Two-Channel IIR QMF Banks with Approximately Linear-Phase Analysis and Synthesis Filters

M. M. Ekanayake & K. Premaratne
Department of Electrical & Computer Engg.
University of Miami
Coral Gables, FL 33124

Abstract

In this paper, a computationally simple method to obtain IIR analysis and synthesis filters that possess negligible phase distortion is presented. It is shown that for a certain class of FIR filter banks, the polyphase components are approximately all-pass. By applying the balanced reduction procedure to the polyphase components of such a FIR prototype filter all-pass IIR polyphase components are obtained. The resulting IIR designs have negligible phase distortion. Design examples indicate that the derived IIR filter banks can be implemented more efficiently than the FIR prototypes. Results show that the IIR filters are better than linear phase perfect reconstruction systems when implemented in a finite wordlength processor.

1 Introduction

Quadrature mirror filters (QMF's) are used in many speech and communications applications. A two-channel QMF bank is shown in Fig. 1 where $H_0(z)$, $H_1(z)$ are the transfer functions of the analysis bank filters and $F_0(z)$ and $F_1(z)$ are the synthesis filters. The reconstructed signal, in general, suffers from aliasing distortion (ALD), amplitude distortion (AMD) and phase distortion (PHD) due to the fact that the analysis and synthesis filters are not ideal. A common requirement in most applications is that the reconstructed signal $\hat{x}(n)$ should be as close to $x(n)$ as possible in some well defined sense. However other constraints are usually imposed to reduce nonlinear distortions, such as coding errors and transmission channel distortions, that cannot be directly evaluated. One such constraint that is usually imposed is that the analysis/synthesis filters be linear-phase [1].

There are two types of Finite Impulse Response (FIR) filter banks which have the linear phase property. In Type I systems only ALD and PHD are eliminated while AMD is minimized. Type II systems have the Perfect Reconstruction (PR) property where all ALD, AMD and PHD are eliminated.

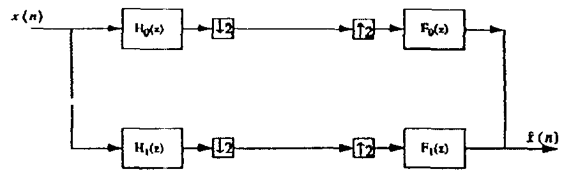

Figure 1: Two-channel QMF bank

In Infinite Impulse Response (IIR) filter banks usually only ALD and AMD are eliminated. PHD is minimized by using a separate AP equalizer network once the signal is reconstructed [1]. Due to the lack of linear phase property in these IIR filter banks they are not suitable for linear phase applications. In [2] the eigenfilter approach has been used to design IIR filter banks with approximate linear phase characteristics in the passband. However since these filters exhibit exessive PHD in the transition band a postprocessing equalizer network is necessary at the output. An alternative approach to design IIR filter banks that has not been investigated so far is to eliminate ALD and AMD and minimize PHD. As we shall show, PHD can be minimized by constraining the filters to have approximately linear phase in the entire baseband.

In this paper, a computationally efficient method for the design of IIR filter banks, where AMD and ALD are eliminated and PHD is minimized, is presented. Each of the analysis and synthesis filters are designed to have approximately linear phase characteristics in the entire baseband. Hence the overall PHD of the filter bank is also negligible.

1058-6393/95 $4.00 © 1995 IEEE

2 Design of Linear-Phase IIR Filter Banks

Design of optimum filters involves the matching of its frequency response with that of a prescribed frequency response. In IIR filter designs usually an FIR filter is used as a prototype. Hence the properties of the FIR prototype filter must be close to that of the desired properties of the IIR filter.

In IIR filter banks, it is desirable to have the polyphase components as AP filters, since with this choice, AMD is eliminated. Moreover, AP filters can be implemented very efficiently and have very low coefficient sensitivity. Hence one of the desired properties of the FIR prototype filter is that its polyphase components be approximately AP. The polyphase representation of Type I filter bank analysis filters $H_0(z)$ and $H_1(z)$ is given by

$$H_0(z) = E_0(z^2) + z^{-1}E_1(z^2) \tag{2.1a}$$
$$H_1(z) = E_0(z^2) - z^{-1}E_1(z^2) \tag{2.1b}$$

To eliminate ALD [1] the synthesis filters are given by $F_0(z) = H_0(z)$ and $F_1(z) = -H_0(-z)$.

Theorem 2.1 *The polyphase components of a Type I system are approximately AP with*

$$|E_0(e^{j\omega})| = |E_1(e^{j\omega})| \approx \frac{1}{2} \tag{2.2}$$

Proof:
It can be shown [1][3] that for a Type I system

$$|E_0(e^{j\omega})| = |E_1(e^{j\omega})| \quad \text{and} \tag{2.3}$$
$$|T(e^{j\omega})| = 2|E_0(e^{j2\omega})||E_1(e^{j2\omega})| \tag{2.4}$$

where $\hat{X}(z) = T(z)X(z)$. In Type I systems optimization is done to reduce AMD so that $|T(e^{j\omega})| \approx \frac{1}{2}$. Hence the claim follows. ■

However for Type II filter banks this is not true. Hence a Type I filter bank is the best candidate as the prototype filter.

The polyphase representation of IIR filter bank analysis filters $\bar{H}_0(z)$ and $\bar{H}_1(z)$ is given by

$$\bar{H}_0(z) = A_0(z^2) + z^{-1}A_1(z^2) \tag{2.5a}$$
$$\bar{H}_1(z) = A_0(z^2) - z^{-1}A_1(z^2) \tag{2.5b}$$

where $A_0(z)$ and $A_1(z)$ are allpass and thus AMD is eliminated. To ensure that ALD is eliminated, we choose the synthesis filters of the IIR filter bank in the same way as for Type I FIR filters.

We now show that minimizing PHD can be achieved by constraining the analysis filters to be approximately linear phase in the entire baseband.

Theorem 2.2 *For a two-channel IIR filter bank with AP polyphase components, the analysis filter $H_0(z)$ will be approximately linear phase iff $T(z)$ is approximately linear phase.*

Proof:

Let $A_0(e^{j\omega}) = e^{j\phi_0(\omega)}$ and $A_1(e^{j\omega}) = e^{j\phi_1(\omega)}$

$$\begin{aligned} \text{Hence } H_0(z) &= \frac{1}{2}[A_0(z^2) + z^{-1}A_1(z^2)] \\ &= \frac{1}{2}[e^{j2\phi_0(\omega)} + e^{-j\omega}e^{j2\phi_1(\omega)}] \\ &= cos(\phi_0 - \phi_1 + \frac{\omega}{2})e^{j(\phi_0+\phi_1-\frac{\omega}{2})} \end{aligned} \tag{2.6}$$

$$\begin{aligned} \text{Also} \quad T(e^{j\omega}) &= 2e^{-j\omega}A_0(e^{j2\omega})A_1(e^{j2\omega}) \\ &= e^{j(2\phi_0(\omega)+2\phi_1(\omega)-\omega)} \end{aligned} \tag{2.7}$$

Therefore the claim follows from (2.6) and (2.7). ■

The most straightforward method to design the IIR filter based on the prototype filter is to optimize the coefficients of $A_0(z)$ and $A_1(z)$ to match their frequency responses to those of $E_0(z)$ and $E_1(z)$ respectively. However this method turns out to be computationally inefficient and sometimes numerically unstable. An alternative method that has been used for the design of approximately linear phase IIR filters is the Balanced Reduction (BR) procedure. This method yields a nearly optimum IIR filter from a FIR filter. Therefore the method we adopt to design the IIR filter bank is:

- Choose a Type I filter as the prototype.
- Apply the BR procedure firstly to $E_0(z)$ to obtain $A_0(z)$ and secondly to $E_1(z)$ to obtain $A_1(z)$.

3 The Balanced Reduction Procedure

As proposed by Moore [4], the BR procedure is a very attractive procedure to derive a reduced order model from a given high order system. In this method, the given state space (s.s) formulation is transformed into a coordinate system wherein each state is as reachable as it is observable. This transformed system is called *balanced*, and by deleting the least reachable and observable states, a reduced model of the original results. For the purpose at hand, since the higher-order system is nearly AP, the lower order subsystem that we obtain is also nearly AP.

3.1 The Balanced Realization

Let (A, B, C, D) be a minimal realization of a stable transfer function $H(z)$ of order m. The two positive definite matrices P and Q, which are called the reachability and observability gramians can be found by solving the pair of Lyapunov equations given by

$$P - APA^T = BB^T \tag{3.1a}$$

$$Q - A^TQA = C^TC \tag{3.1b}$$

The Hankel singular values of the system $\sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_m^2$ are defined as the eigenvalues of the positive definite matrix PQ. A non-singular similarity transformation T yields the similar system $(\hat{A}, \hat{B}, \hat{C}, D)$ where

$$\hat{A} = TAT^{-1}, \quad \hat{B} = TB, \quad \hat{C} = CT^{-1} \tag{3.2}$$

The transfer function and its Hankel singular values are invariant under a nonsingular similarity transformation. It is well known [4] that there exists a non-singular matrix T such that the similar system $(\hat{A}, \hat{B}, \hat{C}, D)$ obtained as in (3.2) is a balanced realization in the sense that the corresponding gramians are diagonal and identical, that is $\hat{P}$ and $\hat{Q}$, of the system $(\hat{A}, \hat{B}, \hat{C}, D)$ takes the form

$$\hat{Q} = \hat{P} = \Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_m) \tag{3.3}$$

Any realization satisfying (3.3) is called a balanced realization of $H(z)$. For an efficient method to compute the balanced realization, see the routine 'dbalreal' in the MATLAB control system toolbox [5].

3.2 The Balanced Approximation of the Polyphase Components

The first step in the BR procedure is to obtain a balanced realization of $H(z)$. The key to the reduction procedure is the matrix Σ. Let Σ be decomposed into two parts;

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \tag{3.4}$$

$$\Sigma_1 = diag(\sigma_1, \ldots, \sigma_{\hat{m}}) \text{ , } \Sigma_2 = diag(\sigma_{\hat{m}+1}, \ldots, \sigma_m)$$

We can represent the balanced realization according to the partition (3.4) as

$$\hat{A} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}, \hat{B} = \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \end{bmatrix}, \hat{C} = \begin{bmatrix} \hat{C}_1 & \hat{C}_2 \end{bmatrix} \tag{3.5}$$

The stable system [4] $(\hat{A}_{11}, \hat{B}_1, \hat{C}_1, D)$ of order $\hat{m}$ represents a good lower-order approximation of the original system if $\sigma_{\hat{m}} \gg \sigma_{\hat{m}+1}$.

Starting with a nearly AP FIR filter, by applying the balanced reduction procedure, we obtain a reduced order IIR filter which is also nearly AP. Therefore the numerator polynomial is nearly a mirror image of the denominator. Now since we are looking for AP polyphase components, we must force the nearly AP transfer function to be AP. Hence we choose the numerator polynomial as the mirror image of the denominator polynomial.

The balanced approximation is not the optimal lower order approximation, but we choose it because of its computational advantage. Also the AP approximation to a nearly AP IIR filter is not optimal. In practice this method yields filters which are nearly optimal. Little or no improvement can be obtained with optimization.

4 Design Examples

In this section, we demonstrate the application of the proposed method to design an IIR filter bank with each analysis filter having a transition bandwidth, $\Delta f = 0.086$ and a stop-band attenuation of 65dB.

Example:
The design of FIR filters with the desired properties referred to in this paper have been presented earlier by Johnston [6]. We choose a filter which satisfies the above specifications from the appendix of [6]. This turns out to be a filter of length 64, which is referred to as the 64D filter.

We now find the balanced realization for each polyphase component. The following were observed for the gramians.

$$\sigma_{0,1} = 0.50000 \ \sigma_{0,16} = 0.50000 \ \sigma_{0,17} = 0.0001$$
$$\sigma_{1,1} = 0.50000 \ \sigma_{1,15} = 0.50000 \ \sigma_{1,16} = 0.0001$$

Therefore the orders of the reduced 0^{th} and 1^{st} polyphase components are 16 and 15 respectively. After truncating according to (3.4) the two denominator polynomials $\check{d}_0(z)$ and $\check{d}_1(z)$ shown in Table 1 were obtained. We choose the numerator polynomial of each polyphase component to be the mirror image of the respective denominator polynomials. The magnitude response and the group delay of the IIR filter $\check{H}_0(z) = 0.5[\check{A}_0(z^2) + z^{-1}\check{A}_1(z^2)]$ is shown in Fig2. The computer time taken for the BR procedure is approximately 0.65 seconds on a DEC 5000 workstation using the MATLAB control system toolbox.

Table 1: Polyphase component denominator polynomials for the IIR filter in the example.

n	$d_0(n)$	$d_1(n)$
0	1.00000000000000	1.00000000000000
1	0.24510479134270	-0.24510331217525
2	-0.08540446106043	0.14547975580225
3	0.04372217148846	-0.10031250245546
4	-0.02473518071621	0.07246314200969
5	0.01413313848422	-0.05288462490705
6	-0.00769194123910	0.03829127599374
7	0.00369791714860	-0.02719069757361
8	-0.00128353807212	0.01876589218446
9	-0.00009405863941	-0.01245789864033
10	0.00068111709418	0.00794962259068
11	-0.00089067178133	-0.00477740314111
12	0.00082093773822	0.00267592696004
13	-0.00062700877872	-0.00136646062031
14	0.00039360412931	0.00062174841299
15	-0.00019341002057	-0.00023563382949
16	0.00008049208892	-

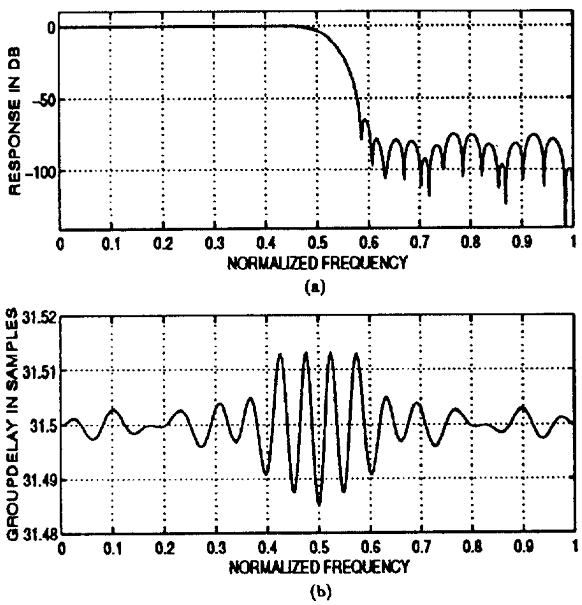

Figure 2: Frequency response of the IIR filter (a) Magnitude response and (b) groupdelay response

4.1 Implementation in Finite Wordlength

The IIR filter designed in the example, the corresponding Type I filter and the 64 length Type II filter designed in [7] were implemented in floating point arithmetic with 12-bits in the mantissa. The Type II filter was implemented with the 1-multiplier lattice structure. The frequency responses of the three filters in Fig. 3. shows that the Type I and IIR filters have good stopband characteristics whareas the Type II filter stopband characteristics have deteriorated.

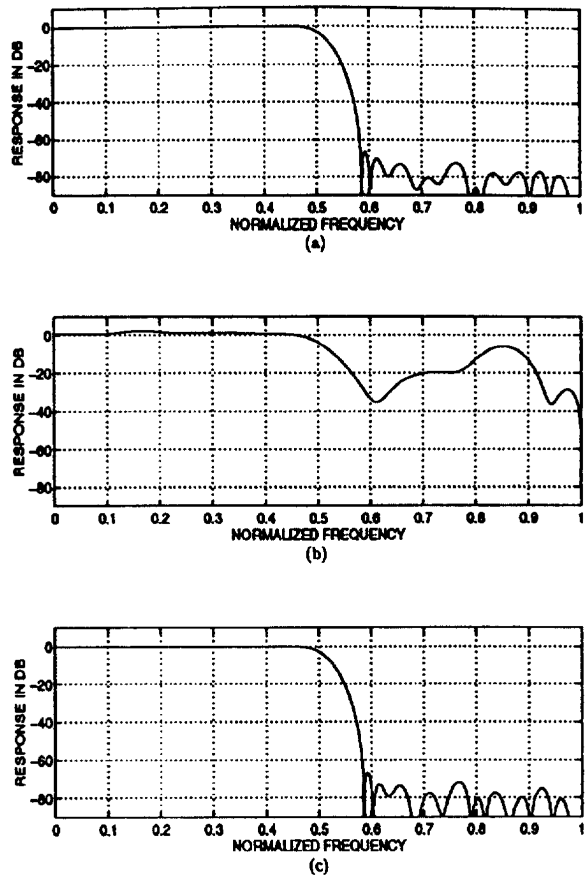

Figure 3: Magnitude response of the filters in FWL. (a) Type I, (b) Type II and (c) IIR.

Although theoretically, the Type II filter is a PR system, with finite word length (FWL) it becomes a non PR system. This is clearly seen by the magnitude response of the Type II system overall transfer function $|T(e^{j\omega})|$ shown in Fig. 4.

To investigate the reconstruction error when these filters are implemented in a FWL processor, a white noise signal with a flat spectrum and random phase was input to the filter bank and the signal was reconstructed at the output. The signal-to-noise ratio was computed using

$$SNR = \frac{\sum_{n=0}^{N-1} x^2(n)}{\sum_{n=0}^{N-1} (x(n) - \hat{x}(n))^2} \tag{4.1}$$

The results are tabulated in Table 2.

4.2 Comparison with FIR filters

An AP filter of order p with real coefficients can be implemented with p multipliers and $2p$ adders

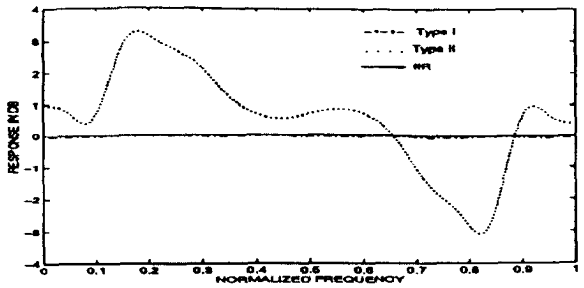

Figure 4: Magnitude response of the overall transfer function $T(e^{j\omega})$

[1]. However with the polyphase implementation a polyphase component of order p is computing at half rate and hence needs only $p/2$ multiplications per unit time (MPU's) and p additions per unit time (APU's). Hence our analysis bank needs only $(16+15)/2 = 15.5$ MPU's and $(31*2+2)/2 = 32$ APU's. Table 2 gives a comparison of the IIR filter with FIR filters. The infinite wordlength (IWL) implementation is actually with 26 bits in the mantissa. The data for the Type II filter is based on [7].

Table 2: Comparison of IIR with FIR filter banks.

Feature		Type I FIR	Type II FIR	IIR
No of MPU's for Analysis bank		32	17	15.5
No of APU's for Analysis Bank		32	49	32
Average group delay (samples)		63	63	63
Stop Band Atten.	- IWL	65 dB	42 dB	65 dB
	- FWL	68 dB	06 dB	66.5 dB
AMD	- IWL	0.002 dB	0 dB	0 dB
	- FWL	0.006 dB	3.320 dB	0 dB
Delay distortion (samples)	- IWL	0	0	±0.0125
	- FWL	0	0	±0.0309
SNR	- IWL	144 dB	179 dB	143 dB
	- FWL	125 dB	21 dB	125 dB

Based on the information in the Table 2 the IIR is better than the Type I and Type II filters in every respect excepting the group delay distortion. When compared with the Type I filters, the IIR filter would be much better when efficient implementation is necessary since the MPU count is very much less for the latter while the same stopband attenuation is achieved. The poor SNR and poor stopband attenuation of the Type II filters indicate that they have to be implemented with the 2-multiplier lattice structures. Hence again the IIR filter bank will be much more efficient compared to a Type II system. The price paid for the better efficiency is the small group delay distortion. However since the group delay is very small, for most applications this is bound to be acceptable.

5 Conclusion

We have presented an algorithm to design IIR filter banks based on FIR prototypes. The analysis filters of the IIR filter bank have approximately linear-phase. It was shown that the Type I FIR filter family was the most suitable prototype filter. Design examples demonstrated the application of the algorithm and indicated the computational advantage of the IIR filter bank compared to FIR designs. The group delay distortion of the filter bank is so small that for most applications an IIR implementation seems to be better than an FIR implementation due to the computational advantage.

References

[1] P.P. Vaidyanathan, *Multirate Systems and Filter Banks.* Englewood Cliffs, NJ: Prentice Hall, 1993.

[2] T.Q. Nguyen, T.I. Laakso and R.D. Koilpillai, "Eigenfilter approach for the desgn of allpass filters approximating a given phase response," *IEEE Trans. Signal Proc.*, Sept. 1994.

[3] M.M. Ekanayake, *An Algorithm for the Design of Approximately Linear Phase IIR Filter Banks,* Master of Science Thesis, Department of Electrical & Computer Engineering, Universilty of Miami, Coral Gables, FL 33124, May 1994.

[4] B.C. Moore, "Principal Component Analysis in Linear Systems : Controllability, Observability and Model Reduction", *IEEE Trans. Auto. Control*, vol. AC-26, No 1, pp17-32, Feb. 1981.

[5] Pro-MATLAB Users Guide, *MathWorks, Inc.*

[6] J.D. Johnston, "A filter family designed for use in quadrature mirror filter banks,"in *Proc. IEEE Int. Conf. ASSP*, Apr. 1980, pp 291-294.

[7] T.Q. Nguyen and P.P. Vaidyanathan, "Two-channel perfect reconstruction FIR QMF structures which yield linear phase FIR analysis and synthesis filters," *IEEE Trans. Acoust. Speech Signal Proc.*, vol. 37, pp 676-690, May 1989.