocked by the delay element in the ith and jth dimensional output paths. Another systolic transformation as shown in Fig. 4(c) is used to construct a one-dimentional (1-D) IIR digital filter [16]. Nevertheless, how to construct a 2-D digital filter has not been discussed in [16]. For the sake of reducing the critical period and maintaining local broadcast in the jth-dimensional paths, we apply a new systolic transformation as shown in Fig. 4(d) to a 2-D IIR digital-filter design. In other words, the square brackets of (7) can be mapped to the structure of this new systolic transformation. We emphasize that although there exists the same/similar second-order relationship among Fig. 4(a)–(d), the systolic transformation in Fig. 4(d) is different from the one in Fig. 4(b) and 4(c) due to the different splitting of the delay element in Fig. 4(a). Utilizing this modified reordering scheme and a new systolic transformation, we can obtain a new 2-D systolic local broadcast digital filter architecture as shown in Fig. 5 with less critical period compared with Shanbhag's structure [12]. On the other hand, the critical path of the subblock, as plotted in dotted circle in Fig. 5 can be further improved by tree method. Therefore, the critical period only requires one multiplication and three additions. In addition, users merely set $b_{i,j}$ to zero such that a new systolic FIR digital filter architecture can be obtained.

B. 2-D Systolic Cascade-Form Digital Filter

The realization of digital filters by cascading the second-order IIR digital filters has many desirable features, such as less sensitivity to coefficient quantization error and better roundoff noise performance than the noncascade-form realization while in the fixed-point operation [17]. Under the same assumption $N_1 = N_2 = N$ as stated in Section II-A, the cascade-form transfer

$$\begin{aligned} Y &= \sum_{i=0}^{N}\sum_{j=0}^{N} a_{i,j} z_1^{-i} z_2^{-j} X + \sum_{i=0}^{N}\sum_{j=0}^{N} b_{i,j} z_1^{-i} z_2^{-j} Y \\ &= \sum_{j=0}^{N} a_{0,j} z_2^{-j} X + \sum_{j=0}^{N} b_{0,j} z_2^{-j} Y \\ &\quad + \sum_{i=1}^{N} z_1^{-i} z_2^{i\times P} \left[\left(\sum_{j=0}^{N} a_{i,j} z_2^{-j} \right) \left(z_2^{-i\times P} X \right) + \left(\sum_{j=0}^{N} b_{i,j} z_2^{-j} \right) \left(z_2^{-i\times P} Y \right) \right] \end{aligned} \tag{5}$$

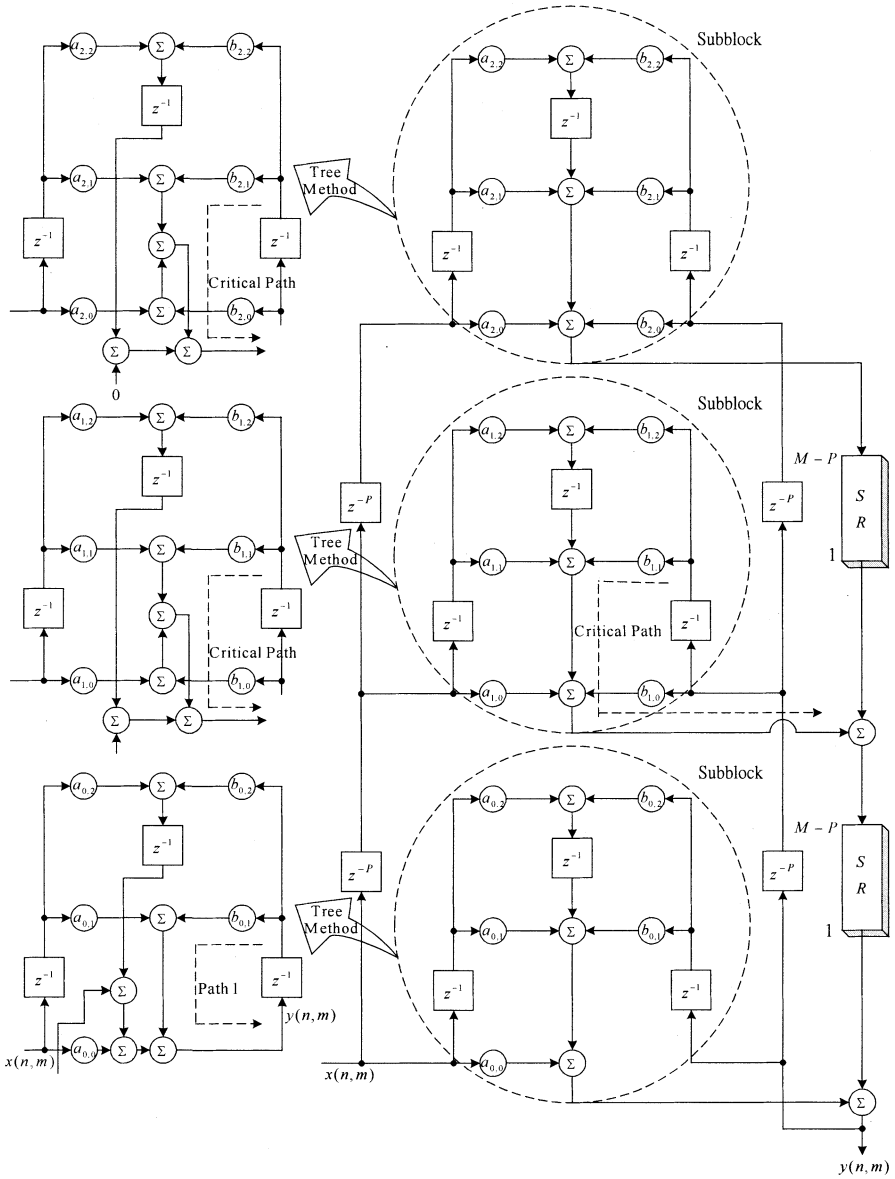

Fig. 5. A new proposed IIR digital filter applying a modified reordering scheme and a new systolic transformation for $N = 2$.

function based on a 2×2th-order IIR digital filter can be written as

$$H(z_1, z_2) = \prod_{l=1}^{N_s} H_l(z_1, z_2) = \prod_{l=1}^{N_s} \frac{\sum_{i=0}^{2}\sum_{j=0}^{2} a_{i,j,l} z_1^{-i} z_2^{-j}}{1 - \sum_{i=0}^{2}\sum_{j=0}^{2} b_{i,j,l} z_1^{-i} z_2^{-j}} \quad (9)$$

$$N_s = \left\lfloor \frac{N+1}{2} \right\rfloor \quad (10)$$

where $b_{0,0,l} = 0$ and the floor operator $\lfloor \bullet \rfloor$ denotes the maximum integer less than or equal to $\bullet$. We insert one unit delay at the output of each lth stage in order to achieve a systolic architecture and avoid the degradation in the critical period. Thus, (9) combined with X and Y can be modified as

$$Y = z^{-1} H_{N_s}(z_1, z_2)\, z^{-1} H_{N_s-1}(z_1, z_2) \cdots z^{-1} H_1(z_1, z_2) X. \quad (11)$$

The amount of inserting delays is equal to N_s. These inserting delays for the 2-D systolic cascade-form architecture would not affect the magnitude response but just results in the latency of N_s. For example, let $N = 4$ and the resulting 2-D systolic cascade-form IIR digital filter architecture is revealed in Fig. 6, where $N_s = 2$ and the detailed block diagrams of PE_0 and PE_1 are shown in Figs. 7(a) and 7(b), respectively. Note that the proposed cascade-form architecture certainly has the local broadcast characteristic.

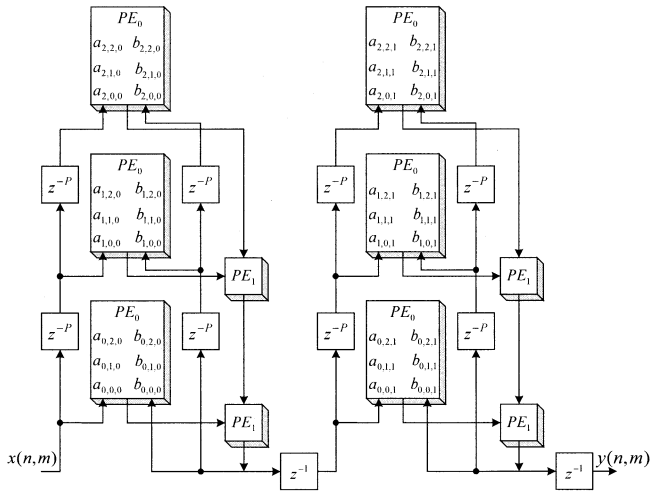

Fig. 6. Systolic array of a 2-D cascade-form IIR digital filter for $N = 4$.

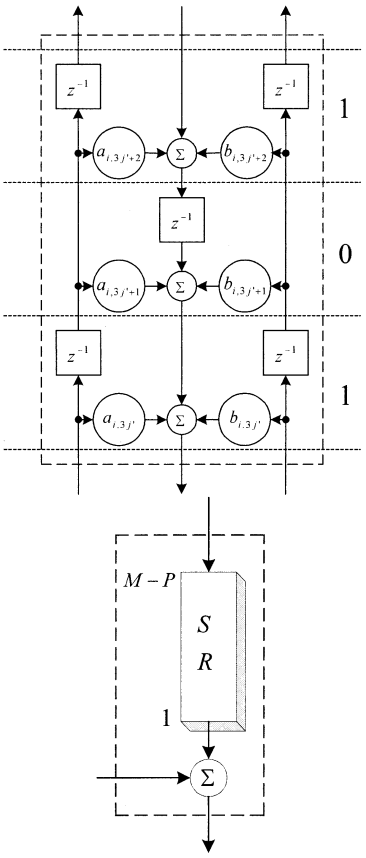

Fig. 7. Detailed block diagrams of two types of PEs. (a) PE_0. (b) PE_1.

III. Error Analysis of New Digital Filter Architectures

In this section, we investigate error analysis for finite word-length arithmetic in the proposed 2-D systolic IIR and FIR digital filters. Here, we adopt most notations as defined in [12], [18], [19] to analyze quantization error. For convenience of error analysis, the 2-D local broadcast $N \times N$th order digital filter can be equivalently plotted in Fig. 8. Let $PE_{i,j'}$ signify the position of PE_0 in the ith row and the j'th column of Fig. 8, where $0 \le i \le N$ and $0 \le j' \le \lfloor N/3 \rfloor$. Observing the block diagram in Fig. 8, the two indices j' and j are related as

$$j' = \left\lfloor \frac{j}{3} \right\rfloor. \tag{12}$$

If X represents the true value of a variable, then its quantized value would be represented by $\overline{X}$. Let $\alpha_{i,j}$ and $\beta_{i,j}$ denote coefficient quantization errors in the representation of finite word-length $a_{i,j}$ and $b_{i,j}$, respectively. Also, the input and output quantization errors are denoted as $e_X(n,m)$ and $e_Y(n,m)$, respectively, while the input $x(n,m)$ and the final output $y(n,m)$ operate in finite word-length arithmetic. We derive the error expression at the output of $PE_{i,j'}$ as shown in Fig. 8 for $0 \le j' \le \lfloor N/3 \rfloor$. In other words, we consider PEs for type PE_0 as shown in Fig. 7(a). Let $y_{i,j'}(n,m)$ denote the true output of $PE_{i,j'}$ and it can be represented as

$$\begin{aligned} y_{i,j'}(n,m) =& a_{i,3j'+2}x\left(n-\delta_1, m-3j'-2-\xi_1\right) \\ &+ a_{i,3j'+1}x\left(n-\delta_1, m-3j'-1-\xi_1\right) \\ &+ a_{i,3j'}x\left(n-\delta_1, m-3j'-\xi_1\right) \\ &+ b_{i,3j'+2}y\left(n-\delta_1, m-3j'-2-\xi_1\right) \\ &+ b_{i,3j'+1}y\left(n-\delta_1, m-3j'-1-\xi_1\right) \\ &+ b_{i,3j'}y\left(n-\delta_1, m-3j'-\xi_1\right) \\ &+ y_{i,j'+1}(n, m-1) \end{aligned} \tag{13}$$

where

$$\delta_1 = \left\lfloor \frac{i \times P}{M} \right\rfloor, \quad \text{for } 1 \le P \le M-1 \tag{14a}$$

$$\xi_1 = (i \times P) \bmod M, \quad \text{for } 1 \le P \le M-1 \tag{14b}$$

where mod denotes the modulus operation. Therefore, the quantized value of $y_{i,j'}(n,m)$ is given by

$$\begin{aligned} \overline{y}_{i,j'}(n,m) =& \left[\overline{a}_{i,3j'+2}\overline{x}\left(n-\delta_1, m-3j'-2-\xi_1\right)\right] r_q \\ &+ \left[\overline{a}_{i,3j'+1}\overline{x}\left(n-\delta_1, m-3j'-1-\xi_1\right)\right]_q \\ &+ \left[\overline{a}_{i,3j'}\overline{x}\left(n-\delta_1, m-3j'-\xi_1\right)\right]_q \\ &+ \left[\overline{b}_{i,3j'+2}\overline{y}\left(n-\delta_1, m-3j'-2-\xi_1\right)\right]_q \\ &+ \left[\overline{b}_{i,3j'+1}\overline{y}\left(n-\delta_1, m-3j'-1-\xi_1\right)\right]_q \\ &+ \left[\overline{b}_{i,3j'}\overline{y}\left(n-\delta_1, m-3j'-\xi_1\right)\right]_q \\ &+ \overline{y}_{i,j'+1}(n, m-1) - s_{i,j'} \end{aligned} \tag{15}$$

where $s_{i,j'}$ is defined as the storage error caused by storing the output of an adder in a register. In [12], [19], it has been discussed that the rounding operation with a $(t+1)$-bit register leads to $s_{i,j'} \le 2^{-t-1}$ for fixed-point data arithmetic. Representing the error at the output of $PE_{i,j'}$ by $f_{i,j'}(n,m)$ and neglecting second-order terms involving $\alpha_{i,3j'+k_1}e_X\left(n-\delta_1, m-3j'-k_1-\xi_1\right)$ and $\beta_{i,3j'+k_1}e_Y\left(n-\delta_1, m-3j'-k_1-\xi_1\right)$ for $k_1 = 0$, 1 and 2, we obtain

$$\begin{aligned} f_{i,j'}(n,m) =& y_{i,j'}(n,m) - \overline{y}_{i,j'}(n,m) \\ =& \sum_{k_1=0}^{2} \left(p_{i,3j'+k_1}\left(n-\delta_1, m-3j'-k_1-\xi_1\right)\right. \end{aligned}$$

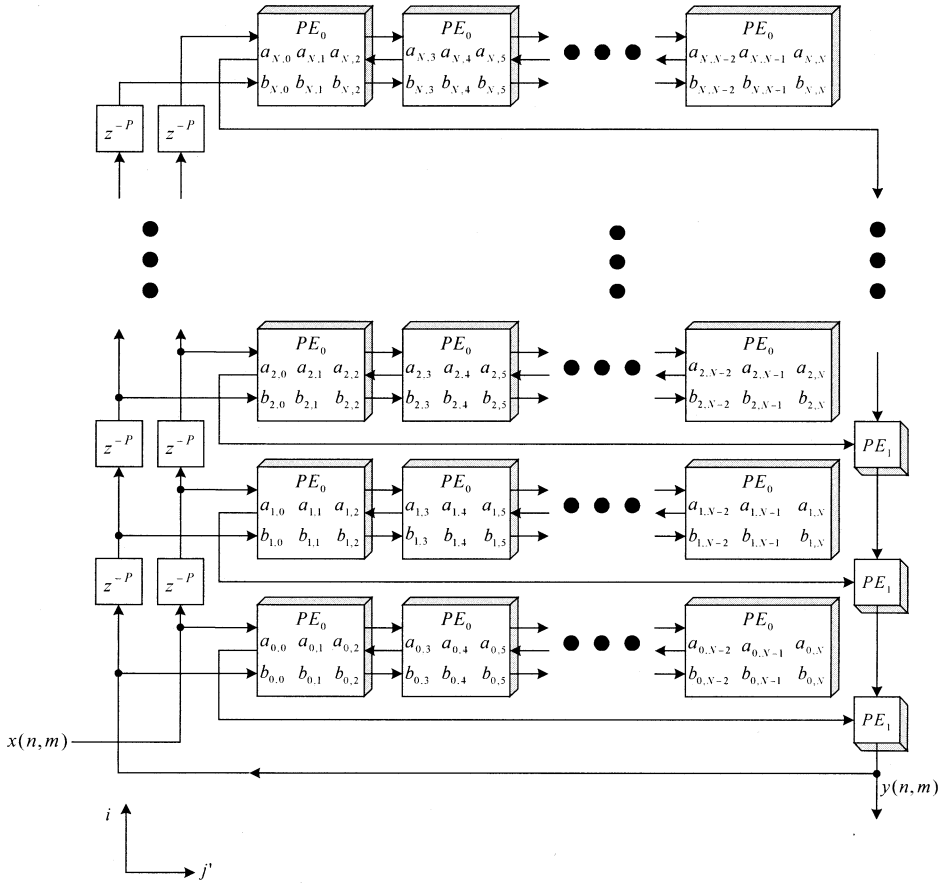

Fig. 8. Block diagram of the proposed 2-D $N \times N$th order digital filter architecture for error analysis.

$$
\begin{aligned}
&+ r_{i,3j'+k_1}(n-\delta_1, m-3j'-k_1-\xi_1) \\
&+ \alpha_{i,3j'+k_1} x(n-\delta_1, m-3j'-k_1-\xi_1) \\
&+ \beta_{i,3j'+k_1} y(n-\delta_1, m-3j'-k_1-\xi_1) \\
&+ a_{i,3j'+k_1} e_X(n-\delta_1, m-3j'-k_1-\xi_1) \\
&+ b_{i,3j'+k_1} e_Y(n-\delta_1, m-3j'-k_1-\xi_1)) \\
&+ f_{i,j'+1}(n, m-1) + s_{i,j'} \qquad (16)
\end{aligned}
$$

where $p_{i,j}(n,m)$ and $r_{i,j}(n,m)$ referred to as multiplication roundoff errors are defined, respectively, as

$$p_{i,j}(n,m) = \overline{a}_{i,j}\overline{x}(n,m) - \left[\overline{a}_{i,j}\overline{x}(n,m)\right]_q \qquad (17a)$$

$$r_{i,j}(n,m) = \overline{b}_{i,j}\overline{y}(n,m) - \left[\overline{b}_{i,j}\overline{y}(n,m)\right]_q. \qquad (17b)$$

Now, we accumulate the error from $j'=0$ to $j'=\lfloor j/3 \rfloor$ and thus the combined error at the output of $PE_{i,0}$ can be treated as (18) shown at the bottom of the next page. Summing $f_{i,0}(n,m)$ in (18) for $0 \le i \le N$, we obtain the final combined error as

$$
\begin{aligned}
f(n,m) &= \sum_{i=0}^{N} f_{i,0}(n-\delta_2, m-\xi_2) + \sum_{i=1}^{N} s_i \\
&= m_E + c_E + i_E + s_E \\
&\quad + \sum_{i=0}^{N}\sum_{j=0}^{N} b_{i,j} e_Y(n-i, m-j) \qquad (19)
\end{aligned}
$$

where

$$\delta_2 = \left\lceil \frac{i \times (M-P)}{M} \right\rceil, \quad \text{for } 1 \le P \le M-1 \qquad (20a)$$

$$\xi_2 = (i \times (M-P)) \bmod M, \text{ for } 1 \le P \le M-1 \qquad (20b)$$

$$
\begin{aligned}
m_E = \sum_{i=0}^{N}\sum_{j=0}^{N} &(p_{i,j}(n-i, m-j) \\
&+ r_{i,j}(n-i, m-j)) \qquad (20c)
\end{aligned}
$$

$$
\begin{aligned}
c_E = \sum_{i=0}^{N}\sum_{j=0}^{N} &(\alpha_{i,j} x(n-i, m-j) \\
&+ \beta_{i,j} y(n-i, m-j)) \qquad (20d)
\end{aligned}
$$

$$i_E = \sum_{i=0}^{N}\sum_{j=0}^{N} a_{i,j} e_X(n-i, m-j) \qquad (20e)$$

$$s_E = \sum_{i=0}^{N}\sum_{j'=0}^{\lfloor N/3 \rfloor} s_{i,j'} + \sum_{i=1}^{N} s_i. \qquad (20f)$$

The ceiling operator $\lceil \bullet \rceil$ denotes a minimum integer that is greater than or equal to $\bullet$ and m_E, c_E, i_E, and s_E represent roundoff error, coefficient-quantization error, input-quantization error and storage error, respectively. Note that the integer variable P does not affect quantization error and, in fact, its main aim is to provide several local broadcast architectures.

TABLE I
COMPARISON RESULT AMONG DIFFERNT SIZES OF PES

Size of *PE*	10 (1st order)	101 (2nd order)	1011 (3rd order)	10111 (4th order)
Critical Period	$T_m + 3T_a$	$T_m + 3T_a$	$T_m + 4T_a$	$T_m + 4T_a$
Storage Error	$\sum_{i=0}^{N}\sum_{j'=0}^{\lfloor N/2\rfloor} s_{i,j'} + \sum_{i=1}^{N} s_i$	$\sum_{i=0}^{N}\sum_{j'=0}^{\lfloor N/3\rfloor} s_{i,j'} + \sum_{i=1}^{N} s_i$	$\sum_{i=0}^{N}\sum_{j'=0}^{\lfloor N/4\rfloor} s_{i,j'} + \sum_{i=1}^{N} s_i$	$\sum_{i=0}^{N}\sum_{j'=0}^{\lfloor N/5\rfloor} s_{i,j'} + \sum_{i=1}^{N} s_i$

TABLE II
COMPARISON RESULTS AMONG IIR DIGITAL FILTER ARCHITECTURES

Parameters	Sid-Ahmed [10]	SCH3 of S-G-A [11]	Shanbhag [12]	This Work
Global Broadcast	Input and Output	Input and Output	Input and Output	No
Storage Error	$\sum_{i=0}^{N}\sum_{j=0}^{N} s_{i,j}$	$\sum_{i=0}^{N}\sum_{j=0}^{N} s_{i,j}$	$\sum_{i=0}^{N}\sum_{j'=1}^{\lceil N/2\rceil} s_{i,j'} + \sum_{i=1}^{N} s_i$	$\sum_{i=0}^{N}\sum_{j'=0}^{\lfloor N/3\rfloor} s_{i,j'} + \sum_{i=1}^{N} s_i$
Critical Period	$T_m + (2+\lceil \log_2(N+1)\rceil)T_a$	$T_m + 2T_a$	$2T_m + 2T_a$	$T_m + 3T_a$
Latency	1	1	0	0
No. of Multipliers	$2(N+1)^2 - 1$	$2(N+1)^2 - 1$	$2(N+1)^2 - 1$	$2(N+1)^2 - 1$
No. of Delay Elements	$(N+1)^2 + 2MN$	$(N+1)^2 + (M+1)N$	$\frac{3N(N+1)}{2} + MN$	$5(\lfloor N/3\rfloor+1)(N+1) + (M+P)N$

By setting all $b_{i,j}$ and $\beta_{i,j}$ to zero, the error expression corresponding to a new FIR digital filter architecture can be expressed as

$$f(n,m) = m_E + c_E + i_E + s_E \tag{21}$$

where

$$m_E = \sum_{i=0}^{N}\sum_{j=0}^{N} p_{i,j}(n-i, m-j) \tag{22a}$$

$$c_E = \sum_{i=0}^{N}\sum_{j=0}^{N} \alpha_{i,j} x(n-i, m-j). \tag{22b}$$

The other two terms i_E and s_E are the same as (20e) and (20f), respectively. If we assume that $s_{i,j'}$ and s_i, for all the architectures under consideration, are all of a similar nature, then it is apparent that our architecture has lower-storage error than existing structures [10]–[12]. This is due to the fact that the proposed architecture in Fig. 8 has fewer PEs and thus, the lower sum of storage error is obtained in (19) and (21). Consequently, the quantization error is reduced.

It is worthy to note that the storage error depends on different types and sizes of PEs. In this paper, the PEs are restricted to the type that is either PE_0 or extended type for PE_0. Under this constraint, we define two useful notations 0 and 1, where 0 and 1 indicate that the delay elements as shown in Fig. 7(a) are at top-to-down and bottom-to-up signal paths, respectively. Thus, we can easily use the digital sequence to represent different sizes of PEs. For example, PE_0 proposed in Section II can be represented as 101. While the larger size of PE is used to construct the 2-D digital filter, the lower-storage error is achieved; however, the critical period is sacrificed as listed in Table I. With minimum-critical period and low-storage error in mind, we select the second-order PE denoted as 101 (i.e., PE_0) for our design.

$$\begin{aligned} f_{i,0}(n,m) &= \sum_{j'=0}^{\lfloor N/3\rfloor}\left(\sum_{k_1=0}^{2}\Big(p_{i,3j'+k_1}(n-\delta_1, m-3j'-k_1-\xi_1) + r_{i,3j'+k_1}(n-\delta_1, m-3j'-k_1-\xi_1)\right. \\ &\quad + \alpha_{i,3j'+k_1} x(n-\delta_1, m-3j'-k_1-\xi_1) + \beta_{i,3j'+k_1} y(n-\delta_1, m-3j'-k_1-\xi_1) \\ &\quad \left. + a_{i,3j'+k_1} e_X(n-\delta_1, m-3j'-k_1-\xi_1) + b_{i,3j'+k_1} e_Y(n-\delta_1, m-3j'-k_1-\xi_1)\Big) + s_{i,j'}\right) \\ &= \sum_{j=0}^{N}\Big(p_{i,j}(n-\delta_1, m-j-\xi_1) + r_{i,j}(n-\delta_1, m-j-\xi_1) \\ &\quad + \alpha_{i,j} x(n-\delta_1, m-j-\xi_1) + \beta_{i,j} y(n-\delta_1, m-j-\xi_1) \\ &\quad + a_{i,j} e_X(n-\delta_1, m-j-\xi_1) + b_{i,j} e_Y(n-\delta_1, m-j-\xi_1)\Big) + \sum_{j'=0}^{\lfloor N/3\rfloor} s_{i,j'}. \end{aligned} \tag{18}$$

TABLE III
COMPARISON RESULTS AMONG FIR DIGITAL FILTER ARCHITECTURES

Parameters	Sid-Ahmed [10]	SCH3 of S-G-A [11]	Shanbhag [12]	This Work
Global Broadcast	Input	Input	Input	No
Storage Error	$\sum_{i=0}^{N}\sum_{j=0}^{N} s_{i,j}$	$\sum_{i=0}^{N}\sum_{j=0}^{N} s_{i,j}$	$\sum_{i=0}^{N}\sum_{j'=1}^{\lceil N/2\rceil} s_{i,j'} + \sum_{i=1}^{N} s_i$	$\sum_{i=0}^{N}\sum_{j'=0}^{\lfloor N/3\rfloor} s_{i,j'} + \sum_{i=1}^{N} s_i$
Critical Period	$\max\{(T_m + T_a), (\lceil \log_2(N+1)\rceil)T_a\}$	$T_m + T_a$	$T_m + 2T_a$	$T_m + 2T_a$
Latency	1	1	0	0
No. of Multipliers	$(N+1)^2$	$(N+1)^2$	$(N+1)^2$	$(N+1)^2$
No. of Delay Elements	$(N+1)^2 + MN$	$(N+1)^2 + MN$	$N(N+1) + MN$	$3(\lfloor N/3\rfloor+1)(N+1) + MN$

Moreover, the quantization error for a 2-D systolic cascade-form digital filter is also investigated in similar fashion. The resulting error expression at the lth stage is briefly summarized as

$$f^{(l)}(n,m) = m_{E,l} + c_{E,l} + i_{E,l} + s_{E,l} + \sum_{i=0}^{2}\sum_{j=0}^{2} b_{i,j,l} f^{(l)}(n-i, m-j), \quad \text{for } l = 1, 2, \ldots, N_s \tag{23}$$

where

$$m_{E,l} = \sum_{i=0}^{2}\sum_{j=0}^{2} \left(p_{i,j,l}(n-i, m-j-1) + r_{i,j,l}(n-i, m-j)\right) \tag{24a}$$

$$c_{E,l} = \sum_{i=0}^{2}\sum_{j=0}^{2} \left(\alpha_{i,j,l} y^{(l-1)}(n-i, m-j-1) + \beta_{i,j,l} y^{(l)}(n-i, m-j)\right) \tag{24b}$$

$$i_{E,l} = \sum_{i=0}^{2}\sum_{j=0}^{2} a_{i,j,l} f^{(l-1)}(n-i, m-j-1) \tag{24c}$$

$$s_{E,l} = \sum_{i=0}^{2}\sum_{j'=0}^{\lfloor 2/3\rfloor} s_{i,j',l} + \sum_{i=1}^{2} s_{i,l}. \tag{24d}$$

In similar fashion $f^{(l-1)}(n,m)$ can be evaluated and then $f^{(l-1)}(n, m-1)$ can be obtained utilizing $(m-1)$ instead of m. Finally, (23) can be solved recursively. Note that $b_{0,0,l} = \beta_{0,0,l} = 0$, $f^{(0)}(n,m) = e_X(n,m)$ and $f^{(N_s)}(n,m) = e_Y(n,m)$.

IV. COMPARISON RESULTS OF IIR AND FIR DIGITAL FILTERS

In this section, we make an effort to compare our architecture with existing architectures [10]–[12]. Comparison results of 2-D IIR digital filters are tabulated in Table II in terms of storage error, critical period, latency, the number of multipliers and delay elements and, importantly, whether the input and output signals locally broadcast in these structures. In Table II, the proposed hybrid of two schemes completely eliminates the ith and jth dimensional global broadcast paths. Besides, if it is assumed that $s_{i,j'}$, $s_{i,j}$ and s_i in Table II are all of a similar nature, then the latter scheme (i.e., a new systolic transformation) leads to lower-storage error than that in [10]–[12]. Let T_m and T_a represent the operation time required for one multiplication and one addition, respectively. So as to minimize periods in critical paths as shown in Figs. 1, 2, 3, and 5, we properly apply tree method to those structures and then separately evaluate the periods. As a result, we detect that this work has higher throughput than Sid-Ahmed's [10] and Shanbhag's structure [12] but requires extra one addition than S-G-As structure applying SCH3 [11]. In general cases, since M is much larger than N, the number of delay elements in Table II is dominated by the product-term MN. Hence, the number of delay elements in this work for small value P is almost equal to that of [11] as well as [12] and less than that of [10]. In the same way, comparison results as listed in Table III among 2-D FIR digital filter architectures can be obtained after setting $b_{i,j}$ to zero. From Tables II and III, it turns out that the proposed architecture has local broadcast as well as lower-storage error and maintains zero latency under the acceptable critical period.

V. CONCLUSION

A new systolic architecture for the implementation of 2-D IIR and FIR digital filters has been proposed by a modified reordering scheme and a new systolic transformation. Applying the hybrid of two schemes, better performance showing local broadcast, lower-quantization error, zero latency, and the satisfactory critical period without sacrificing other hardware characteristics can be achieved. In addition, we extend the new architecture to a 2-D systolic cascade-form digital filter and offer quantization error analysis related to the proposed architectures.

ACKNOWLEDGMENT

The authors wish to thank Chip Implementation Center (CIC), National Science Council, Taiwan, R.O.C., for financial support. The authors would also like to thank the associative editor and anonymous referees for their valuable suggestions to this paper.

REFERENCES

[1] A. M. Tekalp, *Digital Video Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1995, ch. 14.

[2] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[3] M. A. Sid-Ahmed, *Image Processing: Theory, Algorithms and Architectures*. New York: McGraw-Hill, 1995.
[4] H. T. Kung, "Why systolic architectures?," *IEEE Comput.*, vol. C-25, pp. 37–46, Jan. 1982.
[5] N. H. E. Weste and K. Eshraghian, *Principles of COMS VLSI Design: A Systems Perspective*, 2nd ed. Reading, MA: Addison-Wesley, 1993, ch. 3–5.
[6] L. D. Van, "Design of Efficient VLSI Architectures: Multiplier, 2-D Digital Filter and Adaptive Digital Filter," Ph.D. dissertation, Dept. of the Electrical Engineering, National Taiwan Univ., Taipei, Taiwan, R.O.C., 2001.
[7] S. Y. Kung, *VLSI Array Processors*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
[8] P. Pirsch, *Architectures for Digital Signal Processing*. New York: Wiley, 1998, ch. 5.
[9] K. K. Parhi, *VLSI Digital Signal Processing Systems: Design and Implementation*. New York: Wiley, 1999, ch. 7.
[10] M. A. Sid-Ahmed, "A systolic realization for 2-D digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 560–565, Apr. 1989.
[11] 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.
[12] N. R. Shanbhag, "An improved systolic architecture for 2-D digital filters," *IEEE Trans. Signal Processing*, vol. 39, pp. 1195–1202, May 1991.
[13] L. D. Van, C. C. Tang, S. Tenqchen, and W. S. Feng, "A new VLSI architecture without global broadcast for 2-D systolic digital filters," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, Geneva, Switzerland, May 2000, pp. 547–550.
[14] L. D. Van, S. Tenqchen, C. H. Chang, and W. S. Feng, "A new 2-D digital filter using a locally broadcast scheme and its cascade form," in *Proc. IEEE Asia Pacific Conf. on Circuits Syst.*, Tianjin, China, Dec. 2000, pp. 579–582.
[15] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1984, ch. 4.
[16] Jayadema, "A new systolic design for digital IIR filters," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 653–654, May 1990.
[17] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989, ch. 6.
[18] B. G. Mertzios and A. N. Venetsanopoulos, "Combined error at the output of two-dimensional recursive digital filters," *IEEE Trans. Circuits Syst.*, vol. CAS-31, pp. 888–891, Oct. 1984.
[19] D. Raghuramireddy and R. Unbehauen, "Error analysis of a systolic realization of 2-D filters," *IEEE Trans. Signal Processing*, vol. 40, pp. 478–482, Feb. 1992.

Lan-Da Van (S'98–A'01–M'02) was born in Miaoli, Taiwan, R.O.C., on October 9, 1972. He received the B.S., (Honors) and the M.S. degrees from Tatung Institute of Technology, Taipei, Taiwan, R.O.C., in 1995 and 1997, respectively. He received the Ph.D. degree from National Taiwan University (NTU), Taipei, Taiwan, R.O.C., in 2001, all in electrical engineering.

From 1997 to 2001, he was a Research Assistant at NTU. Since 2001, he has been a Researcher at the Chip Implementation Center (CIC), National Science Council, Hsinchu, Taiwan, R.O.C. His current research interests are in VLSI architectures, algorithms, and chips for digital signal processing (DSP), baseband communication systems, and computer applications. This includes the design of high-performance/low-power/low-area adaptive digital filters. He has published more than 11 IEEE journal and conference papers in these areas.

Dr. Van was a recipient of the Chunghwa Picture Tube (CPT) and Motorola Scholarships in 1996 and 1997, respectively. He was an elected Vice-Chairman and Chairman of the IEEE NTU Student Branch in 1999 and 2000, respectively. In 2002, he received an IEEE award for outstanding leadership and service to theIEEE NTU Student Branch. He was a referee of IEEE ISCAS. Presently, he serves as a reviewer for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I: FUNDAMENTAL THEORY AND APPLICATIONS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: ANALOG AND DIGITAL SIGNAL PROCESSING, and IEEE TRANSACTIONS ON MULTIMEDIA.