Efficient Multirate Realization for Narrow Transition-Band FIR Filters

Tor A. Ramstad

Department of Electrical and Computer Engineering
The Norwegian Institute of Technology
N-7034 Trondheim-NTH, Norway

Tapio Saramäki

Department of Electrical Engineering
Tampere University of Technology
P. O. Box 527, SF-33101 Tampere, Finland

Abstract **– A filtering technique for reducing the computational complexity in FIR filters at any bandwidth is advanced. This technique uses a combination of multirate and complementary filters. It is shown that the resulting computational complexity is almost independent of the transition bandwidth of the filter, but depends somewhat on the cut-off frequency.**

I. Introduction

FIR filters have several advantages over IIR filters such as guaranteed stability, possibility of linear phase, and usually shorter necessary wordlengths. However, for narrow transition-band filters, the high FIR filter order makes them unsuited in many applications.

There are several ways of reducing the complexity of FIR filters through the use of multistage implementations. For narrow-band filters, two methods are IFIR (interpolated FIR) [1],[2] and multirate structures [3]. For narrow transition-band filters with arbitrary bandwidth, these methods cannot be used in a straightforward manner. However, Jing and Fam [4] have suggested a very elegant technique which is able to cope with any bandwidth. Their method is a combination of IFIR and complementary filtering, and leads to a dramatic reduction in the computational complexity.

The method we suggest is an extension of the theory in [4], where multirate techniques are included to reduce the computational complexity even further. It will be shown that the multiplication rate is fairly independent of the transition bandwidth, but depends somewhat on the position of the transition band.

The paper is organized as follows: First, we consider the basic processing structures and a procedure to combine these to obtain an efficient overall realization. Then, we work out a system for certain bandwidths where we apply the same filters in a repetitive manner, and give an example of this. This is followed by a discussion on how to obtain arbitrary bandwidths.

II. Basic Principles

We first review the two filtering techniques upon which our method rely, namely complementary filters and multirate filtering. In the discussion we distinguish between two filter classes: narrow-band filters where the passband occupies a bandwidth less than $f_s/4$ with f_s being the sampling frequency and broad-band filters where the passband occupies more bandwidth than $f_s/4$.

A. Complementary Filters

If a broad-band filter is to be designed, we try to transform this into an equivalent structure containing a narrow-band filter in order to easily apply decimation/interpolation techniques. If $H(z)$ represents a broad-band linear-phase filter, we can construct an equivalent problem in terms of a narrow-band complementary filter $H_C(z)$ [5] as

$$H(z) = z^{-D} - H_C(z), \tag{1}$$

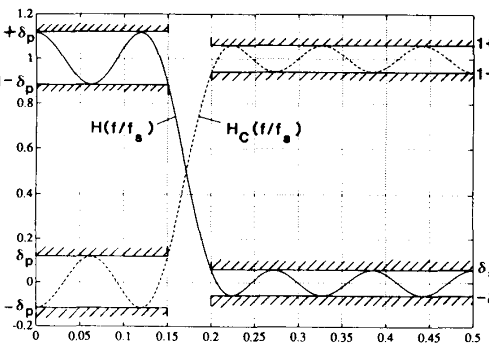

Fig. 1. *Frequency responses of two complementary filters.*

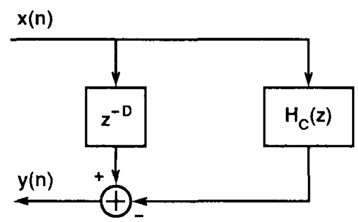

Fig. 2. *Realization of $H(z)$ using $H_C(z)$ and a parallel delay.*

where z^{-D} is the delay of the two filters. The lengths of these filters are $2D+1$ and the impulse responses possess an even symmetry. If the impulse response coefficients of $H_C(z)$ are $h_C(n)$, then the zero-phase frequency responses of $H(z)$ and $H_C(z)$ can be written as

$$H_C(f/f_s) = h_C(D) + \sum_{n=1}^{D} 2h_C(D-n)\cos(2\pi n f/f_s) \tag{2a}$$

$$H(f/f_s) = 1 - H_C(f/f_s). \tag{2b}$$

According to Eqn. (2b), the sum of these two zero-phase responses is unity. Therefore, these filters are called complementary filters.

If it is desired to design a broad-band lowpass filter $H(z)$ with zero-phase response approximating unity on $[0, f_p]$ with ripple δ_p and zero on $[f_c, f_s/2]$ with ripple δ_s, then the synthesis is converted to the design of a narrow-band highpass filter $H_C(z)$ with zero-phase response approximating zero on $[0, f_p]$ with ripple δ_p and unity on $[f_c, f_s/2]$ with ripple δ_s, as shown in Fig. 1. The filter structure which implements $H(z)$ in terms of $H_C(z)$ is shown in Fig. 2. With this complementing technique, all filters (except when $f_c \approx f_s/4$) can be made narrowband.

ISCAS'88

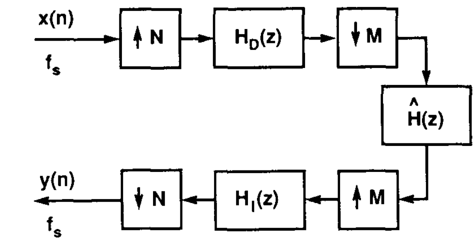

Fig. 3. *Two-rate realization of a narrow-band filter.*

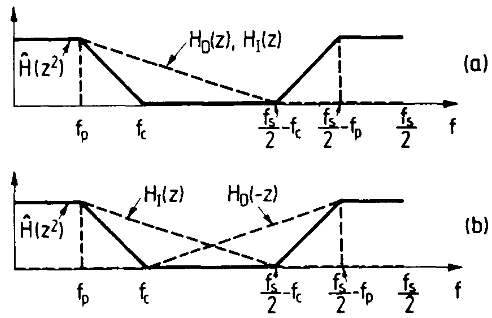

Fig. 4. *Transfer functions for the aliased and unalised components in the narrow-band case for $r = 2$. (a) Terms in $F_1(z)$. (b) Terms in $F_2(z)$.*

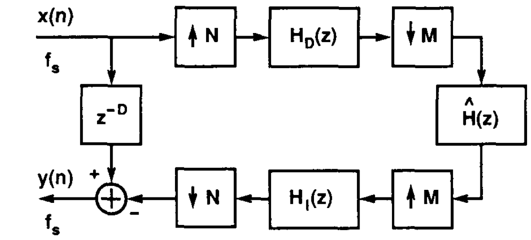

Fig. 5. *Two-rate realization of a broad-band filter.*

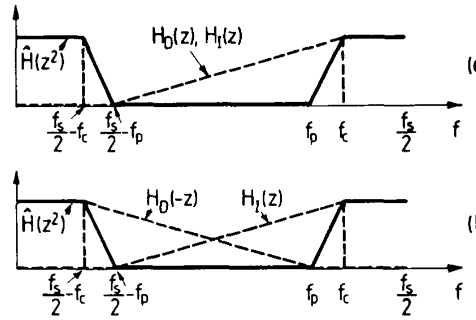

Fig. 6. *Transfer functions for the aliased and unalised components in the broad-band case for $r = 2$. (a) Terms in $F_1(z)$. (b) Terms in $F_2(z)$.*

B. Multirate techniques

We now consider the narrow-band case. An efficient realization for this case is shown in Fig. 3, where the sampling rate is first increased by a factor of N, then follows an anti-aliasing filter $H_D(z)$ and, finally, the sampling rate is reduced by a factor of M. $\widehat{H}(z)$ operates at the low sampling frequency f_s/r, where $r = M/N$. After this filter, the sampling frequency is again increased to the input sampling frequency through an inverse operation where the filter $H_I(z)$ removes all the undesired image components.

This contribution concentrates mainly on the case $r = 2$ ($N = 1, M = 2$). In this case, the relation between the z-transforms of the input signal $x(n)$ and the output signal $y(n)$ is

$$Y(z) = F_1(z)X(z) + F_2(z)X(-z), \tag{3a}$$

where

$$F_1(z) = \widehat{H}(z^2)H_D(z)H_I(z) \tag{3b}$$

$$F_2(z) = \widehat{H}(z^2)H_D(-z)H_I(z). \tag{3c}$$

Here, $F_1(z)$ is a conventional transfer function from the input to the output, whereas $F_2(z)X(-z)$ is the aliased term due to the sampling rate alteration. Figure 4 illustrates the terms in $F_1(z)$ and $F_2(z)$. The desired overall design with edges f_p and f_c is achieved by selecting the passband and stopband edges of $\widehat{H}(z)$ to be

$$\hat{f}_p = f_p, \quad \hat{f}_c = f_c \tag{4}$$

and the edges of both $H_D(z)$ and $H_I(z)$ to be

$$f_p^{(D)} = f_p, \quad f_c^{(D)} = f_s/2 - f_c. \tag{5}$$

Because of the periodicity, $\widehat{H}(z^2)$ has, in addition to the passband $[0, f_p]$ and the transition band $[f_p, f_c]$, an extra transition band $[f_s/2-f_c, f_s/2-f_p]$ and an extra passband $[f_s/2-f_p, f_s/2]$. As seen from Fig. 4(a), the terms $H_D(z)$ and $H_I(z)$ included in $F_1(z)$ preserve the lower passband region and attenuate the extra transition band and passband regions. The resulting passband ripple is at most the sum of the passband ripples of the three filters and the stopband ripple is on $[f_c, f_s/2 - f_c]$ at most the stopband ripple of $\widehat{H}(z^2)$ and on $[f_s/2 - f_c, f_s/2]$ the product of the ripples of $H_D(z)$ and $H_I(z)$. When considering the terms of $F_2(z)$ in Fig. 4(b), it is observed that $H_I(z)$ attenuates the upper passband and transition band regions of $\widehat{H}(z^2)$ and $H_D(-z)$ attenuates the lower ones. Therefore, the maximum value of the aliased components is limited by the maximum of the stopband ripples of the three filters.

Since the input sampling rate of the center filter $\widehat{H}(z)$ is

$$\hat{f}_s = f_s/2, \tag{6}$$

the relative transition bandwidth of $\widehat{H}(z)$ is two times that of the overall filter $H(z)$. This means that the order of $\widehat{H}(z)$ is approximately halved as compared with $H(z)$ since the order of an FIR filter is roughly inversely proportional to the transition bandwidth. Also accounting for the reduced sampling rate, the expected multiplication rate in $\widehat{H}(z)$ is one fourth of that of $H(z)$. The total multiplication count includes the processing in $H_D(z)$ and $H_I(z)$.

Returning to the broad-band case, we can combine the realizations of Fig. 2 and Fig. 3 to get the structure of Fig. 5. According to the above discussion, $H_C(z)$ is narrow-band highpass filter with stopband edge f_p and passband edge f_c. For $r = 2$, the input-output relation becomes

$$Y(z) = [z^{-D} - F_1(z)]X(z) - F_2(z)X(-z), \tag{7}$$

where $F_1(z)$ and $F_2(z)$ are given by (3b) and (3c). The desired result is obtained by designing $\widehat{H}(z)$ to be a lowpass filter with passband and stopband edges

$$\hat{f}_p = f_s/2 - f_c, \quad \hat{f}_c = f_s/2 - f_p \tag{8}$$

and by designing both $H_D(z)$ and $H_I(z)$ to be highpass filters with stopband and passband edges

$$f_c^{(D)} = f_s/2 - f_p, \quad f_p^{(D)} = f_c. \tag{9}$$

In this case, the periodic filter $\widehat{H}(z^2)$ has two passband regions $[0, f_s/2-f_c]$ and $[f_c, f_s/2]$ and one stopband region $[f_s/2-f_p, f_p]$,

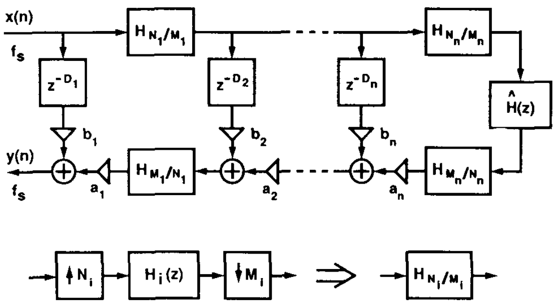

Fig. 7. *Proposed multistage multirate realization of an FIR filter.*

as shown in Fig. 6(a). The basic difference compared to the above narrow-band case is that now $H_D(z)$ and $H_I(z)$ included in $F_1(z)$ preserve the upper passband $[f_c, f_s/2]$ of $\hat{H}(z^2)$ and attenuate the lower passband and transition band regions. Correspondingly, the resulting passband and stopband regions of $z^{-D} - F_1(z)$ are $[0, f_p]$ and $[f_c, f_s/2]$, as is desired. The passband ripple of $z^{-D} - F_1(z)$ is at most the product of the stopband ripples of $H_D(z)$ and $H_I(z)$ on $[0, f_s/2 - f_p]$ and at most the stopband ripple of $\hat{H}(z^2)$ on $[f_s/2 - f_p, f_p]$. The stopband ripple is at most the sum of the passband ripples of the three filters. When considering the terms of $F_2(z)$ in Fig. 6(b), the basic diffrerence compared to the above narrow-band case is that now $H_I(z)$ attenuates the lower passband and transition band regions of $\hat{H}(z^2)$ and $H_D(-z)$ attenuates the upper ones. Also, in this case, the sampling frequency of the center filter is half of the sampling rate of the overall filter and the relative transition bandwidth is two times that of the overall filter.

C. Multistage Realizations

To further reduce the number of computations per input sample, $\hat{H}(z)$ can be implemented using either the structure of Fig. 3 or the structure of Fig. 5, depending on whether its edges are smaller or larger than $\hat{f}_s/2$. Later on, we denote the structures of Figs. 3 and 5 by structure I and structure II, respectively. The process can be continued until the transition bandwidth of the center filter becomes large enough and, correspondingly, its order becomes small enough. The resulting overall structure is shown in Fig. 7, where (a_i, b_i) is either $(1, 0)$ or $(-1, 1)$. The former case corresponds to the case where the center filter in the procedure of constructing the overall filter is narrowband. In this case, the delay term is not needed. In the latter case, the center filter is constructed using structure II.

D. Half-Band Decimators and Interpolators

For constructing $H_I(z)$ and $H_D(z)$, half-band filters [6] are particularly efficient. These filters are characterized by the facts that the passband and stopband edges f_p and f_c are related by $f_c = f_s/2 - f_p$ and the passband and stopband ripples are the same. This implies that every second coefficient in these filters is equal to zero except for the central coefficient of value 1/2. Therefore, by exploiting the coefficient symmetry and assuming that the central coefficient is implemented without general multiplier, we can implement with R multipliers a filter of length

$$Q = 4R - 1. \tag{10}$$

From the above relation between the passband and stopband edges of half-band filters, it follows that the passband edge of both $H_D(z)$ and $H_I(z)$ has to be selected to be

$$f_p^{(D)} = f_c \tag{11}$$

for structure I. For structure II, the half-band filter is a highpass filter with stopband edge $f_c^{(D)}$ as given in (9). The input sampling rate of the kth decimation stage and the output sampling rate of the kth interpolation stage is $2^{-(k-1)}$ times that of the overall filter. By exploiting the facts that only every second output of the decimator needs to be computed and only every second input to the interpolator is non-zero, the overall multiplication rate of an n-stage design becomes

$$M_n = \sum_{k=1}^{n} 2^{-(k-1)} R^{(k)} + 2^{-n} R, \tag{12}$$

where $R^{(k)}$ is the number of multipliers in the kth filter stage and R is the number of multipliers in the remaining center filter. It can be shown that by ignoring the aliased terms, the passband and stopband ripples of the overall design are at most

$$\delta_p = \sum_{k \in S} 2\delta^{(k)} + \hat{\delta} \tag{13a}$$

$$\delta_s = \sum_{k \notin S} 2\delta^{(k)} + \tilde{\delta}, \tag{13b}$$

where S is the set of stages in front of which there is an even number of true delay terms or no delay terms and $\delta^{(k)}$ is the ripple of the kth filter stage. $\hat{\delta}$ ($\tilde{\delta}$) is the passband (stopband) ripple of the center filter if there is an even number of delay terms in front of it. Otherwise, it is the stopband (passband) ripple.

III. Repetitive Structures

If we consider a design problem where the transition bandwidth is very small ($f_c - f_p \approx 0$), certain cases lead to very simple structures. Of particular interest is the case where the required broad-band filter has a relative cutoff frequency f_c/f_s. After running through one of the stages of structure II, we end up with $\hat{H}(z)$ where the same relative cutoff frequency applies ($\hat{f}_c/\hat{f}_s = f_c/f_s$). The condition for this to be fulfilled is that $(f_s/2 - f_c)r = f_c$, which resolves to

$$f_c = \frac{r}{2(1+r)} f_s, \tag{14}$$

where r is the decimation factor. r should be chosen either as an integer or a rational number.

The multiplication count per sample for an n-stage system can be estimated to

$$M_n = M_H r^{-2n} + R \sum_{k=1}^{n} r^{-(k-1)} = M_H r^{-2n} + R \frac{1 - r^{-n}}{1 - r^{-1}}, \tag{15}$$

where M_H is the number of multiplications per sample in a single-filter implementation and R is the number of different multipliers in each decimator and interpolator.

As a limiting case as n grows to infinity

$$M_\infty = R \frac{r}{r-1}. \tag{16}$$

This is a remarkable result! It states that a zero-length transition band can be obtained using a finite number of multiplications per sample. There are, however, certain factors which obscure this result. According to the considerations in the previous section, the resulting overall stopband attenuation does not only depend on the attenuation in the decimators/interpolators, but also on the number of stages. As a consequence, a somewhat larger R is necessary when more stages are added to the structure.

Example: Repetitive use of structure II with $r = 2$. An interesting example is the case when $r = 2$ and we apply structure II in a repetitive manner. In this case, we can use highpass half-band filters with stopband edge of 1/3 for all the decimation and interpolation stages. For $R = 5$ (length=19), the stopband attenuation is 59.5 dB, and as R is increased, the stopband attenuation increases approximately 10 dB per an extra multiplier.

Here, we want to consider the following case. We use n stages of structure II in the structure of Fig. 7. We assume that

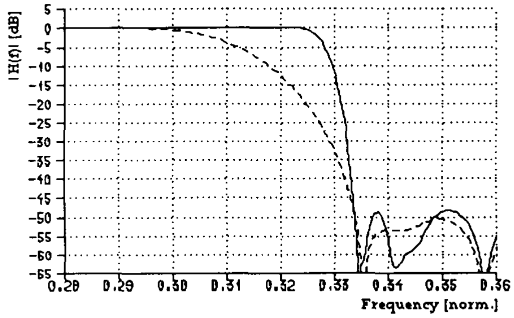

Fig. 8. *Transition regions for a two-stage (dashed line) and four-stage (solid line) filter system based on struture-II half-band filters.*

the center filter filter $\widehat{H}(z)$ has the same length as the decimation/interpolation filters. However, $\widehat{H}(z)$ is not a half-band filter and requires therefore twice the number of multipliers. The total multiplication rate with n stages is therefore

$$M_n = R\sum_{k=1}^{n} 2^{-(k-1)} + 2R \cdot 2^{-n} = 2R. \qquad (17)$$

For this particular case the multiplication rate is independent of the number of stages!

Figure 8 gives the transition regions of two filters, one using two decimation stages (dashed line) and one with four stages (solid line), both composed of length 19 filters. The filters require 10 multiplications per input sample. The memory required is 216 and 900 for the two cases, respectively, including the internal storage in the subfilters. The transition bandwidths of the two filters are approximately 0.04 and 0.01 on a normalized scale, respectively. There is about a 2 dB difference in the minimum stopband attenuation of the two filters.

To meet the tighter specifications, a one-stage FIR and an elliptic filter would require 120 and 14 multiplications, respectively, if we exploit the coefficient symmetries in both cases.

IV. Arbitrary Bandwidth Filters

The design technique described in Section II cannot be used when $f_p < f_s/2$ and $f_c > f_s/2$. Also, if f_c or f_p is close to $f_s/2$, then the order of a half-band filter to meet the resulting criteria becomes very high because of a very narrow transition band. Similarly, the design procedure of Section II cannot be repeated if the edges of the center filter become close to one fourth of its sampling rate. To avoid these problems, one alternative is to use $r = 4/3$ ($N = 3$, $M = 4$). In this case, the decimation filter can be constructed using two low-order half-band filters with the input sampling rates of these filters being $3f_s$ and $3f_s/2$. The advantage of this alternative is that the resulting passband and stopband edges of the new center filter are shifted by a factor of 4/3 with respect to its sampling rate. More details about this will be given in a full length paper to be published.

Theoretically, using $r = 2$ and $r = 4/3$ and structures I and II, we can implement any lowpass filter with zero bandwidth using a finite multiplication rate. Figure 9 gives plots of the number of multiplications per input sample versus the passband edge f_p for the case where $f_c \equiv f_p$ and the number of filter stages is not limited. The curves are given for the cases where the minimum stopband attenuation of each decimation and interpolation stage is 120 dB (uppermost curve), 100 dB, 80 dB, and 60 dB (lowermost curve). In constructing these plots, $r = 4/3$ was used when the cut-off frequency of the center filter was within 0.9–1.1 times one fourth of its sampling rate. In all practical cases, the transition bandwidth is not zero and the sampling rate of the center filter is after ten stages so low that its contribution to the overall multiplication rate is neglibigle. Assuming that the number of structures I and II is about the same, this means, according to the discussion of Section II, that the resulting stopband ripple is at most ten times that of each filter stage. The curves of Fig. 9 give thus practical upper estimates for the overall multiplication rate in the cases where the stopband attenuation of the overall design is lower by 20 dB and the transition bandwidth is extremely narrow. As seen from Fig. 9, the implementation of a filter having a 100-dB attenuation requires in the worst case only 45 multiplications per input sample!

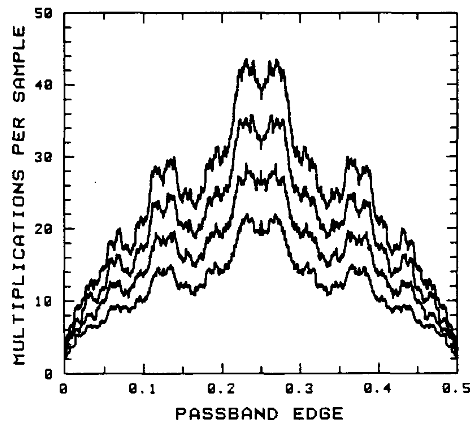

Fig. 9. *Plots of the number of multiplications per input sample versus the passband edge f_p for filters with zero transition bandwidth. The curves starting from the uppermost are for the cases where the minimum stopband attenuation for each filter stage is 120, 100, 80, and 60 dB, respectively.*

V. Conclusion

This paper has shown that the multiplication rate in FIR filters is limited if multirate techniques are applied. Any reasonable filter specification can be met using of the order of 45 multiplications per sample. The price to be paid to obtain this is extra memory. One should also notice that the design procedure is quite simple. All the subfilter are of low order. Consequently, the necessary coefficient wordlengths are modest.

Acknowledgement

This work was partly financed by the University of California Micro Grant. We would like to express our thanks to Professor Sanjit Mitra who made this possible.

References

[1] Y. Neuvo, C.-Y. Dong, and S. K. Mitra: "Interpolated finite impulse response filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 563–570, June 1984.

[2] T. Saramäki, Y. Neuvo, and S. K. Mitra: "Design of computationally efficient interpolated FIR filters," *IEEE Trans. Circuits Syst.*, vol. CAS-35, pp. 70–88, Jan. 1988

[3] L. R. Rabiner and R. E. Crochiere, "A novel implementation for narrow-band FIR digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, pp. 457–464, Oct. 1975.

[4] Z. Jing and A. T. Fam, "A new structure for narrow transition band, lopass digital filter design," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 362–370, April 1984.

[5] T. A. Ramstad, "Branch filtering using FIR and IIR complementary structures," in *Signal Processing*, Academic Press 1980, edited by V. Cappellini and A. G. Constantinides.

[6] M. G. Bellanger, J. L. Daguet, and G. P. Lepagnol, "Interpolation, extrapolation, and reduction of computation speed in digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, pp. 231–235, Aug. 1974.