A Class of Fast Cyclic Convolution Algorithms Based on Block Pseudocirculants

Marvi Teixeira and Domingo Rodriguez

Abstract—**Pseudocirculant matrices have been studied in the past in the context of FIR filtering, block filtering, polyphase networks and others. For completeness, their relation to cyclic convolution, stride permutations, circulant matrices, and to certain permutations of the Fourier matrix is explicitly established in this work. Within this process, a class of highly regular fast cyclic convolution algorithms, based on block pseudocirculant matrices, is obtained.**

I. Introduction

THE availability of multiprocessor architectures makes desirable the development of algorithms that break a large block cyclic convolution into smaller cyclic convolutions within an appropriate multiprocessor structure. Not only is the computational complexity to be minimized, but most importantly, the algorithm structure is sought to match the underlying target architecture. In the past, this problem has been addressed, among others, by Agarwal and Burrus [1] and Pitassi [2]. More recently, certain multirate structures related to pseudocirculant matrices have been studied in the context of FIR filtering, block filtering, and others [3], [4]. For completeness, we are explicitly establishing the relation of those pseudocirculant matrices to cyclic convolution, stride permutations, circulant matrices, and to certain permutations of the Fourier matrix. Within this process, a highly regular cyclic convolution algorithm, which is suitable for VLSI and parallel implementation, is obtained.

II. Description of the Algorithm

We start by considering the cyclic convolution of $x[n]$ and $h[n]$, both of length N sequences, which is given by

$$y[n] = \sum_{k=0}^{k=N-1} x[k]h[n-k]. \tag{1}$$

The cyclic convolution theorem allows us to write

$$Y[k] = X[k]H[k] \tag{2}$$

where

$$Y[k] = DFT\{y[n]\}, \quad X[k] = DFT\{x[n]\} \quad \text{and}$$
$$H[k] = DFT\{h[n]\}.$$

Manuscript received January 20, 1995; revised February 22, 1995. The associate editor coordinating the review of this paper and approving it for publication was Prof. A. E. Yagle.

M. Teixeira is with the Electrical Engineering Department, Polytechnic University of Puerto Rico, San Juan, PR 00919 USA.

D. Rodriguez is with the Electrical Engineering Department, University of Puerto Rico, Mayagüez, PR 00681-5000 USA.

IEEE Log Number 9411211.

Provided that r is a factor of N, we can perform a radix r decimation in time on each of the factors in (2), obtaining

$$\sum_{l=0}^{r-1}\sum_{n=0}^{N/r-1} y[rn+l]W_N^{(rn+l)} = \left(\sum_{l=0}^{r-1}\sum_{n=0}^{N/r-1} x[rn+l]W_N^{(rn+l)}\right) \cdot \left(\sum_{l=0}^{r-1}\sum_{n=0}^{N/r-1} h[rn+l]W_N^{(rn+l)}\right). \tag{3}$$

This can also be written in matrix form [5], in which case, we would have a multiplication of matrix polynomials (in W). After performing the indicated polynomial multiplication, we can compare both sides through the indeterminate coefficients method. Writing the resulting equations in matrix form [5] gives (4), which appears at the bottom of the next page, which is the product of a block pseudocirculant matrix H_p times a vector $[X_i]$. The vectors entries X_i and Y_i are decimated sections (of length N/r) of the original sequences $x[n]$ and $y[n]$. The block pseudocirculant matrix of size r by r has as its entries the circulant matrices H_i and, over the main diagonal, the circulant matrices multiplied by cyclic shift operators $S_{N/r}H_i$. All entries are of size $N/r \times N/r$. The cyclic shift operator $S_{N/r}$ can be written in matrix form as

$$S_{N/r} = \begin{bmatrix} 0 & 0 & 0 & \cdot & 0 & 1 \\ 1 & 0 & 0 & \cdot & 0 & 0 \\ 0 & 1 & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & 1 & 0 \end{bmatrix}. \tag{5}$$

The whole process involves decimating the sequences by r and using the subsequences to perform shorter cyclic convolutions that are followed, when appropriate, by cyclic shifts and, finally, a reconstruction stage. A direct realization of the algorithm, which can be easily achieved looking at its matrix formulation, would take r^2 sections or subconvolutions. This type of direct realization is far from optimal (in terms of number of sections) and should be used only if it matches the underlying multiprocessor architecture. A reduction in the number of sections, however, can be achieved through a factorization of the pseudocirculant matrices. The cases $r = 2$ and $r = 3$ are illustrated next.

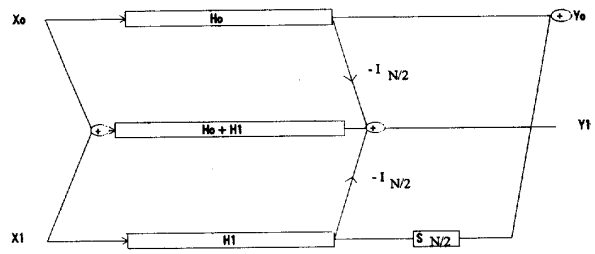

Fig. 1. N point cyclic convolution; decimation by 2.

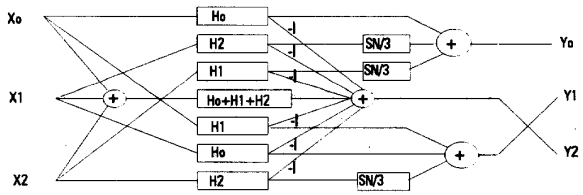

Fig. 2. N point cyclic convolution; decimation by 3.

For $r = 2$, we have

$$\begin{bmatrix} Y_0 \\ Y_1 \end{bmatrix} = \begin{bmatrix} H_0 & S_{N/2}H_1 \\ H_1 & H_0 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \end{bmatrix} \tag{6}$$

which can be decomposed using a technique illustrated in [5] as

$$\begin{bmatrix} Y_0 \\ Y_1 \end{bmatrix} = \begin{bmatrix} I_{N/2} & 0 & S_{N/2} \\ -I_{N/2} & I_{N/2} & -I_{N/2} \end{bmatrix} \cdot \begin{bmatrix} H_0 & 0 & 0 \\ 0 & H_0 + H_1 & 0 \\ 0 & 0 & H_1 \end{bmatrix} \cdot \begin{bmatrix} I_{N/2} & 0 \\ I_{N/2} & I_{N/2} \\ 0 & I_{N/2} \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \end{bmatrix}. \tag{7}$$

The corresponding block diagram realization can be found in Fig. 1.

For $r = 3$, we have

$$\begin{bmatrix} Y_0 \\ Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} H_0 & S_{N/3}H_2 & S_{N/3}H_1 \\ H_1 & H_0 & S_{N/3}H_2 \\ H_2 & H_1 & H_0 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix}. \tag{8}$$

Similarly, we have (9), which appears at the top of the next page. The same technique can be used for a higher radix. The block diagram realization can be found in Fig. 2.

This factorization yields $r(r-1)+1$ sections, which is an improvement with respect to the direct realization and still within a highly regular structure. The cyclic subconvolutions in each section can be decimated, at the previous or different rate, depending on their length and our overall implementation strategy. A further reduction in the number of sections can be achieved by looking at (3) as a polynomial multiplication and by using the Chinese remainder theorem, as it is done in [6]. This is at the expense of increasing the structural complexity.

III. Block Pseudocirculant Matrices and Their Relation to Stride Permutations, Circulant Matrices, and the Fourier Matrix

The matrix formulation of (3) can be achieved using the Fourier matrix F_N, the stride permutation matrices, $P_{N,r}$, and the sequences $x[n], y[n]$ written in vector form [7]:

$$F_N P_{n,r}^{-1} P_{N,r} y = (F_N P_{N,r}^{-1} P_{N,r} x) \cdot (F_N P_{N,r}^{-1} P_{N,r} h) \tag{10}$$

where $P_{N,r}x$ and $P_{N,r}y$ are the input and output vectors $[Xi], [Yi]$ found in (4) that are decimated by r, and $F_N P_{N,r}^{-1}$ is a permutation of the Fourier matrix. From (10), it follows that [5]

$$P_{N,r} y = P_{N,r} H_N P_{N,r}^{-1} P_{N,r} x. \tag{11}$$

The previous expression, however, is identical to (4), allowing us to write

$$H_p = P_{N,r} H_N P_{N,r}^{-1} \tag{12}$$

which effectively relates the block pseudocirculant matrices H_p obtained within a decimation scheme in the context of cyclic convolution to the circulant matrices H_N and to the stride permutation matrices $P_{N,r}$. It can be further shown that these pseudocirculants matrices are diagonalized by $F_N P_{N,r}^{-1}$ [5].

IV. Relation to Previous Work

The case $r = 2$ was studied in [5] for sequences of length a power of two and shown to belong to a class of algorithms related to the Walsh transform. Neither a generalization nor a relation to the pseudocirculant matrices were provided at that time. Letting $W_N^{kn} \rightarrow z^{-n}, S \rightarrow z^{-1}$, we establish a correspondence that partially relates this work with those found in [3], [4], [6], and others, which, in the context of FIR filtering, have also exposed these underlying multirate structures based on block pseudocirculants.

$$\begin{bmatrix} Y_0 \\ Y_1 \\ Y_2 \\ \cdot \\ Y_{r-2} \\ Y_{r-1} \end{bmatrix} \begin{bmatrix} H_0 & S_{N/r}H_{r-1} & S_{N/r}H_{r-2} & \cdot & S_{N/r}H_2 & S_{N/r}H_1 \\ H_1 & H_0 & S_{N/r}H_{r-1} & \cdot & S_{N/r}H_3 & S_{N/r}H_2 \\ H_2 & H_1 & H_0 & \cdot & S_{N/r}H_4 & S_{N/r}H_3 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ H_{r-2} & H_{r-3} & H_{r-4} & \cdot & H_0 & S_{N/r}H_{r-1} \\ H_{r-1} & H_{r-2} & H_{r-3} & \cdot & H_1 & H_0 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ \cdot \\ X_{r-2} \\ X_{r-1} \end{bmatrix} \tag{4}$$

$$= \begin{bmatrix} I_{N/3} & S_{N/3} & S_{N/3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{N/3} & I_{N/3} & S_{N/3} \\ -I_{N/3} & -I_{N/3} & -I_{N/3} & I_{N/3} & -I_{N/3} & -I_{N/3} & -I_{N-3} \end{bmatrix} \begin{bmatrix} H_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & H_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & H_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & H_0+H_1+H_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & H_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & H_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & H_2 \end{bmatrix}$$

$$\cdot \begin{bmatrix} I_{N/3} & 0 & 0 \\ 0 & I_{N/3} & 0 \\ 0 & 0 & I_{N/3} \\ I_{N/3} & I_{N/3} & I_{N/3} \\ I_{N/3} & 0 & 0 \\ 0 & I_{N/3} & 0 \\ 0 & 0 & I_{N/3} \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix} \tag{9}$$

V. Conclusions

Block pseudocirculant matrices were studied in the context of cyclic convolution, including their relation to stride permutations of the Fourier and circulant matrices. Within this process, a highly regular class of fast algorithms that are suitable to handle large size cyclic convolutions through multiprocessor or VLSI implementation, was derived.

References

[1] R. C. Agarwal and C. S. Burrus, "Fast one dimensional digital convolution by multidimensional techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, no. 1, pp. 1–10, Feb. 1974.

[2] D. A. Pitassi, "Fast convolution using the Walsh transform," in *Proc. Conf. Applications Walsh Functions*, Washington, DC, Apr. 1971, pp. 130–133.

[3] Z. Mou and P. Duhamel, "Short-length FIR filters and their use in fast nonrecursive filtering," *IEEE Trans. Signal Processing*, vol. 39, no. 6, pp. 1322–1332, June 1991.

[4] P. P. Vaidyanathan and S. Mitra, "Polyphase networks, block digital filtering, LPTV systems, and alias-free QMF banks: A unified approach based on pseudocirculants," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 3, pp. 381–391, Mar. 1988.

[5] M. Teixeira and D. Rodriguez, "A new method mathematically links fast Fourier transform algorithms with fast cyclic convolution algorithms," in *Proc. 37th Midwest Symp. Circuit Syst.*, Lafayette, LA, Aug. 1994.

[6] Z. J. Mou and P. Duhamel, "A unified approach to the fast FIR filtering algorithms," in *Proc. ICASSP '88*, 1988, pp. 1914–1917.

[7] R. Tolimieri and C. Lu, *Algorithms for Discrete Fourier Transform and Convolution*. New York: Springer-Verlag, 1989.