2-D Digital Filter Architectures without Global Broadcast and Some Symmetry Applications

I-Hung Khoo & Hari C. Reddy*
Dept. of Electrical Engineering
California State University Long Beach
Long Beach, CA, U.S.A.
*and College of Computer Science, National Chiao-Tung University, Hsinchu, Taiwan, R.O.C

Lan-Da Van & Chin-Teng Lin
Department of Computer Science
National Chiao-Tung University
Hsinchu, Taiwan, R.O.C

***Abstract*— Four new 2-D filter VLSI architectures without global broadcast are presented. The first is a transposed systolic structure which requires fewer delay elements compared to the original systolic structure in [1]. By combining the sub-blocks of the original with the new transposed structure, two additional systolic structures are obtained to realize transfer functions with separable denominators, which require fewer multipliers. These separable denominator structures have important symmetry applications. A structure which possesses diagonal symmetry is then shown which requires even fewer multipliers.**

I. INTRODUCTION

Two-dimensional (2-D) digital filters find applications in many digital signal processing areas such as image processing and seismic data processing. Although 2-D digital filters can be simulated on a general purpose computer, for applications involving high data rate, such as real time image processing, dedicated computing structures are needed in order to meet the high throughput demands. Networks using systolic arrays are popular candidates for VLSI applications that are computationally intensive. Systolic arrays are easily realized as an application-specific integrated circuit (ASIC) due to the regularity and modularity of the processing elements (PEs) involved, and the local data communication among them. Thus, having global broadcast of signals is undesirable for systolic designs. In [1], a 2-D systolic digital filter architecture was presented which eliminated the global broadcast of the input and output signals in previous architectures [2,3]. In this paper, a transpose of the structure in [1] is presented which requires fewer delay elements while maintaining the same speed and latency (defined as the elapsed clock cycles between the application of the first input sample and the appearance of the first output sample). It is then shown that by combining the sub-blocks of the original structure in [1] with the new transposed structure, two additional systolic structures can be obtained which realize transfer functions with separable denominator. Both the separable denominator structures require fewer multipliers, and one of them also has the advantage of a lower critical period which means higher speed. Beside these advantages, another motivation for deriving the separable denominator structures is for the symmetry applications. The frequency responses of most 2-D filters possess some form of symmetry which can be used to further reduce the number of multipliers needed in the implementation [4]. Here, a structure possessing diagonal symmetry is derived from the separable denominator structures. This new structure requires fewer multipliers and has a higher speed than a diagonal symmetry structure recently presented in [5].

II. PRELIMINARIES

A general 2-D IIR transfer function can be represented as in (1), where $b_{00}=0$, N_1xN_2 is the order of the filter, and X and Y are respectively the input and output of the filter. The equation can also represent an FIR transfer function if we set $b_{ij}=0$ for all i and j.

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

Now assuming that the filter is used to process a square image of size MxM and the pixel values in the image are fed to the filter in raster-scan mode, i.e. the input sequence is $x(0,0)$, $x(0,1)$, …, $x(0,M\text{-}1)$, $x(1,0)$, $x(1,1)$, … etc, then we can replace z_2^{-1} by a single delay register, z^{-1}, and z_1^{-1} by a shift register of length M, z^{-M}, provided $M>N_2$. Without loss of generality, we can assume $N_1=N_2=N$ in discussing the structures.

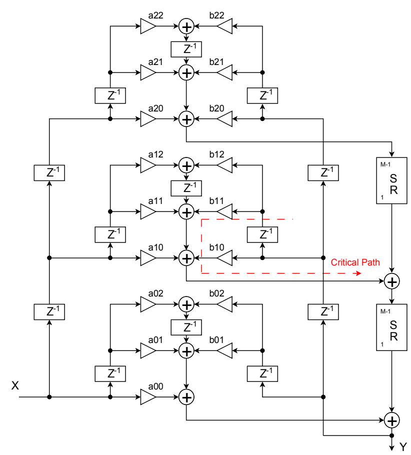

Figure 1. Original systolic structure presented in [1]

A 2x2 systolic IIR filter from [1] is shown in Fig. 1. Notice that the shift registers (SR) are of length $M\text{-}1$ due to the additional delays added at the input and output branches to eliminate the global broadcast. It can easily be verified using Mason's gain formula that the structure, with $z^{-1}=z_2^{-1}$ and $\mathrm{SR}=z_1^{-1}z_2$, possesses the transfer function in (1). The 2x2 structure in Fig. 1 consists of 3 sub-blocks realizing multiplier sets: $\{a_{0j},b_{0j}\}$, $\{a_{1j},b_{1j}\}$, $\{a_{2j},b_{2j}\}$. The structure can be extended to NxN by adding additional sub-blocks/shift registers etc. The critical path (slowest path) of the structure is indicated in red in Fig. 1. The resulting critical period, which determines the speed of the structure, is Tm+3Ta, where Tm and Ta represent the operation time required by the multiplier and adder respectively.

III. New Transposed Structure

By taking the transpose of Fig. 1 and rearranging some of the delays in the sub-blocks, the following new structure is obtained. Its transfer function can easily be verified using the Mason's gain formula.

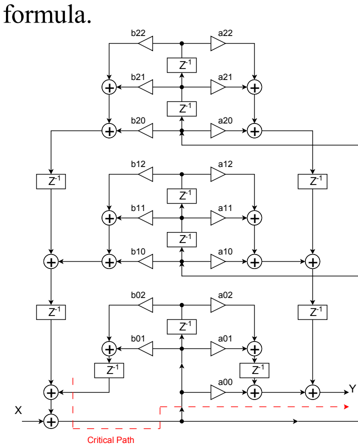

Figure 2. New transposed structure

This new transposed structure has fewer delays than that in Fig. 1, 9+2M versus 11+2M. The critical path is at a different location, but the critical period is still Tm+3Ta as shown in Fig. 3 below with the adders rearranged at the output.

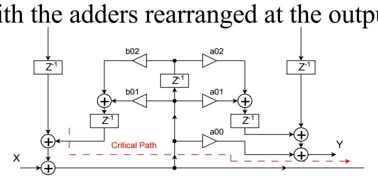

Figure 3. Critical path of transposed structure with adders rearranged

The transposed structure can be easily extended to NxN. This is done by first decomposing Fig. 2 into its basic PEs, as shown below, and using them to form the higher order structure.

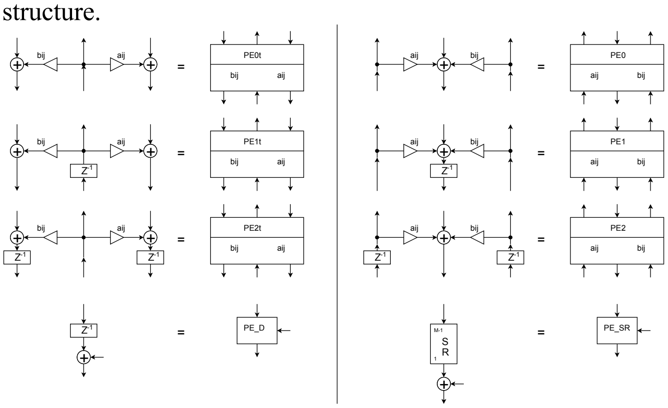

Figure 4. Processing elements

PE0t, PE1t, PE2t are used in the sub-blocks of the transposed structure, while PE_SR realizes the shift register. PE_D consists of a delay element and is used in the input and output branches to eliminate global broadcast. PE0, PE1, PE2 correspond to the original structure in Fig. 1 and will be used later to realize the separable denominator structures.

The resulting *N*x*N* transposed structure is shown in Fig. 5. In order to keep the critical period at Tm+3Ta, the bottom sub-block needs to have PE order: (PE0t,PE2t,PE1t, PE1t, PE1t, PE1t, PE1t, PE1t, PE2t, PE1t, PE1t, PE1t, PE1t, PE1t, PE1t, PE2t,…) while the non-bottom sub-blocks should have PE order: (PE0t, PE1t, PE1t, PE1t, PE1t, PE1t, PE1t, PE2t, PE1t, PE1t, PE1t, PE1t, PE1t, PE1t, PE2t,….).

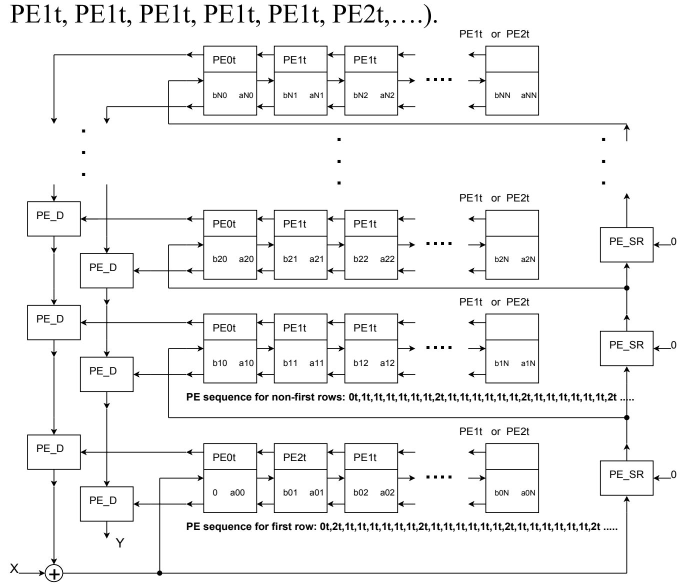

Figure 5. Transposed structure (NxN)

The comparison between the new transposed structure and the original structure in [1] is shown in Table 1. The following graph highlights the advantage, in terms of the number of delays, of the new transposed structure (blue) over the original structure in [1] (red) as the order N increases.

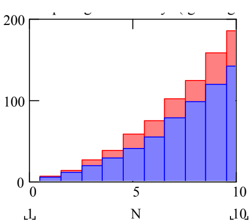

Figure 6. Comparing the number of delays (ignoring SR)

The transposed structure can also be used to realize FIR filters by removing the feedback portion involving the coefficients b_{ij}. The resulting structure, however, has the same number of delay elements as the corresponding FIR structure presented in [1].

IV. New Separable Denominator Structures

A transfer function with separable denominator is shown in (2).

$$H(z_1,z_2)=\frac{Y(z_1,z_2)}{X(z_1,z_2)}=\frac{\sum_{i=0}^{N_1}\sum_{j=0}^{N_2} a_{ij} z_1^{-i} z_2^{-j}}{\left(1-\sum_{i=1}^{N_1} b_{i0} z_1^{-i}\right)\cdot\left(1-\sum_{j=1}^{N_2} b_{0j} z_2^{-j}\right)} \quad (2)$$

This transfer function has several advantages over the general one in (1). Firstly, the stability is easy to check by simply solving for the poles of the two 1-D polynomials. Any unstable pole (z_{p_i}) can then be stabilized by replacing with its inverse pole ($1/z_{p_i}$) without affecting the magnitude response [4]. Alternatively, it can be stabilized using planar least-squares inverse approach. Secondly, the separable denominator requires fewer multipliers to realize. Thirdly, the separable denominator is useful in realizing magnitude responses possessing various symmetries [4]. This is illustrated later with a structure with diagonal symmetry. Structures possessing other symmetries will be reported in future publications.

Two new structures realizing the separable denominator transfer function in (2) can be obtained by combining the sub-blocks in Fig. 1 and Fig. 2. The first version is shown in Fig. 7. It involves using PE0, PE1, PE2 in the bottom sub-block, and PE0t, PE1t, PE2t in the rest of the sub-blocks. Its transfer function can easily be verified using Mason's gain formula. The critical period is Tm+3Ta.

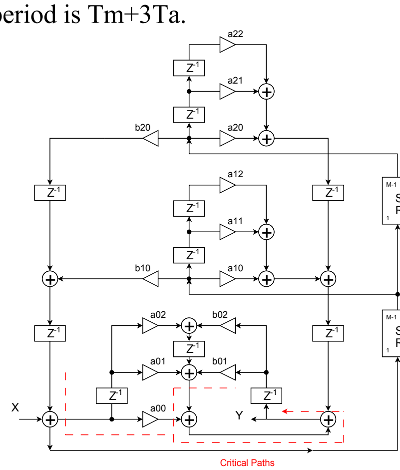

Figure 7. New separable denominator structure (version 1)

By using the PEs listed in Fig. 4, the structure can be extended to NxN as shown in Fig. 8. In order to keep the critical period at Tm+3Ta, the bottom sub-block needs to have PE order: (PE0,PE2, PE1, PE2, PE1, PE2…) while the non-bottom sub-blocks should have the same PE order as the non-bottom sub-blocks in Fig 5.

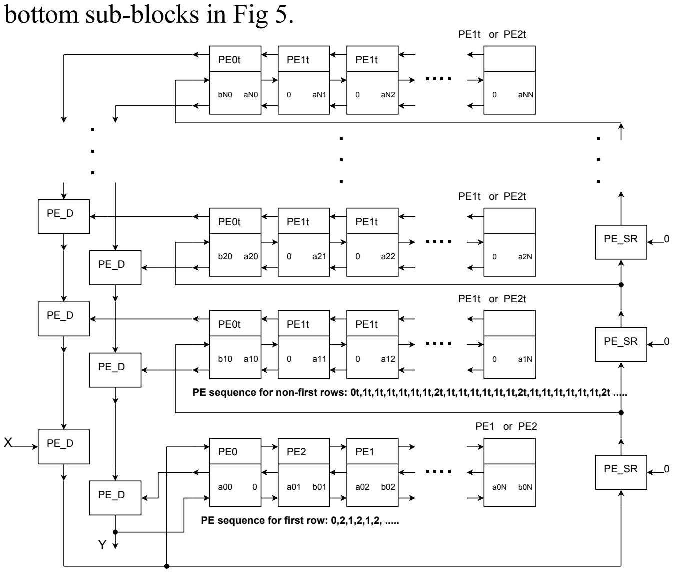

Figure 8. Separable denominator structure (version 1) -- NxN

The second version of the separable denominator structure is shown in Fig. 9. It involves using PE0t, PE1t, PE2t in the bottom sub-block, and PE0, PE1, PE2 in the rest of the sub-blocks. It also realizes the transfer function in (2). The critical period is Tm+2Ta, as shown in Fig. 10 using the Tree method to rearrange the adders. So this structure is faster than version 1. The structure can be extended to NxN as shown in Fig. 11. In order to keep the critical period at Tm+2Ta, the bottom sub-block needs to have PE order: (PE0t, PE2t, PE1t, PE1t, PE2t, PE1t, PE1t, PE2t…) while the non-bottom sub-blocks should have the same PE order as the bottom sub-block in Fig 8.

Both the separable denominator structures require fewer multipliers and delays than the original structure in [1] and the new transposed structure. Version 2 of the separable denominator structure is also fastest among all the structures discussed. It has slightly fewer delay elements compared to version 1. The complete comparison is shown in Table 1.

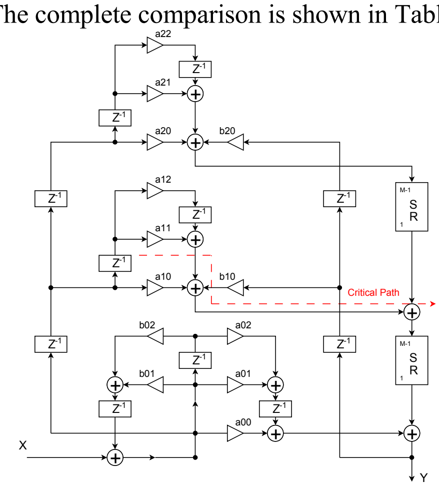

Figure 9. New separable denominator structure (version 2)

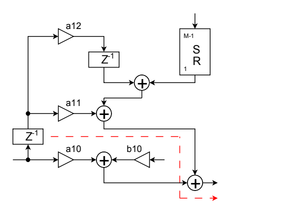

Figure 10. Critical path of separable denominator structure (version 2) using Tree Method

V. NEW DIAGONAL SYMMETRY STRUCTURE

The presence of symmetry in the 2-D frequency response can reduce the design complexity of a 2-D filter. Symmetry in the frequency response induces certain relationship among the filter coefficients which can result in fewer multipliers in the implementation. There are many types of symmetries [4]. Here, we will focus only on diagonal symmetry.

A 2-D magnitude response possesses diagonal symmetry if $\left|H(z_1,z_2)\right| = \left|H(z_2,z_1)\right|$ with $z_1 = e^{j\theta_1}$ and $z_2 = e^{j\theta_2}$, $\forall\ (\theta_1,\theta_2)$ [4]. Assuming the transfer function in (2) is used, this implies the following constraints on the coefficients: $a_{ij} = a_{ji}$ and $b_{0k} = b_{k0}$ for all i, j, k. Applying these constraints to the separable denominator structure (version 2), a new diagonal symmetry structure with separable denominator is obtained as shown in Fig. 12. This structure has the lowest number of multipliers compared to all the structures discussed so far. It also has a critical period of only Tm+2Ta as can be seen in Fig. 13. When compared to another diagonal symmetry structure in [5] (which has a general non-separable denominator), this new structure has fewer multipliers and a lower critical period. To realize higher order transfer functions, cascading can be used, i.e. the output of one 2x2 section is connected to the input of the next section with a delay added in between to control the critical period at the expense of latency. The structure, however, is not systolic. The complete comparison is shown in Table 1.

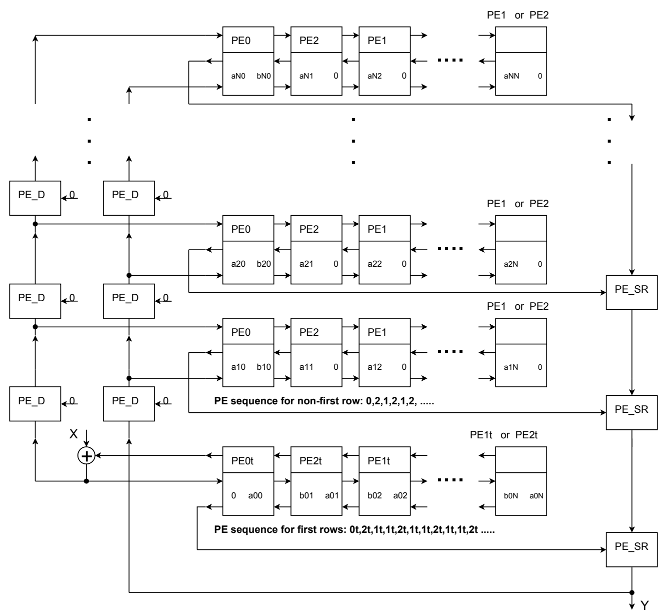

Figure 11. Separable denominator structure (version 2) -- NxN

VI. Conclusion

Four new 2-D filter VLSI architectures without global broadcast are presented. The first is a transposed systolic structure which requires fewer delay elements than [1]. By combining sub-blocks, two additional systolic architectures with structurally imposed separability of the transfer function denominator are obtained, which require fewer multipliers. One of them also has a lower critical period. A structure which possesses diagonal symmetry is then shown which requires even fewer multipliers.

References

[1] L. D. Van, "A new 2-D systolic digital filter architecture without global broadcast", *IEEE Trans. VLSI syst.*, vol. 10, no. 4, pp477-486, Aug. 2002.

[2] M. A. Sid-Ahmed, "A systolic realization for 2-D digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp560-565, Apr. 1989.

[3] S. Sunder, F. El-Guibaly, A. Antoniou, "Systolic implementations of two-dimensional recursive digital filters", *Proc. IEEE Int. Symp. Circuits Syst.*, pp1034-1037, May 1990.

[4] H. C. Reddy, I. H. Khoo, P. K. Rajan, "2-D symmetry: theory and filter design applications", *IEEE Circuits and Systems Magazine*, vol. 3, pp4-33, 2003.

[5] P.Y. Chen, L.D. Van, H.C. Reddy, C.T. Lin, "A new VLSI 2-D diagonal-symmetry filter architecture design", to appear in *Proc. IEEE APCCAS*, Macao, China, Dec 2008.

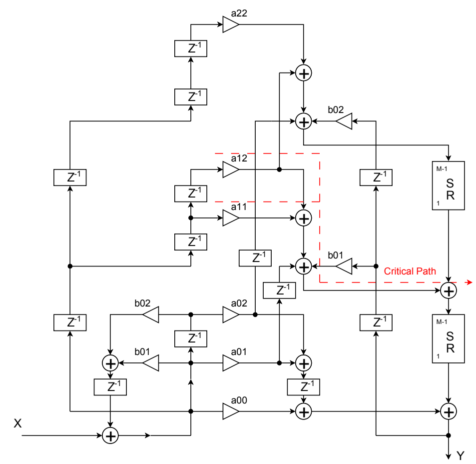

Figure 12. New diagonal symmetry structure with separable denominator

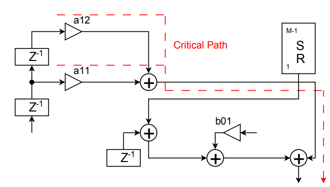

Figure 13. Critical path of diagonal symmetry structure using Tree Method

TABLE I. Comparison of different IIR filter architectures

Parameter	Sid-Ahmed [2]	Van [1]	Chen [5] (diagonal sym.)*	Transposed	Separable denominator V1	Separable denominator V2	Diagonal sym. with separable denominator*
# of multipliers	$2(N+1)^2-1$	$2(N+1)^2-1$	$11N/2$	$2(N+1)^2-1$	$(N+1)^2+2N$	$(N+1)^2+2N$	$10N/2$
Critical period	$T_m+(3+\lfloor\log_2 N\rfloor)T_a$	T_m+3T_a	T_m+3T_a	T_m+3T_a	T_m+3T_a	T_m+2T_a	T_m+2T_a
Latency	1	0	$N/2-1$	0	0	0	$N/2-1$
Global broadcast	Yes	No	No	No	No	No	No
# of delay elements	$(N+1)^2+2MN$	$\left\lfloor\frac{3N+1}{2}\right\rfloor(N+1)+N(M+1)$	$(11+2M)N/2+N/2-1$	$(N+1)N+1+\left\lfloor\frac{N-1}{7}\right\rfloor+\left\lfloor\frac{N}{7}\right\rfloor N+N(M+1)$	$N^2+\left\lfloor\frac{3N+1}{2}\right\rfloor+N(M+1)$	$N^2+N+1+\left\lfloor\frac{N-1}{3}\right\rfloor+N(M+1)$	$(11+2M)N/2+N/2-1$

*Assume N is even. Note: $\lfloor\cdot\rfloor$ denotes rounding down to the nearest integer.