

# A MULTI-LAYER 2-D ADAPTIVE FILTERING ARCHITECTURE BASED ON McCLELLAN TRANSFORMATION \*

K. J. Ray Liu and An-Yeu Wu

Electrical Engineering Department and Institute for Systems Research  
University of Maryland, College Park, MD 20742, USA

## ABSTRACT

A fully pipelined systolic array structure for multidimensional adaptive filtering is proposed. It utilizes the well-known McClellan Transformation (MT) to reduce the total number of parameters used in the 2-D filter. A new multi-layer triangular array, which is based on QR-decomposition RLS (QRD-RLS) as well as the projection method, is derived for the "1-D prototype filter" of MT. The hardware complexity for the new architecture is only  $O(N)$ . The system latency is also reduced from  $O(N)$  to  $O(\log_2 N)$ . Because of the fast convergence rate of the QRD-RLS algorithms, it is suitable for real-time image processing such as video signal processing.

## 1 INTRODUCTION

Multidimensional adaptive filtering has been an active area of research recently. Extension to 2-D applications of the well-known 1-D LMS and RLS algorithms have been reported in the literature [1, 2, 3]. However, the slow convergence rate of the LMS-type algorithms as well as the high computational complexity of the RLS algorithms become major drawbacks when both approaches are applied to real-time application such as video signal processing.

Recently, a concurrent algorithm and systolic architecture for multidimensional adaptive filtering was proposed by Shapiro and Staelin [4]. They employed McClellan transformation (MT) [5] as an parameterization method coupled with a QR factorization systolic array to perform adaptive luminance-chrominance separation. However, there are two disadvantages in their scheme: 1) Since Gentleman and Kung's triangular array (triarray) [6] is used, due to the opposite data waveform, the whole system will be "idle" during the operation of back-substitution, 2) The hardware complexity is  $O(N^2)$  which is not feasible for VLSI implementation as the order of the system increases. In this paper, we propose a modified systolic architecture which is based on McWhirter's triarray [7]. It eliminates the need of back-substitution operation, thus, the whole system can be processed in a fully pipelined fashion. In addition, a multi-layer triarray is also derived by using the projection

method. Since the system can be expanded recursively in a "divide-and-conquer" way, only  $O(N)$  hardware is required.

## 2 SYSTOLIC IMPLEMENTATION OF McCLELLAN TRANSFORMATION

Given a 1-D zero-phase FIR filter with support  $-N \leq n \leq N$ , the frequency response can be written as

$$H(\omega) = \sum_{n=0}^N a(n) \cos(\omega n). \quad (1)$$

Since  $\cos(\omega n)$  can be expressed as a Chebyshev polynomial of degree  $n$ ,  $T_n[\cdot]$ , we can rewrite (1) as

$$H(\omega) = \sum_{n=0}^N a(n) T_n[\cos \omega]. \quad (2)$$

By using the transformation of variables [5]

$$F(\omega_1, \omega_2) \rightarrow \cos \omega, \quad (3)$$

we obtain the MT 2-D frequency response

$$H(\omega_1, \omega_2) = \sum_{n=0}^N a(n) T_n[F(\omega_1, \omega_2)]. \quad (4)$$

The MT is a near-optimal design method for 2-D filters. It decomposes the problem into the design of a "1-D prototype" filter,  $a(n)$ , and the "2-D transformation function",  $F(\omega_1, \omega_2)$ . The former defines the frequency response along the 2-D frequency plane. The latter, which is a small 2-D zero-phase FIR filter, maps the 1-D frequencies into contours in the 2-D frequency plane. Therefore, by updating the coefficients of both the 2-D transformation filter and the 1-D prototype filter in an adaptive way, we can perform 2-D adaptive filtering efficiently. Fig.1 shows the block diagram of the systolic MT based on the Chebyshev recursion [8]. Each PE is a linear systolic array which performs the 2-D transformation.  $X_i[n_1, n_2], i = 1, 2, \dots, N$  are the outputs of the  $i$ th PE and  $y[n_1, n_2]$  is the desired 2-D signal.  $W = [w_1, w_2, \dots, w_N]$  is the tap coefficient vector of the 1-D FIR filter. In [4], to solve  $W$ ,  $X_i[n_1, n_2]$ 's are first sent to Gentleman and Kung's triarray to perform QR decomposition based on Givens rotation, and then  $W$  is solved through

\*The work is partially supported by the NSF grant ECD-8803012.

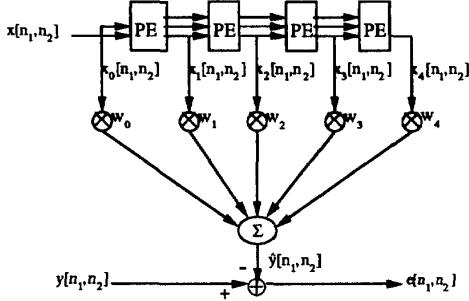


Figure 1: Systolic implementation for McClellan Transform.

a back-substitution linear systolic array. Because the system will become “idle” due to the opposite data flows while solving  $W$ , it is not appropriate for continuous operation, i.e., not applicable for real-time processing.

In some applications such as prediction error filters, adaptive noise cancellers and adaptive beamformers, it is not necessary to compute the least-squares (LS) weighting vector,  $W$ , explicitly. A modified triarray proposed by McWhirter [7], avoids the need to solve  $W$  and generates residuals directly, so it is more suitable for those applications. Similarly, in image restoration, image registration, etc., the estimation error

$$e[n_1, n_2] = y[n_1, n_2] - \hat{y}[n_1, n_2]$$

is more important than the weighting coefficients. Once we calculate the residual,  $e[n_1, n_2]$ , the filtered-out image can be easily obtained by subtracting the estimated error from the desired signal

$$\hat{y}[n_1, n_2] = \sum_i w_i X_i[n_1, n_2] = y[n_1, n_2] - e[n_1, n_2]. \quad (5)$$

Because the McWhirter’s triarray can produce the residuals in real time, the LS error  $e[n_1, n_2]$  can be generated in a fully-pipelined way without any back-substitution operation and with better numerical properties. So now we can modify the MT systolic structure in [4] by employing only one triangular systolic array.

### 3 MULTI-LAYER TRIARRAY STRUCTURE

The two Givens systolic arrays described in [6, 7] do not impose any restriction on the input data structure. As a consequence of this generality, their computational complexity is proportional to  $N^2$ . Recently, some “fast” LS algorithms based on Givens rotation have been proposed by several authors [9, 10, 11]. By exploiting the Toeplitz structure of the input data covariance matrix, a fast QR algorithm for LS estimation can be implemented on a lattice structure with  $O(N)$  complexity. However, for many applications like adaptive sidelobe canceller (SLC), and the image signals  $X_i[n_1, n_2]$ ’s in Fig.1, the fast QR algorithm cannot be used

because no special structure can be exploited. Therefore, we are motivated to derive a reduced-size structure which can achieve a similar performance like full-size QRD-RLS array (FULL-QRD), but only with a computational complexity of  $O(N)$ .

Let us consider a LS problem with only two column vectors

$$AW \approx y$$

where  $A = [a_1, a_2]$ ,  $W = [w_1, w_2]^T$ . The LS optimal weighing coefficient  $\hat{W}$  can be solved by the normal equations:

$$\hat{W} = (A^T A)^{-1} A^T y, \quad \hat{y} = A(A^T A)^{-1} A^T y = A\hat{W}, \quad \hat{e} = y - \hat{y}, \quad (6)$$

where  $\hat{y}$  is the optimal estimation of  $y$  and  $\hat{e}$  is the optimal residual. In the geometric point of view,  $\hat{y}$  is the projection of  $y$  on the subspace spanned by  $a_1, a_2$ , denoted as  $\text{span}\{a_1, a_2\}$ , and the error vector  $\hat{e}$  is perpendicular to that plane. Now instead of projecting  $y$  onto  $\text{span}\{a_1, a_2\}$ , we project  $y$  onto the two smaller subspaces,  $\text{span}\{a_1\}$  and  $\text{span}\{a_2\}$ , and the optimal projection are  $\hat{y}_1, \hat{y}_2$  respectively as shown in Fig.2.

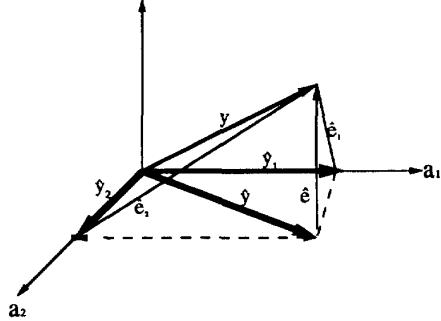


Figure 2: Geometric interpretation for reduced-size LS estimation

Since  $\hat{y}_1 // a_1, \hat{y}_2 // a_2$  in this case, one can easily verify that  $\text{span}\{\hat{y}_1, \hat{y}_2\} = \text{span}\{a_1, a_2\}$ . Thus the optimal projection of  $y$  onto  $\text{span}\{a_1, a_2\}$  can be found by projecting  $y$  onto the new subspace  $\text{span}\{\hat{y}_1, \hat{y}_2\}$ . If now  $A = [A_1, A_2]$  and  $A_1, A_2 \in R^{n \times m/2}$ , we can first obtain  $\hat{y}_1, \hat{y}_2$  by projecting  $y$  onto the two subspaces  $\text{span}\{A_1\}$ ,  $\text{span}\{A_2\}$ . The two subspace projections  $\hat{y}_1$  and  $\hat{y}_2$  can be obtained from (6) as

$$\hat{y}_1 = y - \hat{e}_1, \quad \hat{y}_2 = y - \hat{e}_2. \quad (7)$$

Then we try to estimate  $\hat{y}$  by projecting  $y$  onto the newly constructed small subspace  $\text{span}\{\hat{y}_1, \hat{y}_2\}$ . Through this approach, we can save about 50% hardware to obtain the estimation of  $\hat{e}$ , say  $\hat{e}_{\text{approx}}$ , if  $N$  is large. The corresponding systolic array to implement the above argument is illustrated in Fig.3(a). If we further expand the two half full triarrays on the top of Fig.3(a) recursively, we will get a binary tree-like structure in Fig.3(b). It is easy to verify that this new structure has hardware complexity proportional to  $O(N)$  and a total delay  $O(\log_2 N)$ . Because of its resemblance to a binary tree, we call the expanded architecture as the “multi-layer QRD-RLS array” (ML-QRD).

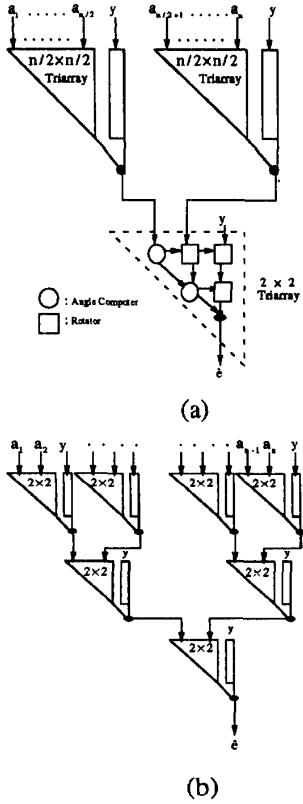


Figure 3: (a) QRD array for two subspaces (b) Multi-layer QRD array

Fig.4 shows the experimental learning curve for both FULL-QRD and the ML-QRD for an 8th order AR process. It is interesting to see that ML-QRD has a faster convergence rate than FULL-QRD since it takes more time to initialize a  $8 \times 8$  tri-array than three cascaded  $2 \times 2$  tri-array. We can also find that the misadjustment is less than 1% MSE which is acceptable compared to the popular LMS-type algorithms. A comparison of hardware cost for FULL-QRD and ML-QRD can be found in Table 1<sup>1</sup>. We can obtain significant savings in multipliers while  $N$  is increasing.

|          | Angle Computer | Rotator    | System latency |
|----------|----------------|------------|----------------|
| FULL-QRD | $N$            | $N(N+1)/2$ | $N$            |
| ML-QRD   | $2(N-1)$       | $3(N-1)$   | $\log_2 N$     |

Table 1: Comparison of hardware cost for FULL-QRD and ML-QRD.

<sup>1</sup>Each Angle Computer and Rotator consists of about 4 multipliers [10].

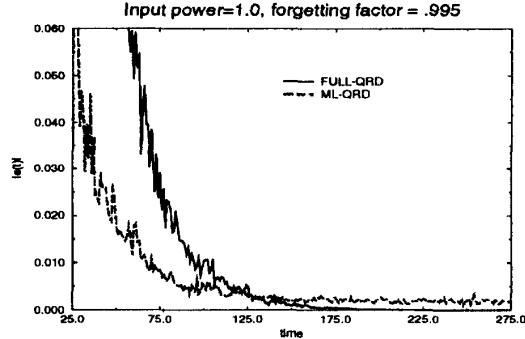


Figure 4: Learning curve of FULL-QRD and ML-QRD for 8th order AR process.

#### 4 SIMULATION RESULTS

Now combining the systolic MT structure and the multi-layer triarray, we can perform multidimensional adaptive filtering with  $O(N)$  hardware complexity and throughput rate equal to one. The performance of the proposed scheme is examined by applying it to a two-dimensional adaptive line enhancer (TDALE) which has been used in [1, 3]. The block diagram is depicted in Fig.5.

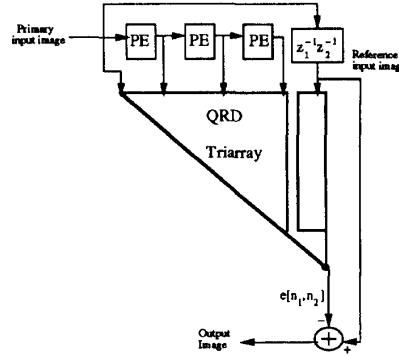


Figure 5: Multidimensional adaptive filter for TDALE

The primary input is the well-known "LENA" image degraded by a white Gaussian noise. A 2-D unit delay  $z_1^{-1} z_2^{-1}$  is used as a decorrelation operator to obtain the reference input from the primary image. The image signal is fed into the system in the raster scanned style, i.e., from left to right from top to bottom. After the input image goes through the triarray, the generated estimated error is subtracted from the reference signal. Thus we can get the filtered image in real-time. The system is fully-pipelined with throughput rate equal to one. Neither statistic operation nor error feedback is used in this implementation.

Some simulation results are shown in Table 2 and the processed images with input SNR equal to 3.7 dB (noise

|                           |      |     |     |
|---------------------------|------|-----|-----|
| Input SNR (dB)            | 10.0 | 3.0 | 0.0 |
| Output SNR in [3]         | 12.0 | 8.0 | 6.0 |
| Output SNR using FULL-QRD | 10.5 | 9.0 | 7.6 |
| Output SNR using ML-QRD   | 10.9 | 9.8 | 8.7 |

Table 2: SNR results for the proposed systolic multidimensional adaptive filtering in the application of restoring noisy images.

variance =1000) are shown in Fig.6. From Table 2, we can find that the performance of our scheme is about the same as the 2-D joint process lattice structure in [3]. It is also interesting to see that the ML-QRD performs even better than the FULL-QRD. As we discussed in the previous section, although the ML-QRD has worse misadjustment, it has faster convergence rate than that of the FULL-QRD. This property makes it preferable in time-varying and nonstationary processing.

## 5 CONCLUSION

In this paper, we derive the multi-layer systolic array for LS estimation with hardware complexity of  $O(N)$ . Its application in real-time multidimensional adaptive filtering is also presented. Because of the superior computational robustness and the fast convergence rate for the RLS algorithm, it is a promising candidate for real-time multidimensional application such as video image restoration, image signal channel equalization and image filtering.

## References

- [1] M. M. Hadhoud and D. W. Thomas, "The two-dimensional adaptive LMS (TDLMS) algorithm," *IEEE Trans. Circuits Syst.*, vol. 5, pp. 485–494, May 1988.
- [2] A. M. Sequeira and C. W. Therrien, "A new 2-D fast RLS algorithm," in *Proc. IEEE ICASSP*, pp. 1401–1404, 1990. Albuquerque.
- [3] H. Youlal, Malika Janati-I, and M. Najim, "Two-dimensional joint process lattice for adaptive restoration of images," *IEEE Trans. Image Processing*, vol. 1, pp. 366–378, July 1992.
- [4] J. M. Shapiro and D. H. Staelin, "Algorithms and systolic architecture for multidimensional adaptive filtering via McClellan transformation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 2, pp. 60–71, Mar 1992.
- [5] R. M. Mersereau, W. F. G. Mecklenbrauker, and T. F. Quatieri, Jr., "McClellan transformations for two-dimensional digital filtering: I - Design," *IEEE Trans. Circuits Syst.*, vol. 23, pp. 405–422, July 1976.
- [6] W. M. Gentleman and H. T. Kung, "Matrix triangulation by systolic arrays," *Proc. SPIE, Real-Time Signal Processing IV*, vol. 298, pp. 298–303, 1981.
- [7] J. G. McWhirter, "Recursive least-squares minimization using a systolic array," *Proc. SPIE, Real-Time Signal Processing VI*, vol. 431, pp. 105–112, 1983.
- [8] J. H. McClellan and D. S. K. Chan, "A 2-D FIR filter structure derived from the Chebyshev recursion," *IEEE Trans. Circuits Syst.*, vol. 24, pp. 372–378, July 1977.
- [9] J. M. Cioffi, "The fast adaptive ROTOR's RLS algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 631–653, April 1990.
- [10] F. Ling, "Givens rotation based least squares lattice and related algorithms," *Trans. Signal Processing*, vol. 39, pp. 1541–1551, July 1991.
- [11] I. K. Proulder, J. G. McWhirter, and T. J. Shepherd, "Computationally efficient QR decomposition approach to least squares adaptive filtering," *IEE Proceedings-F*, vol. 138, pp. 341–353, Aug. 1991.

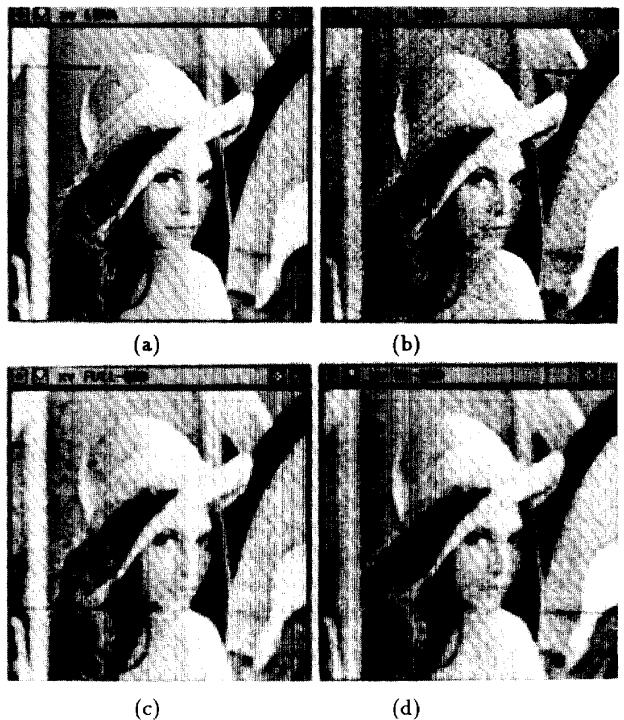


Figure 6: (a) Original LENA image. (b) Noisy input image with SNR=3.7 dB (variance = 1000). (c) Output of TDALE with full array (SNR=9.2 dB). (d) Output of TDALE with multi-layer array (SNR=10.0 dB).