VERY HIGH SPEED REAL-TIME IIR DIGITAL FILTER STRUCTURES: SUITABLE FOR VLSI IMPLEMENTATION

Kaushal K. Dhar

ETF C115, Communication Technology Laboratory
Swiss Federal Institute of Technology
IKT-ETH Zentrum, CH 8092 Zurich, Switzerland.
Tel. No. : (+41) 1 256 5265
Fax No. : (+41) 1 262 0943, E-mail : dhar@nari.ikt.ethz.ch

I. ABSTRACT

The techniques to realize very high speed, linear, time invariant and recursive digital filters, such that the achieved throughput rates are not limited by the atomicity of the basic operation of the device used to implement them, are presented here. The usual constraint in realizing very high speed IIR filters is that only one delay is tolerated in the feedback path. In order to overcome it, the given transfer function is suitably modified. The resulting scheme consist of several processing paths, which can be implemented by multi-processing systems, that use commercial processors. Additionally, the present structures are made up of modules, which are repeated many times over, with mainly local interconnections. Hence, they possess substantial potential to be implemented on VLSI's.

II. INTRODUCTION

Several high speed signals like those in radar and spread spectrum system [5]-[6] require IIR digital filtering. However, unlike their FIR counterparts, the IIR digital filters can not be implemented at arbitrary speeds by using normal pipelining techniques which create input/output (I/O) delay of several sampling intervals. The feed back loop which realizes the recursive part of the transfer function can not tolerate the delay of more than one sampling interval because that will alter the transfer function. However, some pipelining methods, to implement IIR transfer functions at high speeds, have also been suggested [10]--[11]. The number of pipelining stages that can be used to implement an arithmetic operation like multiplication are typically limited to 10. These implementations operate at speeds which are bound by the rate at which the basic arithmetic operations, referred to as atomic operations, can be executed in pipelined fashion [7]-[8]. Several schemes have been suggested in order to implement very high speed IIR filters using block-processing approach [4] and [12]. They improve the theoretical value of sample period bound [9] beyond the limit determined by the atomic operations. The hardware complexity of a systolic realization based on block processing approach [12], in terms of number of multiplier/adder units, is much larger than what can be reasonably expected from the speedup factor it achieves over a single multiplier/adder realization. Several schemes have also been proposed to implement very high speed IIR digital filters on VLSI chips. In [13], the block processing approach has been combined with pipelining and it results in an upper bound of sampling rate $f_{max} = \frac{1}{B} MAX[\frac{\Delta_l}{S_l}]$, where the maximum is over all the loops of the computation graph and the terms Δ_l and S_l are the latency associated with the computation in the loop l and the corresponding number of logical delay operators in that loop. The loop corresponding to which the ratio $\frac{\Delta_l}{S_l}$ is maximum is called the critical loop. The term B refers to the block size and if the S_l's are the same for all the loops and are equal to S, then the number of pipeline stages is S. The overall speedup factor over a single processor realization is ideally SB, though the actually achieved value is substantially lower. The required hardware complexity although linear with respect to the filter size M is still too large. The result is that the achieved multi-processor efficiency is lower than a satisfactory value. Moreover, the I/O delay is also very large which may be a limitation in certain coherent telecommunication [5]-[6] and video signal processing applications, where the signal phase is critical.

III. PRESENT OBJECTIVE

Although the above mentioned methods can implement very high speed IIR filters by means of VLSI technology, yet a matter of continued significance is to further reduce the required hardware complexity and the I/O delay for a certain desired speedup over a single processor realization. The aim of this presentation is to suggest alternate efficient schemes which attain the following objectives: (i). To be able to perform filtering at an arbitrarily high sampling rate not limited by the filter order M. (ii). To have the speedup gained over a single processor realization to be commensurate with the resulting hardware complexity. This implies that the achieved multi-processor efficiency $\eta_P = \frac{C_1 S_P}{C_P} = \frac{C_1}{C_P} \cdot \frac{f_{sP}}{f_{s_1}}$ be as close to unity as possible. Here, f_{s_1} and f_{sP} are the sampling frequencies at

0-7803-1254-6/93$03.00 © 1993 IEEE

which the signal can be processed by a sequential implementation employing a single processing element, and a parallel implementation respectively. The speedup factor is S_P. The terms C_1 and C_P are the cost functions representing the hardware complexities associated with the sequential and the parallel implementations respectively. (iii). To propose flexible structures such that they can either be integrated as VLSI chips or implemented by means of multi-processor architectures using several of the commercially available digital signal processing (DSP) components. It should be possible to process signals having sampling periods which are even smaller than the time required by the aforementioned DSP components to perform an atomic operation. Here, we assume that the available device technology is such that the pipelining technique by itself is not adequate to perform the processing at the desired sampling rate.

IV. PRESENT APPROACH

In this presentation, we propose very high speed IIR digital filters utilizing the structures based on delayed multi-path scheme [1]-[2], which use N-path sections of Fig. 1(a). An IIR digital filtering equation, in time domain, is given as below :

$$\begin{aligned} y(nT) &= a_0x(nT) + \cdots + a_fx(nT - fT) \\ &\quad + b_1y(nT - T) + \cdots + b_hy(nT - hT) \end{aligned} \tag{1}$$

The corresponding z-domain equation is given by:

$$\begin{aligned} H_I(z^{-1}) &= \frac{N(z^{-1})}{D(z^{-1})} = \frac{N(z^{-1})}{1 + z^{-1}D_1(z^{-1})} \\ &= \frac{a_0 + a_1z^{-1} + \cdots + a_fz^{-f}}{1 - b_1z^{-1} - \cdots - b_hz^{-h}} \end{aligned} \tag{2}$$

where $D_1(z^{-1})$ is a polynomial in z^{-1}. This IIR equation can be directly realized as in Fig. 2(a). Each of the two transfer functions $N(z^{-1})$ and $D_1(z^{-1})$ can process high speed signals if their realizations use N-path elements having N parallel processing paths. If the time taken to process along the i^{th} path is t_i, the maximum uniform sampling interval at which an input signal can be processed, is $T_s = \frac{t_m}{N}$ such that $t_m = MAX[t_i]$ and the maximum is over the path indices i. The term N is a design parameter. As far as the realization of the numerator part is concerned, it can be done by any one of the delayed multi-path structures presented in [2]. However, the use of multi-path section results in N sample interval I/O delay. This precludes its use in realizing the denominator of the given IIR rational transfer function $H(z^{-1})$, which determines the feed back loop part of the direct realization scheme, because the feedback loop allows only one delay. Hence, if we try to utilize N-path elements in the feedback loop, the realized transfer function is different from the desired one. The desirable approach now is to modify the transfer function such that the feedback loop, in the direct realization form, can accommodate N sample interval delay and in the ensuing process two features are kept in mind : (a). The increase in complexity is as small as possible. (b). The modified transfer function is still stable.

There are several ways to modify the IIR transfer function such that the denominator polynomial becomes a function of only z^{-N}. A method given in [4] is used here. An all-pole filter of the Fig. 2(b), having the transfer function $H_p(z^{-1})$, can also be represented by a Block-recursive filter structure $H_B(z^{-1})$, of the Fig. 2(c), which consists of an all zero block $H_1(z^{-1})$ in the forward path, and another all zero block $G(z^{-1})$ along with an arbitrary delay of L sample intervals, in the feedback path. The value L depends on the desired block size, and hence can be chosen as a parameter of interest. A method to derive $H_1(z^{-1})$ and $G(z^{-1})$ from the given $H_p(z^{-1})$ and a desired L has been provided in [4]. However, an alternate approach is employed here, which to us appears more straightforward. A single sample delayed version of (1) is given as:

$$\begin{aligned} y(nT - T) &= a_0x(nT - T) + a_1x(nT - 2T) + \cdots \\ &\quad + a_fx(nT - (f + 1)T) + b_1y(nT - 2T) \\ &\quad + \cdots + b_hy(nT - (h + 1)T). \end{aligned} \tag{3}$$

Another form of the digital filter equation is obtained by substituting (3) into (1), which then shows that, as far as the recursive part is concerned, $y(nT)$ is dependent only on $y(nT - cT)$, where $c \geq 2$. It implies that now the recursive loop can tolerate a delay of 2, instead of just one, as was the case in (1), because the z-domain representation of the new equation has a denominator of the type $(1 + z^{-2}D_2(z^{-1}))$, where $D_2(z^{-1})$ is a polynomial in z^{-1}. Extending this argument further, we can obtain a k samples delayed version, to be mentioned as the k-delayed version hereinafter, of $y(nT)$ i.e., $y(nT - kT)$, given as:

$$\begin{aligned} y(nT - kT) &= a_0x(n - k)T + a_1x(n - k - 1)T \\ &\quad + \cdots + a_fx(n - f - k)T \\ &\quad + b_1y(n - k - 1)T + \cdots \\ &\quad + b_hy(n - h - k)T. \end{aligned} \tag{4}$$

Arguing along the present lines, we can make $y(nT)$ dependent only on $y(nT - cT)$, where $c \geq (k+1)$. Continuing this way, an IIR transfer function of (2) can be converted into a desired form, which is the $(N - 1)$-delayed version, given as

$$H_I^{(N-1)}(z^{-1}) = \frac{N_1(z^{-1})}{1 + z^{-N}D_N(z^{-1})} \, . \tag{5}$$

Looking at the denominator now, we see that the z-domain transfer function, to be realized inside the recursive loop, is $z^{-N}D_N(z^{-1})$. Hence, if we try to realize the transfer function $D_N(z^{-1})$, using the delayed multi-path approach, we end up with $z^{-N}D_N(z^{-1})$ because a delay of N sample intervals is inherently provided by the N-path realization. However, in this kind of approach to transform the IIR transfer function, we have to be aware of stability aspects and the additional computational burden created. In the given h^{th} order difference equation (1),

we can say, without the loss of generality, that $h \geq f$. The L-delayed version of (2) becomes

$$H_I^L(z^{-1}) = \frac{[a_0 + .. + a_f z^{-f}][1 + a_1^L z^{-1} + .. + a_L^L z^{-L}]}{1 - b_{L+1}^L z^{-L-1} - .. - b_{L+h}^L z^{-L-h}}. \quad (6)$$

Looking at the terms which are in the denominator of $H_I^{(L)}(z^{-1})$, it is clear that the computational complexity of the recursive part of (6) is no different from that of (2). Hence, the time needed to perform the feedback path processing in the Fig. 2(c), remains the same as that of h multiplications and summation over $h-1$ quantities. However, now we have to perform $f+L+1$ multiplications and a summation over $f+L$ quantities along the forward path, in place of $f+1$ multiplications and the summation over f quantities. This requires an additional $2L$ computation operations which need a time of $2LT_a$. We present a structure, given in Fig. 3(a), which is representative of the application of the present approach to realize the IIR filters. Figure 3(b) represents the structure of each of the modules of Fig. 3(a). As a matter of fact, it only realizes the all-pole part i.e., $\frac{1}{1+z^{-N}D_N(z^{-1})}$ of (5), by computing the intermediate z-domain signal $Y_1(z^{-1}) = \frac{X(z^{-1})}{1+z^{-N}D_N(z^{-1})}$. This is because the numerator part can be simply realized by one of the schemes suggested in the [2], and subsequently cascaded to it. This structure is by no means the only one possible because many more can be organized on the basis of the schemes given for the FIR case in [2].

V. DISCUSSION

In this presentation, we suggested improved structures, based on the delayed multi-path scheme to realize high speed digital filters in the time domain, in order to process communications signals. They offer a promising advantage, that the speedup factor obtained is not limited by the order of the filter to be realized. Hence, even the small order filters can be implemented such that very high throughput rates are achieved, but still the architectural complexity is commensurate with the speedup gained over the SISD implementation. The result is that reasonably good figures of the multi-processor efficiencies are obtained even when the number of processing paths is very large. These realizations are constituted of several identical modules, which have mostly local interconnections. This feature strongly qualifies them for an implementation in the form of VLSI's [3].

Let us discuss some of the implementation aspects here. Each of the processing paths receives the input signal sample train at a rate decimated by N. However, the two elements which are required to have response times determined by the actual input sampling rate, are the input $1{:}N$ demultiplexer, and the output $N{:}1$ multiplexer. For an example, if we are to process input samples arriving at the rate of 100 MSPS (Mega samples per second), and we use 1:10 demultiplexer at the input side, then each one of the resulting processing paths will have to process at the rate of 10 MSPS, while the multiplexer and the demultiplexer must be specified at 100 MSPS.

REFERENCES

[1] K. Hayashi, K.K. Dhar, K. Sugahara and K. Hirano, "Design of high-speed digital filters suitable for multi-DSP implementation," IEEE Trans. Circuits & Syst., vol-CAS-33, pp. 202-217, Feb. 1986.

[2] Kaushal K. Dhar, " Realization of very high speed real-time FIR digital filters, " Proc. of the IEEE International Conference on Acoustic, Speech and Signal Processing, ICASSP-92, pp. IV-297 –IV-300, 1992, San Francisco, California, USA.

[3] K.K. Dhar, "Multi-processor architectures for very high speed recursive digital processing of real-time communication signals," Swiss Patent awarded.

[4] H.B. Voelcker and E.E. Hartquist, "Digital filtering via block recursion," IEEE Trans. Audio Electroacoustics, pp. 169-176, June 1970.

[5] Urs Fawer and Kaushal K. Dhar, " Digital implementation of a coherent diversity-receiver, " The 1991 IEEE Int. Workshop on Microelectronics in Communications, March 10-13, 1991, Interlaken, Switzerland.

[6] Urs Fawer, " A coherent spread-spectrum diversity--receiver with AFC for multipath fading channels, " IEEE Transactions on Communications, in press.

[7] Alfred Fettweis, "Realizability of digital filter networks," Arch. Elek. Ubertragung, Vol. 30, pp. 90-96, Feb. 1976.

[8] Markku Renfors and Yrjo Neuvo, "Maximum sampling rate of digital filters under hardware speed constraints," IEEE Trans. Circuits & Syst., Vol. CAS-28, No. 3, pp. 196-202, March 1981.

[9] D.A. Schwartz and T.P. Barnwell, "Optimal implementation of flow graphs on synchronous multiprocessors," Proc. of the IEEE Int. Conf. on Acoustic, Speech, and Signal Processing, pp. 1384-1387, 1985.

[10] Herschel H. Loomis and Bhaskar Sinha, "High-speed recursive digital filters realization," Circuits, Systems, Signal Processing, Vol. 3, No. 3, pp. 267-294, 1984.

[11] Wyong Sung and Sanjit K. Mitra, "Implementation of digital filter algorithms using pipelined vector processors," Proceedings of the IEEE, Vol. 75, No. 9, Sept. 1987.

[12] H.H. Lu, Edward A. Lee and David G. Messerschmitt, "Fast recursive filtering with multiple slow processing elements," IEEE Transactions on Circuits and Systems, Vol. CAS-32, No. 11, Nov. 1985.

[13] Keshab K. Parhi and David G. Messerschmitt, "Pipeline interleaving and parallelism in recursive digital filter-part I: Pipelining using scattered look--ahead and decomposition," IEEE Tranactions on Acoustic, Speech and Signal Processing, Vol. 37, No. 7, July 1989.

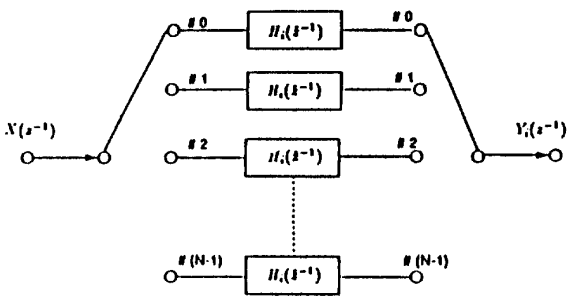

Fig. 1(a) A digital N-path element.

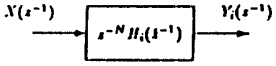

Fig. 1(b) An equivalent block representation of digital N-path element.

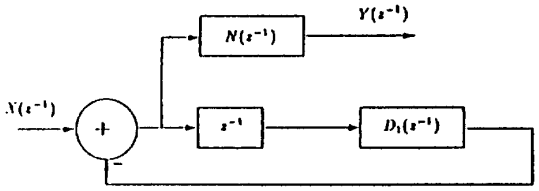

Fig. 2(a) A structure for direct realization of IIR filters.

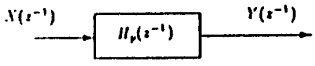

Fig. 2(b) An all-pole filter $H_p(z^{-1})$.

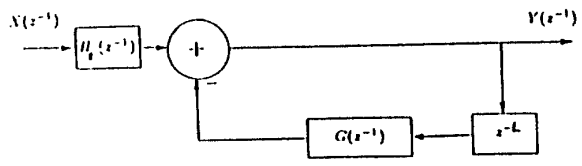

Fig. 2(c) A block-recursive realization $H_R(z^{-1})$ of $H_p(z^{-1})$.

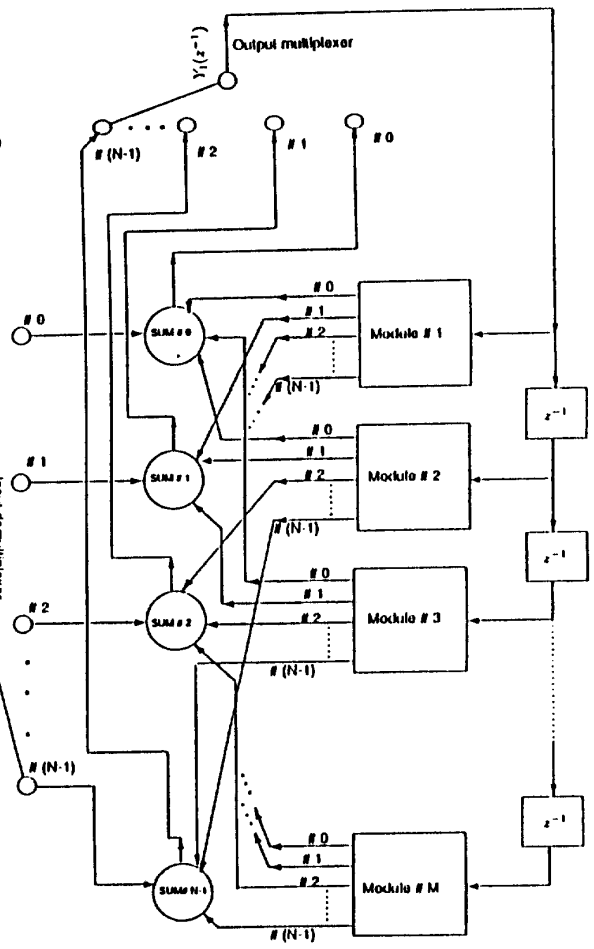

Fig. 3(a) Realization of an all-pole IIR section $\frac{1}{1+z^{-N}D_N(z^{-1})}$.

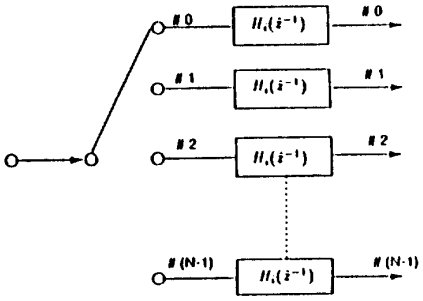

Fig. 3(b) Constituent "Module" of the filter structure of Fig. 3(a).