A VLSI ARCHITECTURE FOR DIGITAL FILTERS USING COMPLEX NUMBER-THEORETIC TRANSFORMS*

I.S. Reed, C.-S. Yeh

Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089-0272

T.K. Truong

Communication System Research
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91103

ABSTRACT

In this paper a parallel architecture is developed to realize a digital filter. First a systolic array is used to compute a 248-point complex number-theoretic transform (CNT). Next an algorithm is developed to realize a digital filter that uses 248-point CNT's and a generalization of the overlap-save method. This algorithm solves the conflict between long transform lengths and a wide dynamic range associated with the number-theoretic transform. Finally this algorithm is mapped to a parallel architecture. This architecture is simple, regular and expandable, and, hence, is suitable for VLSI implementation.

INTRODUCTION

The digital complex number convolution is a useful tool for many digital processing applications [1]. Some important applications are the processing of spaceborne high-resolution synthetic aperture radar (SAR), image processing, communication modems, speech processing.

Number-theoretic transforms were first defined by Radar [3]. He defined the Mersenne prime transform (MPT) and the Fermat number transform (FNT) [3]. Reed and Truong [4] extended the MPT to a complex number-theoretic transform (CNT) by taking transforms over the Galois field $GF(q^2)$ of q^2 elements, where $q = 2^P-1$ is a Mersenne prime for P = 2,3,5,7,13,17,19,31,61,... This Galois field $GF(q^2)$ is analogous to the field of complex numbers. Recently, Nussbaumer [5] observed that $(1+\hat{i})$ is an element of order 8P in $GF(q^2)$ where $\hat{i}$ is a root of irreducible polynomial $p(x)=x^2+1$ over GF(q).

A PIPELINE STRUCTURE FOR COMPUTING A 248-POINT CNT

Let $GF(q^2)$ be a Galois field, where $q=2^P-1$ is a Mersenne prime and let the integer d divide q^2-1. Also let the element $\gamma \varepsilon GF(q^2)$ generate the cyclic subgroup of d elements, $G_d=\{\gamma,\gamma^2,\ldots,\gamma^{d-1},\gamma^d=1\}$, in the multiplicative group of $GF(q^2)$. The CNT transform pair between two d-point sequences $\{a_n\}$ and $\{A_k\}$ over G_d is defined as follows:

$$A_k = \sum_{n=0}^{d-1} a_n \gamma^{kn} \tag{1a}$$

$$a_n = (d)^{-1} \sum_{k=0}^{d-1} A_k \gamma^{-kn} \tag{1b}$$

where $0 \le n, k \le d-1$, and $a_n, A_k \varepsilon GF(q^2)$. In eq. (1) and the following discussion all values are taken modulo q. Also all exponents are assumed to be taken modulo d, since $\gamma^d=1$. It is shown [4] that the cyclic convolution of two sequence is obtained by taking the inverse CNT of the product of the CNT's of these two sequences.

It is shown [5] that $(1+\hat{i})$ is an element of order 8P=248 in $GF(q^2)$ for $q=2^{31}-1$. That is, d=248 is least positive integer such that $(1+\hat{i})^d \equiv 1 \pmod q$. In the following the q,γ and d in eq. (1) are chosen to be $2^{31}-1$, $(1+\hat{i})$ and 248, respectively. From eq. (1) one observes that the operations needed to compute a 248-point CNT with $\gamma=(1+\hat{i})$ require only bit rotations and additions.

Note that $\gamma^{-124}=\gamma^{124}=-1$ and $\gamma^{248}=1$. Eq. (1a), with d=248, can be rewritten as follows:

$$\begin{aligned} A_k &= \sum_{n=0}^{123} (a_n + a_{n+124}\gamma^{124k})\gamma^{nk} \\ &= \sum_{n=0}^{123} y_n \gamma^{nk} = \sum_{n=0}^{123} y_n \gamma^{-(124-n)k} \gamma^{124k} \end{aligned} \tag{2}$$

where $\gamma=1+\hat{i}$ and $y_n=a_n+a_{n+124}\gamma^{124k}=a_n+(-k)^k a_{n+124}$. Since $\gamma^{-(k+124)}=\gamma^{-k}\gamma^{-124}=-\gamma^{-k}$, Eq. (2) can be expressed in a recursive formula as follows:

$$A_k = ((\ldots(y_0\gamma^{-k}+y_1)\gamma^{-k}+y_2)\gamma^{-k}+\ldots)\gamma^{-k}+y_{123})\gamma^{-k}(-1)^k \tag{3a}$$

$$A_{k+124} = ((\ldots(y_0(-\gamma^{-k})+y_1)(-\gamma^{-k})+\ldots)(-\gamma^{-k}) + y_{123})(-\gamma^{-k})(-1)^k \tag{3b}$$

ICASSP 83, BOSTON CH1841-6/83/0000-0923 $1.00 © 1983 IEEE

where $0 \le k \le 123$. Note that, in eq. (3), $\gamma^{-k} = -\gamma^{124}\gamma^{-k} = -\gamma^{124-k}$.

Fig. 1 shows a structure for computing a 248-point CNT. Similar structures were proposed in [6] for a different purpose. In Fig. 1 z^{-124} denotes a 124-step delay element. Input data $a_0, a_1 \ldots$ are fed successively into the system from data-in pins DIN. PT_k is a set of pass transistors commonly controlled by control signal $f_{(k)+1}$, where $0 \le k \le 247$ and (k) denotes the residue of k modulo 124. L_k for $0 \le k \le 123$ is a set of dual inner product cells.

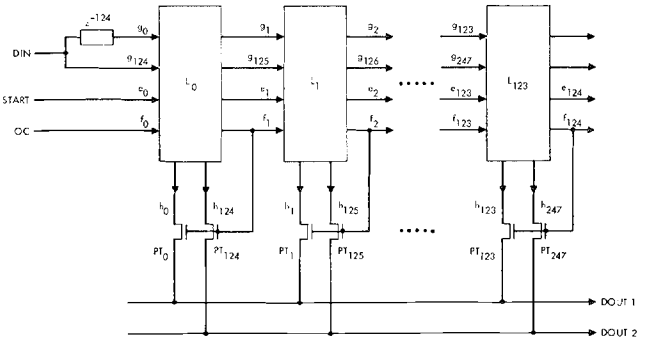

Figure 1. A Pipeline Structure for Computing a 248-point CNT

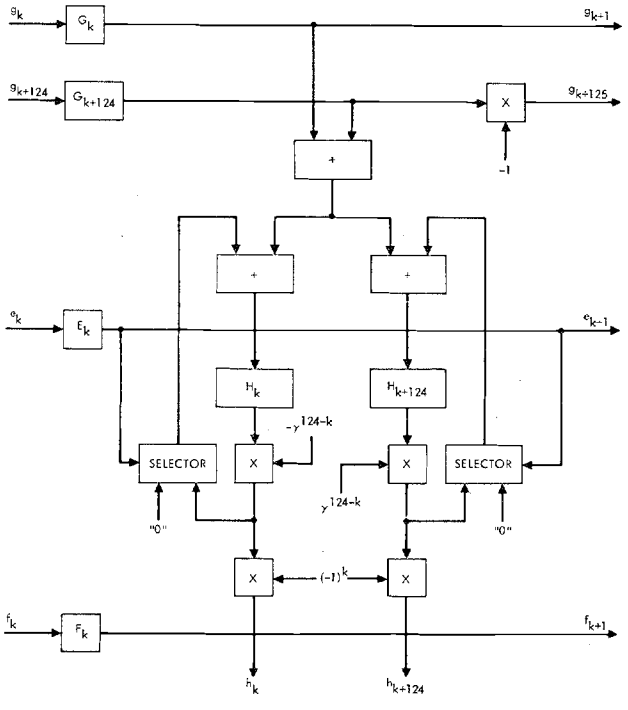

Figure 2. A Block Diagram of Dual Inner Product Cell L_k with $\gamma = 1 + \hat{i}$

Fig. 2 shows a block diagram of L_k that computes both A_k and A_{k+124}. Radar proposed a similar method to compute A_k in [3]. $(-1)^k a_{n+124}$ is generated recursively by $(-1)((-1)^{k-1}a_{n+124})$. In Fig. 2 registers G_k and G_{k+124} store a_n and $(-1)^k a_{n+124}$, respectively, for some n. The computations of A_k and A_{k+124} in L_k are evoked by control signal e_{k+1} obtained from flip-flop E_k. The control signal e_k was derived initially in the system from input pin START and propagated from cell to cell. The accumulated results of A_k and A_{k+124} are stored in registers H_k and H_{k+124} and are available at h_k and h_{k+124}, respectively. The h_k and h_{k+124} are transmitted out of the circuit from data-out pins DOUT1 and DOUT2, respectively, when control signal f_{k+1} in flip-flop F_k is set to 1. The control signal f_k comes from input-pin OC (Output Control). At any time instance at most one of f_i's is set to 1 for $0 \le i \le 123$.

All operations in Fig. 2 are taken modulo $q = 2^{31}-1$. $(-a_n)$ mod q is obtained simply by taking the 1's complement of a_n, where $q = 2^{31}-1$. Addition modulo q can be realized by the circuit in [3].

The multiplication of a_n by $(1+\hat{i})^k$ is computed as follows: Let $k = k_0 \cdot 2^0 + k_1 \cdot 2^1 + k_2 \cdot 2^2 + k_3 \cdot 2^3$, where $0 \le k_3 \le 15$ and $k_0, k_1, k_2 = 0$ or 1. Note that $(1+\hat{i})^2 = 2\hat{i}$, $(1+\hat{i})^4 = -2^2$, $(1+\hat{i})^8 = 2^4$. Therefore, $a_n(1+\hat{i})^k = a_n(1+\hat{i})^{k_0}\cdot(2\hat{i})^{k_1}\cdot(-2^2)^{k_2}\cdot(2^4)^{k_3} = a_n(1+\hat{i})^{k_0}\cdot(\hat{i})^{k_1}\cdot(-1)^{k_2}\cdot(-2)^{[k/2]}$, where $[k/2] = k_1 + k_2 \cdot 2^1 + k_3 \cdot 2^2$ is the largest integer less than or equal to k/2. The multiplication of an integer by a power of 2 modulo q can be realized readily by the standard barrel shifter circuit, given in [7].

A DIGITAL FILTER ARCHITECTURE USING THE CNT

The generalized overlap-save method [2] was developed to solve the conflict between long transform lengths and a wide dynamic range of the number-theoretic transform. In this section this method is used in conjunction with the CNT circuit in the previous section to realize a digital filter. The dynamic range [4] of the CNT in the previous section is $\sqrt{(q-1)/(2d)} \cong \sqrt{2} \cdot 2^{10}$. This value is sufficiently large for a number of applications.

Let $\{a_n\}$ and $\{b_m\}$ be the input data sequence and filter sequence of a finite impulse response (FIR) digital filter, respectively, where $0 \le n \le N-1$ and $0 \le m \le M-1$. The output sequence $\{c_k\}$ of the filter is the linear convolution of $\{a_n\}$ and $\{b_m\}$, where $0 \le k \le N+M-1$ [1]. The following algorithm outlines the generalized overlap-save method for computing $\{c_k\}$ using the CNT.

(1) Partition $\{b_m\}$ into subfilter $\{b_m^{(j)}\}$ for $j \ge 1$ with $b_m^{(j)} = b_m$ for $124(j-1) \le m \le 124j-1$, and $b_m^{(j)} = 0$ otherwise.

(2) Compute the linear convolution $\{c_k^{(j)}\}$ of $\{a_n\}$ and $\{b_m^{(j)}\}$ using the standard overlap-save method as follows:

(2.1) Section $\{a_n\}$ into 248-point subsequences $\{a_n^{(t)}\}$ for $t \ge 0$ with 124 points

of $\{a_n\}$ overlapped between two consecutive subsequences. That is, $a_n^{(t)}=a_n$ for $124(t-1)\leq n\leq 124(t+1)-1$, and $a_n^{(t)}=0$ otherwise, where $a_n=0$ for $n<0$.

(2.2) Compute the cyclic convoltuion $\{z_k^{(t)}\}$ of $\{a_n^{(t)}\}$ and $\{b_m^{(j)}\}$ by using the CNT.

(2.3) $\{c_k^{(j)}\}$ is obtained by discarding the first half and saving the second half of each convolution $\{z_k^{(t)}\}$ computed in (2.2).

(3) Finally, the desired output sequence $\{c_k\}$ equals the arithmetic sum of $\{c_k^{(j)}\}$ for $j\geq 1$.

Fig. 3 illustrates an example of the generalized overlap-save method for m=496. For simplicity the drawing in Fig. 3 shows the data as if it were continuous rather than digital. Also the results of the convolutions are not drawn accurately. Other cases of the generalized overlap-save method are constructed in a similar manner.

Fig. 4 shows an architecture for realizing the generalized overlap-save method for achieving the digital filter in Fig. 3. In Fig. 4 $\{B_k^{(j)}\}$ is the CNT of $\{b_m^{(j)}\}$ multiplied by the factor $(248)^{-1}$ in eq. (1b) for $1\leq i\leq 4$. The $B_k^{(j)}$'s can be precomputed and stored in the system. The products of the multipliers in Fig. 4 are taken modulo $q=2^{31}-1$. The adders in Fig. 4 perform normal binary additions rather than additions modulo q. The inverse CNT circuit used in Fig. 4 is a degenerative version of the structure shown in Fig. 1.

The advantages of the generalized overlap-save method for implementing a FIR digital filter using CNT transforms are the following: (1) No multiplications are required. Only additions and bit rotations are needed. (2) The usual dynamic range limitation for long sequence CNT's is alleviated. (3) The CNT and inverse CNT circuits are utilized 100% of the time. (4) The lengths of the input data and filter sequences can be arbitrary and different.

CONCLUSION

A pipeline structure is developed to compute a 248-point complex number-theoretic transform (CNT). The 248-point CNT requires only additions and bit rotations. The generalized overlap-save method is used in conjunction with this CNT to realize a FIR digital filter with arbitrary input data and filter sequence lengths. An architecture is developed to implement the generalized overlap-save method by a straightforward combination of one 248-point CNT and several 248-point inverse CNT structures. This technique for realizing a digital filter alleviates the dynamic range limitations of the CNT with a long transform length. The architecture of the circuit is simple and regular, and hence suitable for VLSI implementation.

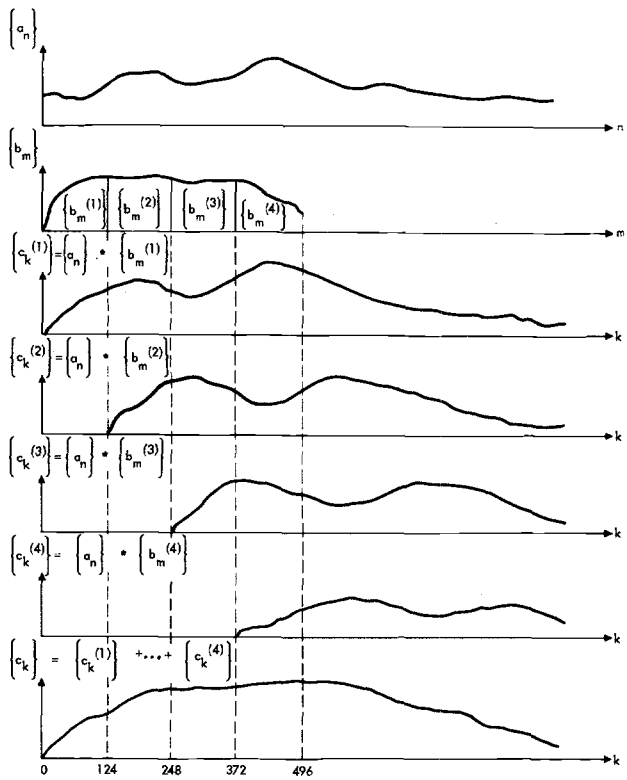

Figure 3. An Example of the Generalized Overlap-Save Method Using 248-point CNT's with m=496

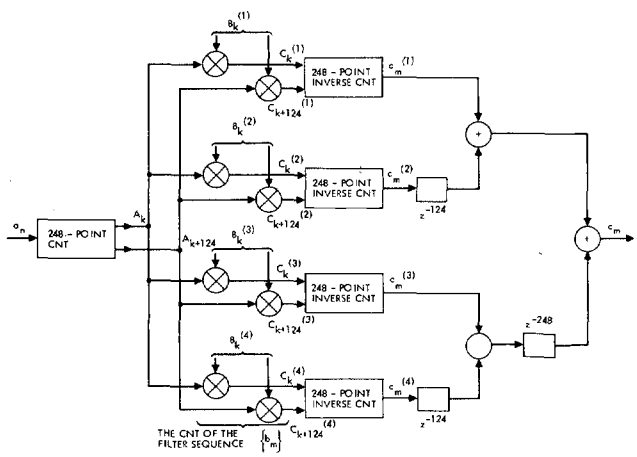

Figure 4. A Realization of a Digital Filter With a Filter Sequence of 496 Points by Using the Generalized Overlap-Save Method and the CNT Technique

REFERENCES

[1] L.R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliff New Jersey, 1975.

[2] T.K. Truong, I.S. Reed, C.-S. Yeh, and H.M. Shao, "A Parallel VLSI Architecture for a Digital Filter of Arbitrary Length Using Fermat Number Transforms," Proceedings of ICCC 82, pp. 574-578, New York, Sept. 28-Oct 1, 1982.

[3] C.M. Rader, "Discrete Convolutions Via Mersenne Transforms," IEEE Trans. Computers, Vol. C-21, No. 12, pp. 1269-1273, Dec. 1972.

[4] I.S. Reed and T.K. Truong, "The Use of Finite Field to Compute Convolutions," IEEE Trans. Information Theory, Vol. IT-21, No. 2, pp. 208-213, March 1975.

[5] H.J. Nussbaumer, "Digital Filtering Using Complex Mersenne Transforms," IBM Journal of Research and Development, Vol. 20, No. 5, Sept. 1976.

[6] H.T. Kung, "Why Systolic Architectures?" IEEE Computer, Vol. 25, No. 1, pp. 37-46, Jan. 1982.

[7] C.A. Mead and L.A. Conway, Introduction to VLSI Systems, Addison-Wesley, Reading, Mass., 1980.