A Class of Cosine-Modulated Causal IIR Filter Banks

Linnéa Svensson, Per Löwenborg, and Håkan Johansson
Department of Electrical Engineering
Linköping University, SE-581 83 Linköping, Sweden
{linneas, perl, hakanj}@isy.liu.se

ABSTRACT

This paper introduces a class of cosine modulated causal IIR filter banks. In the simplest case, the filter banks have near perfect magnitude reconstruction (NPMR). By introducing a phase equalizer one can also obtain near perfect reconstruction (NPR) filter banks. By using filter banks with NPR instead of NPMR (NPR) filter banks have near perfect magnitude reconstruction (NPMR). By introducing a phase equalizer one can also obtain NPR filter banks. Cosine modulated filter banks are known to be very efficient since each of the analysis and synthesis parts can be implemented with the aid of one filter and a discrete cosine transform. By including IIR filters, the complexity and delay can be reduced as compared to the case where only FIR filters are used. Earlier, only PR and NPR cosine modulated filter banks have been considered, both in the FIR case (see e.g. [2], [3], [7]) and IIR case (see e.g. [5], [6]). By using instead the proposed NPMR banks, the complexity can be reduced. A prerequisite is of course that the application at hand allows phase distortion. Finally, we point out that NPR IIR filter banks can alternatively be efficiently implemented using tree-structures [8], [2]. Tree-structured filter banks suffer from some drawbacks though. For example, in the uniform-band case, the number of channels is restricted to be a power of two. This problem is overcome by employing modulated filter banks which do not have such limitations.

1. INTRODUCTION

Maximally decimated filter banks find application in numerous areas [1]–[3]. Over the past two decades, a vast number of papers on the theory and design of such filter banks have been published. The attention has to a large extent been paid to the problem of designing perfect reconstruction (PR) filter banks. In a PR filter bank, the output sequence of the overall system is simply a shifted version of the input sequence. However, filter banks are most often used in applications where small errors (emanating from quantizations etc.) are allowed. Imposing PR on a filter bank is then an unnecessarily severe restriction which may lead to a higher arithmetic complexity than is actually required to meet the specification at hand. To reduce the complexity one should therefore consider non-PR filter banks. For example, it is demonstrated in [4] that the complexity can be significantly reduced by relaxing the PR restriction.

The relation between the input and output of an $M$-channel maximally decimated filter bank can be described by a distortion transfer function and $M-1$ aliasing transfer functions. The distortion function should approximate one in magnitude whereas the aliasing functions must be small in the frequency band of interest (normally the whole band from 0 to $\pi$). In many applications, it is also desired that the distortion function have an approximately linear phase response. In this case the filter bank is a near perfect reconstruction (NPR) filter bank. If the phase response is not approximately linear, we say that it is a near perfect magnitude reconstruction (NPMR) filter bank. (The filter bank is PMR if the magnitude responses of the distortion functions, and given by

$$P(z) = \frac{A(z)}{C(z^{2M})}$$

where $A(z)$ and $C(z)$ are of order $N_A$ and $N_C$, respectively, and given by

$$A(z) = \sum_{n=0}^{N_A} a(n)z^{-n}, \quad C(z) = 1 + \sum_{n=1}^{N_C} c(n)z^{-n}$$

Further, $A(z)$ is a linear-phase FIR filter with symmetric impulse response $a(n) = a(N_A - n)$. The analysis filters $H_0(z)$ and synthesis filters $G_0(z)$ are obtained by modulation of the prototype filter $P(z)$ according to

2. PROPOSED FILTER BANK CLASS

This section gives transfer functions, frequency responses, and some properties of the proposed filter banks.

2.1. Filter Transfer Functions

Let the transfer function of the lowpass prototype filter be given as

$$P(z) = \frac{A(z)}{C(z^{2M})}$$

where $A(z)$ and $C(z)$ are of order $N_A$ and $N_C$, respectively, and given by

$$A(z) = \sum_{n=0}^{N_A} a(n)z^{-n}, \quad C(z) = 1 + \sum_{n=1}^{N_C} c(n)z^{-n}$$

Further, $A(z)$ is a linear-phase FIR filter with symmetric impulse response $a(n) = a(N_A - n)$. The analysis filters $H_0(z)$ and synthesis filters $G_0(z)$ are obtained by modulation of the prototype filter $P(z)$ according to
polyphase components of tor polynomial. A realization of the from which the distortion function is obtained as

From (4) and (6) it follows that

where \( \omega \) and \( \omega \) respectively, are given by

where \( \) and \( \) for .

By using (1)

and a distortion function without bumps.

The magnitude of the distortion and aliasing functions are they are as illustrated for

Further, the impulse responses, \( a_k(n) \) and \( b_k(n) \), for \( A_k(z) \) and \( B_k(z) \), respectively, are given by

The filters \( H_k(z) \) and \( G_k(z) \) thus share the same denominator polynomial. A realization of the filter bank can be seen in Fig. 2, where \( G_0(z) \), \( p = 0, \ldots, M−1 \), denote the polyphase components of \( A(z) \). Further, the impulse responses, \( a_k(n) \) and \( b_k(n) \), for \( A_k(z) \) and \( B_k(z) \), respectively, are given by

\[
a_k(n) = 2a(n)\cos\left(\frac{k+1}{M^2}\left(n - \frac{N_A}{2}\right) + \theta_k\right)
\]

\[
b_k(n) = 2a(n)\cos\left(\frac{k+1}{M^2}\left(n - \frac{N_A}{2}\right) - \theta_k\right)
\]

2.2. Distortion and Aliasing Functions

For the \( M \)-channel maximally decimated filter bank in Fig. 1 the \( z \)-transform of the output signal is given by

\[
Y(z) = \sum_{m=0}^{M-1} V_m(z) X(z W_{M^2}).
\]

\[
V_m(z) = \sum_{k=0}^{M-1} H_k(z) G_k(z)
\]

From (4) and (6) it follows that \( B_k(z) = z^{-N_A} A_k(z^{-1}) \) from which the distortion function is obtained as

The frequency response of the distortion function can thus be written as

\[
V_0(e^{j\omega T}) = \frac{1}{|C(-e^{j2M\omega T})|^2} \sum_{k=0}^{M-1} |A_k(e^{j\omega T})|^2
\]

where \( \Phi_V(\omega T) = -N_A \omega T - 2\Phi_C(2M\omega T) \).

We note that the non-linear phase of \( V_0(e^{j\omega T}) \) is due to \( \Phi_C(2M\omega T) \).

2.3. Some Properties

The proposed filter bank can be viewed as a generalization of regular cosine modulated FIR filter banks [2], to which the former reduces as a special case when \( C(z) = 1 \). The main difference is the introduction of the denominator polynomial \( C(z) \) which improves the frequency selectivity but also makes the phase response of the distortion function nonlinear. Specifically, it can be shown that the following properties remains when going from the FIR to our IIR case.

1) Adjacent-channel aliasing is cancelled. That is, the largest remaining aliasing terms are in the worst case twice the stopband ripple of \( H_k(z) \).

2) The distortion function can be expressed as \( V_0(z) = z^{-N_A} f(z^{-M}) \). This follows immediately since it is true in the FIR case and the only extra factor in our IIR case is \( 1/C(-z^{-2M}) \). In particular, this means that \( |V_0(e^{j\omega T})| \) has period \( \pi M \).

3) From (1) – (6) we get

\[
|H_k(z)|^2 = |P(z W_{M^2})|^2 + |P(z^{-M^2})|^2
\]

for \( z = e^{j\omega T} \) which is utilized when designing the filter bank. Specifically, the two terms on the right-hand side of the above equation must be approximately power complementary in order to obtain lowpass and highpass filters and a distortion function without bumps.

3. FILTER BANK DESIGN

This section introduces a technique to design the proposed IIR filter banks. We consider both NPMR and NPR. 3.1. NPMR Design

Let the specifications of \( H(z) \) be

\[
1 - \delta_c \leq |H_k(e^{j\omega T})| \leq 1 + \delta_c, \quad \omega T \in \Omega_{c,k}
\]

where \( \Omega_{c,k} \) and \( \Omega_{c,k} \), respectively, are the passband and stopband regions of \( H(z) \). Expressed with the aid of \( \Delta, \) they are as illustrated for \( M = 3 \) in Fig. 3.

The magnitude of the distortion and aliasing functions are to meet

\[
1 - \delta_0 \leq |V_0(e^{j\omega T})| \leq 1 + \delta_0, \quad \omega T \in [0, \pi]
\]
Figure 3. Passband and stopband regions for $M = 3$.

$$V_m(e^{j\omega T}) \leq \delta_1, \quad \omega T \in [0, \pi], \quad m = 1, ..., M - 1 \quad (15)$$ respectively. To meet the specifications we divide the optimization into two steps.

**Step 1:** In this step the distortion function is minimized while the magnitude responses of the analysis filters in the passband and stopband regions are kept within the their specified limits. To this end, we solve the following optimization problem:

- **minimize** $\delta$

- **subject to**
  
  $$|H_k(e^{j\omega T})| - 1| \leq \delta_{\epsilon} \delta_{\theta}, \quad \omega T \in \Omega_{c,k}$$
  
  $$|H_k(e^{j\omega T})| \leq \delta_{\epsilon} \delta_{\theta}, \quad \omega T \in \Omega_{s,k}$$
  
  $$|V_0(e^{j\omega T})| - 1| \leq \delta, \quad \omega T \in [0, \pi] \quad (16)$$

The specifications in (13) and (14) are met when $\delta \leq \delta_0$.

**Step 2:** This step minimizes the maximum aliasing magnitude of the $M - 1$ aliasing functions, while keeping the magnitude responses of the analysis filters as well as the distortion function within their specified limits. The following optimization is then to be solved:

- **minimize** $\delta$

- **subject to**
  
  $$|H_k(e^{j\omega T})| - 1| \leq \delta_{\epsilon} \delta_{\theta}, \quad \omega T \in \Omega_{c,k}$$
  
  $$|H_k(e^{j\omega T})| \leq \delta_{\epsilon} \delta_{\theta}, \quad \omega T \in \Omega_{s,k}$$
  
  $$|V_0(e^{j\omega T})| - 1| \leq \delta, \quad \omega T \in [0, \pi] \quad (17)$$

The specifications in (13)–(15) are met when $\delta \leq \delta_1$.

The problems above are non-linear optimization problems and therefore require good initial solutions. The solution of (16) is a good initial solution to (17), since if (13) is satisfied so is (15) with $\delta_{\epsilon} = 2\delta_{\theta}$. A good initial solution to (16) can be found by properly designing the prototype filter $P(z)$. Utilizing (11), and ignoring higher-order terms such as squared stopband ripples, it can be shown that (13) and (14) are satisfied if the prototype filter fulfills

$$1 - \delta_{\epsilon} \leq |P(e^{j\omega T})| \leq 1 + \delta_{\epsilon}, \quad \omega T \in \left[0, \frac{\pi}{2M} - \Delta\right] \quad (18)$$

and

$$E(z) = \frac{F(z)}{C(z^2)} \quad (20)$$

(21)

The functions $F(z)$ and $zC(z^2)$ are the numerator and denominator polynomials, respectively, of an $N_F$th-order half-band FIR filter, where $N_F$ is odd. In other words, $C(z^2)$ is the denominator polynomial less the pole in the origin. The polynomial $F(z)$ is thus an odd-order linear-phase FIR filter whereas $C(z^2)$ is an even-order filter. Further, $D(z)$ is an $N_D$th-order linear-phase FIR filter. Thereby, $A(z)$ will be a linear-phase FIR filter of order $N_A = MN_F + N_D$. The functions $E(z)$ and $D(z)$ are referred to as model and masking filters respectively. The role of $D(z)$ is to remove the $M - 1$ images present due to the factor $z^M$ in the polynomial $E(z^M)$ in (20) [11]. In this way, the prototype filter can be made narrow-band and at the same time be implemented with a low computational complexity, as will be illustrated in Section 4.

A simple way to meet (18) is to design $E(z)$ and $D(z)$ to fulfill

$$|E(e^{j\omega T})| \leq \delta_{\epsilon} / \sqrt{2}, \quad \omega T \in \left[\frac{\pi}{2} + M\Delta, \pi\right] \quad (22)$$

with $E(1) = 1$, and

$$1 - \delta_{\epsilon} \leq |D(e^{j\omega T})| \leq 1 + \delta_{\epsilon}, \quad \omega T \in \left[0, \frac{\pi}{2M} + \Delta\right] \quad (23)$$

The filters $E(z)$ and $D(z)$ are designed with conventional methods [9], [10] to fulfill (22) and (23). This ensures that the requirements (13) and (14) are fulfilled but after optimization (16) in Step 1 the filters are in general overdesigned. Therefore, the requirements on the initial filters, used as starting points in (16), are successively relaxed in order to find the minimum complexity (here computed as $2(N_A + 1)/M + 2N_C$ which is the number of multiplications per input/output sample for the overall filter bank) required to meet (13)–(14). The Step 1 design gives a lower bound on the overall complexity. It may have to be slightly increased though since aliasing is ignored in this step. Constraints on the aliasing functions are added in Step 2, where the optimization problem (17) is solved. If the minimum complexity obtained in Step 1 is too low to meet also the additional aliasing requirements, we successively increase the filter orders and redo Step 1 and Step 2 until we find the minimum complexity required to meet all the specifications (13)–(15).

### 3.2. NPR Design

In the design procedure above, the phase response of the distortion function was ignored which means that it may be too nonlinear for the application at hand. One simple way to improve the phase linearity is to equalize it with the aid of an allpass filter. Since such a filter has a constant magnitude response for all frequencies it will not affect the magnitude response of the distortion and aliasing functions. It thus suffices to consider only the phase response in this step.

Here, it follows from (12) that the nonlinearity in the phase response $\Phi_{\text{eq}}(\omega T)$ emanates from $\Phi_{\text{eq}}(2\pi M T)$ which is the phase response of the denominator polynomial $C(\pi z^2)$. The problem thus reduces to that of
equalizing $\Phi_A(2MoT)$. Further, since this is a function with a periodicity of $\pi/M$, we use an allpass filter with transfer function $H_{AP}(z^{2M})$ as equalizer. The allpass filter is designed so as to minimize the maximum magnitude of the phase error $\Phi_{AP}(oT)$ which is given by

$$\Phi_{AP}(oT) = \Phi_V(oT) + \Phi_A(2MoT) + K o T \quad (24)$$

where $\Phi_A(2MoT)$ is the phase response of $H_{AP}(z)$. The overall filter bank delay after the allpass equalization is $K = N_A + 2MN_{AP}$ where $N_{AP}$ is the order of $H_{AP}(z)$. It is well known that the equiripple solution minimizes $|\Phi_A(oT)|$. There are many different techniques available to find this solution. We use the algorithm in [12].

4. DESIGN EXAMPLE

As a means of demonstrating the proposed design method, a 5-channel cosine modulated filter bank is designed. The specifications of $H_{AP}(z)$ and $V_m(z)$ in (13) – (15) are the following: $\delta^e = 0.05$, $\delta^s = 0.02$, $\delta_0 = 0.01$, $\delta_\delta = 10^{-5}$, and $\Delta = 0.02\pi$.

In the Step 1 design (ignoring aliasing), the specification is met with $N_S = 30$ and $N_C = 2$. In Step 2, constraints are added on the aliasing functions and their maximum is thereby reduced from $\approx 31$ dB to $\approx 66$ dB. To fulfill the specification $\delta^e$ in (15), the order $N_S$ is increased to $N_S = 34$, which results in aliasing terms less than $\approx -123.3$ dB. The magnitude responses of the analysis filters, distortion function, and aliasing functions are plotted in Figs. 4 – 6. The overall filter bank (including the analysis and synthesis parts) requires 18 multiplications per input/output sample plus the cost to implement the cosine transforms. For a regular NPR FIR cosine modulated filter bank (optimized in the same way as our IIR filter bank), the filter part requires $N = 123$ and thus $49.6$ multiplications. We should keep in mind though that the distortion function in our IIR filter bank has a nonlinear phase response whereas in the FIR filter bank is linear. To improve the phase linearity we introduce an allpass phase equalizer according to Section 3.2. Figure 7 plots the maximum phase error as a function of the allpass filter order. For instance, using a 10th-order allpass filter, the maximum phase error becomes $0.055$. The overall multiplication ratio is in this case 28 (18+10) which is still far below that of the FIR filter bank. If very small phase errors are required, the allpass order may have to become quite large in which case a regular FIR filter bank may be in favor.

5. CONCLUSION

In this paper we introduced a class of cosine modulated causal IIR filter banks. In the simplest case, they are NPMR filter banks. By introducing phase errors in the distortion function the arithmetic complexity can be reduced as demonstrated by means of a design example.

REFERENCES


1. The reason why we do not compare our filter banks with existing cosine modulated IIR filter banks is that we simply have not found in the literature reliable techniques for designing those filter banks to satisfy the general specifications that we deal with in this paper.