

New method for designing two-channel causal stable IIR perfect reconstruction filter banks and wavelet bases

J. S. Mao
S. C. Chan
K. L. Ho

The University of Hong Kong
Department of Electrical and Electronic
Engineering
Pokfulam Road
Hong Kong
E-mail: jmao@engr.uvic.ca
klho@eee.hku.hk
scchan@eee.hku.hk

Abstract. A new method for designing two-channel causal stable IIR PR filter banks and wavelet bases is proposed. It is based on the structure previously proposed by Phoong et al. (1995). Such a filter bank is parameterized by two functions $\alpha(z)$ and $\beta(z)$, which can be chosen as an all-pass function to obtain IIR filterbanks with very high stopband attenuation. One of the problems with this choice is that a bump of about 4 dB always exists near the transition band of the analysis and synthesis filters. The stopband attenuation of the high-pass analysis filter is also 10 dB lower than that of the low-pass filter. By choosing $\beta(z)$ and $\alpha(z)$ as an all-pass function and a type-II linear-phase finite impulse response (FIR) function, respectively, the bumping can be significantly suppressed. In addition, the stopband attenuation of the high-pass filter can be controlled easily. The design problem is formulated as a polynomial approximation problem and is solved efficiently by the Remez exchange algorithm. The extension of this method to the design of a class of IIR wavelet bases is also considered. © 2000 Society of Photo-Optical Instrumentation Engineers. [S0091-3286(00)02909-3]

Subject terms: filter bank; wavelet; causal stable; IIR; perfect reconstruction.

Paper 990264 received July 6, 1999; revised manuscript received Mar. 30, 2000; accepted for publication Apr. 7, 2000.

1 Introduction

Perfect reconstruction (PR) multirate filter banks have important applications in signal analysis, coding, and the design of wavelet bases. Figure 1(a) shows the block diagram of a two-channel maximally decimated filter bank. The system is called a PR system if the output $\hat{x}(n)$ is identical to the input $x(n)$ except for some constant scaling and time delay. The theory of PR filter banks has been extensively studied.¹⁻⁴ In finite impulse response (FIR) filter banks, all the analysis filters and the synthesis filters are FIR filters and the PR condition is considerably simplified. More precisely, the filter bank is PR if the determinant of its polyphase matrix is equal to some delay. In IIR filter banks, the entries of the polyphase matrix become rational functions and the system is PR if the determinant of its polyphase matrix is a minimum phase function. In addition to the more complicated PR condition of IIR filter banks, it is also very difficult to ensure that the IIR filters be causal stable. Early attempts typically have noncausal stable filters or causal unstable filters.⁸⁻¹⁰ In Refs. 5 and 6, causal stable IIR PR filter banks are designed by using factorization of the polyphase matrix. However, satisfactory design results are not obtained. Design of causal stable IIR PR filter bank using optimization techniques has also been proposed in Ref. 11 where the Lagrange multiplier and Lagrange-Newton methods are used to perform the optimization. However, satisfactory results are not obtained and the stability of the filters cannot be guaranteed. One of the most successful designs of causal stable IIR PR filter banks is the one reported in Ref. 12, where a new structure for two-

channel IIR PR filter bank was proposed. Such structurally PR filter banks is parameterized by two functions $\alpha(z)$ and $\beta(z)$, which can be chosen as an all-pass function to obtain IIR filter banks with very high stopband attenuation. The design procedure is also very simple. One of the problems with this choice of $\alpha(z)$ and $\beta(z)$ is that a bump of about 4 dB always exists near the transition band of the analysis and synthesis filters. Moreover, the stopband attenuation of the low-pass analysis filter is always 10 dB higher than that of the high-pass filter. To overcome these problems, an improved algorithm using general rational functions for $\alpha(z)$ and $\beta(z)$, instead of the all-pass function, was recently reported in Ref. 14. The bumping problem in the transition band is considerably suppressed. However, due to the use of the general rational functions, the advantages of the all-pass functions such as low implementation complexity and low coefficient sensitivity are lost. Moreover, the design procedure is considerably more complicated than that in Ref. 12 and the stability of the IIR filters cannot be guaranteed, especially for filters with high order. Another method based on the transformation of a FIR prototype filter was proposed in Ref. 15. Due to the use of the transformation, this method will have considerable restriction on the selection of the analysis/synthesis filters and the system delay.

In this paper, we show that when $\beta(z)$ is chosen as an all-pass function, it is still possible to suppress the bumping problem when $\alpha(z)$ is chosen as a linear-phase FIR function. This has also been observed previously in the paper by Kim and Ansari,¹⁶ but no detailed design procedure is

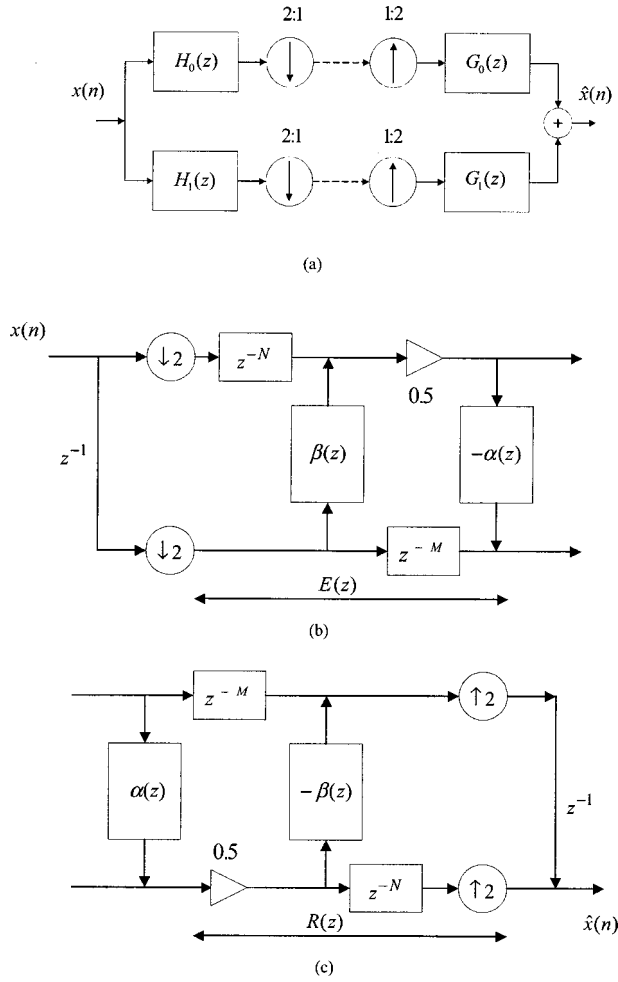


Fig. 1 Two-channel multirate filter bank: (a) maximally decimated filter bank and two-channel PR filter bank proposed in Ref. 12, (b) analysis filter bank, and (c) synthesis filter bank.

given. There are several advantages of this structure. First of all, the advantages of using the all-pass function, such as high stopband attenuation, low implementation complexity, and low coefficient sensitivity, are preserved. Second, the use of a linear-phase FIR function not only helps to reduce the bumping problem, but also provides considerable freedom in choosing the stopband attenuation and the delay of the highpass filter. Moreover, due to the simple structure of the FIR function, the design problem of the high-pass filter can be formulated as a polynomial approximation problem. In particular, it is shown that the least-squares and minimax designs of the high-pass filter can be formulated as the familiar least-squares and the Chebyshev approximation problems, respectively. Both of them can be solved effectively using existing techniques. As a result, causal stable IIR filter banks with approximately linear-phase frequency response and flexible stopband attenuation can be designed easily using the proposed method. Compared with the methods proposed in Refs. 12 and 14, the stopband attenuation of both the low-pass and high-pass filters can be easily controlled without any stability problem. Also, due to the use of the Remez exchange algorithm, the design complexity is extremely low. The generalization of the present approach to the design of a class of IIR wavelet bases is

also studied. By factoring the type II linear-phase FIR function, it is possible to impose the regularity condition on the analysis filters to construct the IIR wavelet bases. Several design examples are given to demonstrate the usefulness of the proposed method. Comparison with other conventional approaches is also given.

The outline of this paper is organized as follows. In Sec. 2, a brief summary of the two-channel structural PR filter bank proposed in Ref. 12 is given. The proposed method and several design examples are given in Sec. 3. Section 4 is devoted to the design of dyadic wavelet bases derived from these two-channel IIR filter banks. Finally, the conclusions are drawn in Sec. 5.

2 Two-Channel Structural PR Filter Banks

Figure 1(a) shows the structure of a two-channel maximally decimated multirate filter bank. It can be shown that^{1,2} the reconstructed signal $Y(z)$ is given by

$$Y(z) = T(z)X(z) + A(z)X(-z), \quad (1)$$

where

$$T(z) = \frac{1}{2} [H_0(z)G_0(z) + H_1(z)G_1(z)], \quad (2)$$

and

$$A(z) = \frac{1}{2} [H_0(-z)G_0(z) + H_1(-z)G_1(-z)]. \quad (3)$$

The aliasing term $A(z)$ can be canceled if the analysis and synthesis filters are chosen as follows:

$$G_0(z) = -H_1(-z), \quad G_1(z) = H_0(-z). \quad (4)$$

The PR filter bank will be perfect reconstruction if the transfer function $T(z)$ is equal to some delay

$$T(z) = cz^{-n_0}. \quad (5)$$

Combining Eqs. (2) and (3) one gets the following PR condition in $H_0(z)$ and $H_1(z)$:

$$T(z) = \frac{1}{2} [H_0(-z)H_1(z) - H_0(z)H_1(-z)] = cz^{-n_0}, \quad (6)$$

where n_0 is an integer and c is a nonzero constant. For IIR filter banks, $H_0(z)$ and $H_1(z)$ are rational functions. In Ref. 12, a class of structurally PR two-channel FIR and IIR filter banks, as shown in Figs. 1(b) and 1(c), was proposed. The polyphase matrix of the filter bank is factored as

$$\begin{aligned} E(z) &= \begin{pmatrix} 0.5 & 0 \\ -0.5\alpha(z) & 1 \end{pmatrix} \begin{pmatrix} z^{-N} & \beta(z) \\ 0 & z^{-M} \end{pmatrix} \\ &= \begin{pmatrix} 0.5z^{-N} & 0.5\beta(z) \\ -0.5z^{-N}\alpha(z) & -0.5\alpha(z)\beta(z) + z^{-M} \end{pmatrix}. \end{aligned} \quad (7)$$

The corresponding expressions for the analysis filters are

$$H_0(z) = \frac{z^{-2N} + z^{-1}\beta(z^2)}{2}, \quad (8)$$

$$H_1(z) = -\alpha(z^2)H_0(z) + z^{-2M-1}. \quad (9)$$

An advantage of such structural system is that it is PR even if the coefficients of $\alpha(z)$ and $\beta(z)$ are quantized. In Ref. 12, $\alpha(z)$ and $\beta(z)$ are chosen to be identical and the delay M is chosen to be $2N-1$. It can be seen from Eqs. (8) and (9) that $H_0(z)$ can be made an ideal low-pass filter if $\beta(z)$ has the following magnitude and phase responses

$$|\beta(e^{j2\omega})| = 1 \quad \forall \omega, \quad (10a)$$

$$\angle \beta(e^{j2\omega}) = \begin{cases} (-2N+1)\omega & \text{for } \omega \in [0, \pi/2] \\ (-2N+1)\omega \pm \pi & \text{for } \omega \in (\pi/2, \pi]. \end{cases} \quad (10b)$$

Another advantage of this structure is that $\beta(z)$ can be chosen as a polynomial or a rational function to obtain structurally PR FIR or IIR filter banks, respectively. In the latter case, the PR filter bank will be causal stable if $\beta(z)$ is causal stable. In Ref. 12, $\beta(z)$ is chosen to be a causal stable all-pass function so that Eq. (10a) is met exactly while Eq. (10b) is approximately satisfied. In the FIR case, $\beta(z)$ is chosen to be a type II linear-phase function with magnitude response as close to unity as possible. Both of these techniques yield filter banks with very high stopband attenuation. As mentioned earlier, such IIR filter banks using the all-pass function has a bump of about 4 dB at $\omega = \pi/2$, no matter how $\beta(z)$ is designed. The stopband attenuation of the high-pass filter $H_1(z)$ is also about 10 dB lower than that of the lowpass filter $H_0(z)$. Instead of using general rational functions for $\alpha(z)$ and $\beta(z)$ as in Ref. 14, $\beta(z)$ is chosen as an all-pass function, while $\alpha(z)$ is chosen as a type II linear-phase FIR filter. The motivation is based on the observation that the bumping problem for this type of structural PR filter bank is much less serious in the FIR case. In fact, by applying model reduction⁷ to the function $\alpha(z)$ in an existing FIR filterbank, causal stable IIR PR filter bank with no bumping can readily be designed.¹⁷ However, the implementation complexity of the resulting IIR filter bank is only slightly lower than that of its linear-phase FIR counterpart. The reason is that the IIR filters after model reduction is in general a stable rational function without any structure, unlike the all-pass function. In this paper, we show that when $\beta(z)$ is chosen as an all-pass function, the bumping problem can be significantly reduced when $\alpha(z)$ is chosen as a linear-phase FIR function. The all-pass function produces filter with very high stopband attenuation. In addition, only N multiplications are required to implement a filter of order N . The all-pass function is also well known for its low sensitivities to coefficient quantization. The linear-phase FIR function $\alpha(z)$ not only helps to reduce the bumping problem but also provides more freedom in choosing the stopband attenuation and the delay of $H_1(z)$. The design of $H_1(z)$ is therefore considerably simplified without any stability problem. In fact, it is shown in the following section that the design of $H_1(z)$ using the

L_2 and L_∞ norm can be formulated as the least-squares and the Chebyshev approximation problems, respectively.

3 Proposed Methods

3.1 Problem Formulation

As mentioned earlier, all-pass-based filter design possesses many attractive properties such as better frequency characteristic, low implementation complexity, and low sensitivities to coefficient quantization. By choosing $\beta(z)$ to be an allpass function as follows:

$$\beta(z) = \frac{\sum_{k=0}^N a_{N,N-k} z^{-k}}{\sum_{k=0}^N a_{N,k} z^{-k}}, \quad (11)$$

it is possible to realize $H_0(z)$ with the aforementioned properties. The design of all-pass-based low-pass filter has previously been addressed in Ref. 13, where the problem is formulated as an eigenvalue problem.¹³

Having assumed that $H_0(z)$ is a reasonably good low-pass filter, we now proceed to formulate the problem of designing $H_1(z)$ with $\alpha(z)$ a type II linear-phase function. It can be seen from Eqs. (8) and (9) that the frequency response of $H_1(z)$ depends on both the lowpass filter $H_0(z)$ and the function $\alpha(z)$.

Let ω_{p0} and ω_{s0} be, respectively, the passband and stopband cutoff frequencies of $H_0(z)$. Similarly, let ω_{p1} and ω_{s1} be the passband and stopband cutoff frequencies of $H_1(z)$, respectively. If $\omega_{p0} = \omega_{s1} = \omega_p$ and $\omega_{s0} = \omega_{p1} = \omega_s$, then the ideal frequency response of $H_1(e^{j\omega})$ is

$$H_d(e^{j\omega}) = \begin{cases} 0 & 0 \leq \omega \leq \omega_p \\ \exp[-j\omega(2M+1)] & \omega_s \leq \omega \leq \pi \end{cases}. \quad (12)$$

The error function $E(\omega)$ of the high-pass filter $H_1(z)$ is defined as

$$E(\omega) = \exp[-j\omega(2M+1)] - \alpha(e^{2j\omega})H_0(e^{j\omega}) - H_d(e^{j\omega}). \quad (13)$$

The weighted distortion measure using the L_p norm of $E(\omega)$ in the passband $I_p = [\omega_s, \pi]$ and stopband $I_s = [0, \omega_s]$ is given by

$$D_p(\alpha) = \int_{I_p} \cup I_s W(\omega) |E(\omega)|^p d\omega, \quad (14)$$

where $W(\omega)$ is a positive weighting function. For simplicity, the $1/p$ 'th power of D_p has been dropped. Usually p is chosen to be two or ∞ , which corresponds to the L_2 and the L_∞ norms, respectively. The optimal $\alpha_p(z)$, with respect to D_p , is therefore given by,

$$\alpha_p(z) = \arg \min_{\alpha} D_p(\alpha). \quad (15)$$

Since the minimization in Eq. (15) involves $\alpha(e^{2j\omega})$, which is periodic with period π , it is different from the conventional Chebyshev approximation. In fact, for a given value of $\omega \in [0, \pi/2]$, $\alpha(e^{2j\omega})$ will affect the values of

$H_1(z)$ at ω as well as $\omega + \pi$. Fortunately, due to the special structure of Eq. (9), further simplification is possible. First of all, it is noted that the magnitude of $\alpha(e^{2j\omega})$ should approximately be equal to 1, except possibly around $\omega = \pi/2$, where it may be even smaller, to obtain filter with good frequency characteristic. It then follows from Eq. (9) that the ripple of $H_1(e^{j\omega})$ in the passband is approximately equal to the stopband attenuation of $H_0(e^{j\omega})$. This enables us to minimize only the stopband attenuation of $H_1(\omega)$ using $\alpha(e^{2j\omega})$, instead of minimizing Eq. (9) over both the pass- and stopbands, and relies on the high stopband attenuation of $H_0(e^{j\omega})$ to achieve small passband ripples. The latter is usually satisfied because $H_0(e^{j\omega})$ is derived from the all-pass function, which can be designed independently. By increasing the length of $\alpha(z)$, the stopband attenuation of $H_1(e^{j\omega})$ can be improved in exchange for a greater delay. We now consider the design of $\alpha(z)$ using the following type II linear-phase FIR filter,

$$\alpha(e^{j\omega}) = \exp(-j\omega M_{\text{odd}}/2) \cos(\omega/2) P(\cos \omega), \quad (16a)$$

where

$$P(\cos \omega) = \sum_{k=0}^L a_k (\cos \omega)^k. \quad (16b)$$

From Eqs. (8), (9), and (16), it can be seen that M_{odd} and hence L should be chosen as $2(M-N)+1$ and $(M_{\text{odd}}-1)/2 = M-N$, respectively. Substituting Eq. (16) into Eq. (13), we have

$$\begin{aligned} E(\omega) &= \exp[-j\omega(2M+1)] \\ &\quad - \exp(-j\omega M_{\text{odd}}) \cos(\omega) P[\cos(2\omega)] H_0(e^{j\omega}), \\ \omega &\in [0, \pi/2]. \end{aligned} \quad (17)$$

For the all-pass-based low-pass filter, the passband is approximately linear phase with transfer function given by

$$H_0(e^{j\omega}) = A(e^{j\omega}) \exp(-j2\omega N), \quad (18)$$

where $A(e^{j\omega})$ is a complex function and is approximately equal to one if $\beta(z)$ is of sufficiently high order. Therefore Eq. (17) can be simplified to

$$\begin{aligned} E(\omega) &= \exp[-j\omega(2M+1)] - \exp[-j\omega(M_{\text{odd}}+2N)] \\ &\quad \times A(e^{j\omega}) \cos(\omega) P[\cos(2\omega)] \\ &= \exp[-j\omega(2M+1)] \\ &\quad \times \{1 - \cos(\omega) P[\cos(2\omega)] \cdot A(e^{j\omega})\}, \\ \omega &\in [0, \pi/2]. \end{aligned} \quad (19)$$

As $P[\cos(2\omega)]$ is a real-valued function, the minimum value of

$$|1 - \cos(\omega) P[\cos(2\omega)] \cdot A(e^{j\omega})|$$

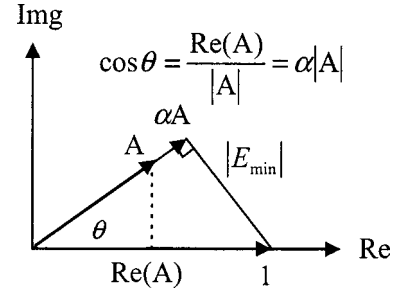


Fig. 2 Geometric interpretation of Eq. (20).

is attained when

$$P_d[\cos(2\omega)] = \frac{\text{Re}[A(e^{j\omega})]}{|A(e^{j\omega})|^2 \cos \omega}, \quad \omega \neq \pi/2. \quad (20)$$

This is best explained geometrically as shown in Fig. 2, where α and A stand, respectively, for $\cos(\omega)P[\cos(2\omega)]$ and $A(e^{j\omega})$. We can see that $|E(\omega)|$ is minimum when αA is the projection of 1 in the direction of A . The corresponding value of α is determined to be $\text{Re}(A)/|A|^2$, which leads to Eq. (20).

Writing $x = \cos(2\omega)$, the design problem becomes a polynomial approximation problem,

$$\begin{aligned} a_{k,\text{opt}} &= \arg \min_{a_k} \int_{I_x} \tilde{W}(x) |P(x) - P_d(x)|^p dx, \\ I_x &= (-1, \tilde{x}_s] \cup [x_s, 1], \quad \tilde{x}_s < x_s = \cos(2\omega_p). \end{aligned} \quad (21)$$

The interval $(0, \tilde{x}_s]$ is an optional disjoint interval to control the values of $P[\cos(2\omega)]$ in the transition band of $H_1(e^{j\omega})$ and $\tilde{W}(x)$ is a positive weighting function. Since $\alpha(e^{j\omega})$ is a type II linear-phase function, it is equal to zero at $\omega = \pi$, i.e., $\alpha(e^{j\pi}) = 0$. The actual value of \tilde{x}_s determines how large the values of $P[\cos(2\omega)]$ are in the interval $(-1, \tilde{x}_s]$ [i.e., in the transition band of $H_1(e^{j\omega})$]. If $p = \infty$, Eq. (21) reduces to the familiar Chebyshev approximation, which can be solved using the Remez exchange algorithm with ideal frequency response given by Eq. (20) and weighting function $\tilde{W}(x)$, given by

$$\tilde{W}_\infty(x) = \cos[0.5 \cdot \arccos(x)] |A\{\exp[j0.5 \arccos(x)]\}|^2. \quad (22)$$

According to the alternation theorem, the optimum solution in Eq. (21) must have a minimum of $L+2$ alternations in I_x . Normally, the values of \tilde{x}_s are very small, therefore, all the alternations appear in the band edges and the interval $[x_s, 1]$.

If $p = 2$, Eq. (21) becomes a least-squares design problem:

$$\begin{aligned} a_{k,\text{opt}} &= \arg \min_{a_k} \int_{I_x} \tilde{W}(x) |P(x) - P_d(x)|^2 dx, \\ I_x &= (-1, \tilde{x}_s] \cup [x_s, 1], \quad \tilde{x}_s < x_s = \cos(2\omega_p). \end{aligned} \quad (23)$$

Let

Table 1 Coefficients of $\alpha(z)$ in Example 1.

n	$\alpha(n)$	n
0	$-6.638650376811762 \times 10^{-003}$	11
1	$1.894646207761688 \times 10^{-002}$	10
2	$-4.256862627194630 \times 10^{-002}$	9
3	$8.811946716409751 \times 10^{-002}$	8
4	$-1.861375907016634 \times 10^{-001}$	7
5	$6.277617720640423 \times 10^{-001}$	6

$$E(\omega) = \int_{I_x} \tilde{W}(x) |P(x) - P_d(x)|^2 dx. \quad (24)$$

Substituting Eq. (16b) into Eq. (24), one gets, after some manipulation, the following

$$E(\omega) = \mathbf{a}^T \mathbf{Q} \mathbf{a} - 2 \mathbf{a}^T \mathbf{r} + c, \quad (25)$$

where

$$\mathbf{a} = [a_0 \ a_1 \ \cdots \ a_L]^T, \quad \mathbf{r} = [r_0 \ r_1 \ \cdots \ r_L]^T,$$

$$[\mathbf{Q}]_{i,j} = \int_{I_x} \tilde{W}(x) x^{i+j} dx, \quad r_i = \int_{I_x} \tilde{W}(x) x^i P_d(x) dx, \quad \text{and}$$

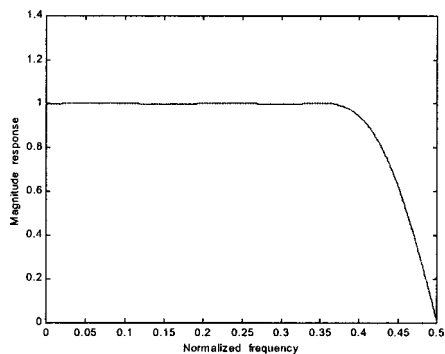
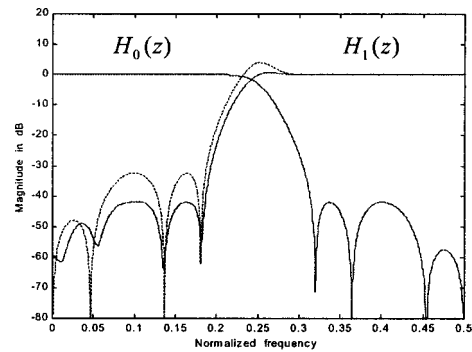
$$c = \int_{I_x} \tilde{W}(x) P_d^2(x) dx.$$

The optimal least-squares solution is therefore given by

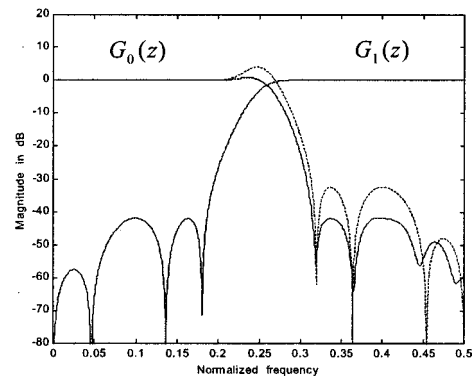
$$\mathbf{a}_{\text{opt}} = \mathbf{Q}^{-1} \mathbf{r}. \quad (26)$$

To avoid calculating the integrals analytically, we can approximate the integral by a summation with sufficient large number of terms. Note that the approach developed here is also valid when $H_0(\omega)$ is a linear-phase function, i.e., it is also applicable to the design of linear-phase FIR PR filter bank using the structure in Figs. 1(b) and 1(c).

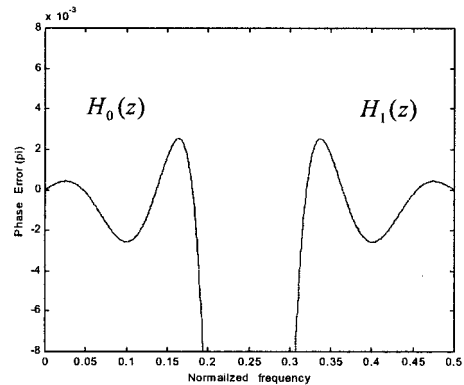
The design procedure can be easily implemented by the signal processing Toolbox of MATLAB. The function REMEZ is used to perform the Remez exchange algorithm ($p = \infty$), while the function FIRLS is used for computing the least-squares solution ($p = 2$).

**Fig. 3** Magnitude response of $\alpha(z)$ in Example 1.

(a)



(b)



(c)

Fig. 4 Frequency responses of Example 1: (a) magnitude responses of $H_0(z)$ and $H_1(z)$; proposed in this paper (solid line) and proposed in Ref. 12 (dashed line); (b) magnitude responses of $G_0(z)$ and $G_1(z)$; proposed in this paper (solid line) and proposed in Ref. 12 (dashed line); and (c) passband phase error of $H_0(z)$ and $H_1(z)$.

3.2 Design Examples

In this section, the proposed design method is evaluated and compared with other conventional methods through several design examples.

3.2.1 Example 1: low-order minimax design

In this design example, the order of the all-pass function $\beta(z)$ in the analysis low-pass filter $H_0(z)$ is chosen as N

=3. To compare the performance of the proposed method with that of Ref. 12, the low-pass filter in example 1 of Ref. 12 is used here. The order of $\beta(z)$ is 3 and its coefficients are $a_{3,1}=0.473$, $a_{3,2}=-0.094$, and $a_{3,3}=0.025$. A type II FIR linear-phase function $\alpha(z)$ with $M_{\text{odd}}=11$ is designed using the minimax approach introduced in Sec. 3.1. The order of the polynomial $P(x)$ is $L=5$ and the delay parameter M is equal to 8. The coefficients and the frequency response of $\alpha(z)$ are shown, respectively, in Table 1 and Fig. 3. Figure 4(a) shows the frequency responses of the analysis filters $H_0(z)$ and $H_1(z)$ designed by the proposed method, while Fig. 4(b) shows the frequency responses of the synthesis filters $G_0(z)$ and $G_1(z)$. Compared with the result in Ref. 12, both the passband and stopband cut-off frequencies of $H_1(z)$ are the same, which are $\omega_{p_1}=0.63\pi$ and $\omega_{s_1}=0.37\pi$, respectively. The stopband attenuation of $H_1(z)$ designed by the proposed method is improved to

about 40 dB, which is about the same as that of $H_0(z)$. The bumping of $H_1(z)$ in the transition band is also significantly suppressed. The deviation from linear phase or the phase error in the passband is plotted in Fig. 4(c). The system delay of the proposed filter bank is slightly increased from 17 to 23 samples. The weighting function $\tilde{W}(x)$ is set to 1 so that the stopband of $H_1(z)$ is not equiripple. It shows that the weighting function in Eq. (22) is useful in achieving an equiripple response in the stopband, as we shall see later in examples 2 and 4. This demonstrates that the bumping problem is effectively suppressed by the proposed method, where a combination of all-pass and linear-phase FIR functions is employed. Also, the stopband attenuation of the high-pass filter can be controlled by using linear-phase function with different orders, in exchange for higher system delay.

The MATLAB source code of example 3.1 is given as follows.

```

N=3; % order of all-pass function  $\beta(z)$ 
% The allpass coefficients are given as:  $a_{3,1}=0.473$ ,  $a_{3,2}=-0.094$ ,  $a_{3,3}=0.025$ .12
% compute  $\beta(e^{j\omega})$  from equation (11)
% compute the lowpass filter  $H_0(e^{j\omega})$  from (8)
Wp=0.371; %  $\omega_p$  passband cutoff frequency of  $H_0(e^{j\omega})$  (normalized by  $\pi$ )
Lwp=round(512*Wp*2); % transform  $2\omega_p$  to integer length (normalized by 512)
La=6; % half length of  $\alpha(z)$ 
delta_w=pi/1024; % frequency spacing  $\Delta\omega$  in  $[0, \pi]$ 
for i=1:Lwp,
    H_delay(i)=H0(i)/exp(-j*delta_w*(i-1)*2*N); % compute  $A(e^{j\omega})$  from (18)
end
for i=1:Lwp,
    H0_(i)=real(H_delay(i))/((abs(H0(i)))^2); % compute  $P_d[\cos(2\omega)](\cos \omega)$ 
end
for i=1:Lwp/2,
    wi(i)=1; % weight in passband of  $\alpha(e^{j\omega})$ 
end
F=[(1:Lwp)/512, 0.99, 1]; % desired frequency points
A=[H0_(1:Lwp)0, 0]; % desired amplitude response
Wi=[wi 0.05]; % weight in frequency axis
alpha=remez(2*La-1,F,A,Wi); % compute  $\alpha(z)$  by Remez Exchange algorithm

```

Vectors H_delay and $H0_$ are used to store, respectively, the samples of $A(e^{j\omega})$ in Eq. (18) and the desired magnitude response of $P_d[\cos(2\omega)]\cos(\omega)$ from $[0, 2\omega_p]$. The stopband cutoff frequency of the desired magnitude response is selected as 0.99π . α is the solution of $\alpha(z)$, and is listed in Table 1.

3.2.2 Example 2: high order minimax design

In this example, the design of $H_1(z)$ with higher order will be illustrated. The order of $\beta(z)$ is chosen as $N=5$ while that of the type-II FIR function is $M_{\text{odd}}=19$. The values of

L and M are 9 and 14, respectively. Note $H_0(z)$ and $\beta(z)$ are designed by the algorithm proposed in Ref. 13. The weighting function of Eq. (22) is used to achieve an equiripple response. Figure 5 shows the frequency responses of the filters $H_0(z)$ and $H_1(z)$. We can see that both the low-pass and the high-pass filters have a stopband attenuation of about 50 dB. Compared with the result in Ref. 14, both methods have comparable cutoff frequencies, but the proposed method has higher stopband attenuation than that of Ref. 14, and the system delay of the proposed design is 39, which is also lower than 43 reported in Ref. 14. Since $\alpha(z)$ is a linear-phase FIR function, the design complexity of the proposed method is also much lower than using the general IIR function reported in Ref. 14.

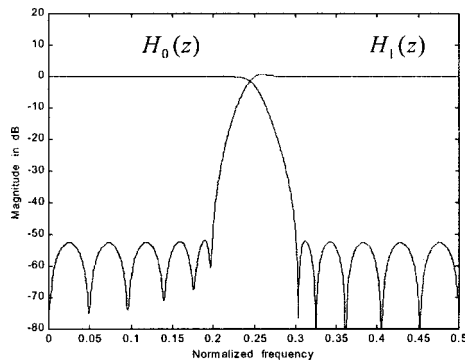


Fig. 5 Magnitude responses of $H_0(z)$ and $H_1(z)$ in Example 2.

3.2.3 Example 3: least-squares design

In this example, the highpass filter $H_1(z)$ is designed using the least-squares approach introduced in Sec. 3.1. The all-pass function $\beta(z)$ and the system delay (i.e., 23) are the same as example 1. The order of the type II function $\alpha(z)$ (see Table 2) is $M_{\text{odd}}=11$, which is identical to that of example 1 for a fair comparison. Figure 6 shows the magnitude responses of the analysis filters, $H_0(z)$ and $H_1(z)$.

3.2.4 Example 4: comparison with maximally flat FIR function

In this example, the proposed method is compared with the one proposed in Ref. 18. In Ref. 18, $H_0(z)$ and $\alpha(z)$ are chosen as the half band and maximally flat FIR filters, respectively. As $\alpha(z)$ is not optimized using any measure, its performance is limited. Without loss of generality, $\beta(z)$ in our system is chosen as an all-pass function, instead of a half band FIR filter, with order $N=8$. In the proposed method, $\alpha(z)$ consists of a type II FIR function with 14 taps and is designed by the Remez exchange algorithm. The magnitude response of the resulting filter $H_1(z)$ is shown as solid line in Fig. 7. The weighting function of Eq. (22) is again used to achieve an equiripple response. The dashed line in Fig. 7 shows the magnitude response of the filter $\hat{H}_1(z)$ when $\alpha(z)$ is chosen as a maximally flat FIR function with the same order. We can see that the filters designed by the proposed method has a much sharper cutoff than that based on the maximally flat FIR function.

The preceding comparison is summarized in Table 5. This includes the results in Examples 1 to 4, and those using allpass-based IIR (Ref. 12), general IIR (Ref. 14) and maximally flat FIR filter bank.¹⁸ Note ω_{p_i} , ω_{s_i} , $i=0, 1$, are,

Table 2 Coefficient of $\alpha(z)$ in Example 3

n	$\alpha(n)$	n
0	$-4.863215050153226 \times 10^{-003}$	11
1	$1.631066387267950 \times 10^{-002}$	10
2	$-3.976458215633622 \times 10^{-002}$	9
3	$8.499478611263678 \times 10^{-002}$	8
4	$-1.841658804678631 \times 10^{-001}$	7
5	$6.267940365478775 \times 10^{-001}$	6

Table 3 Coefficients of allpass filter $\beta(z)$ in Example 5.

n	α_n
0	$1.000000000000000e+000$
1	$4.876862098237123 \times 10^{-001}$
2	$-1.073454651794431 \times 10^{-001}$
3	$4.219586428862606 \times 10^{-002}$
4	$-1.786478722124378 \times 10^{-002}$
5	$8.391063541386605 \times 10^{-003}$

respectively, the cut-off frequencies of analysis filters $H_i(z)$, $i=0, 1$ δ_{s_i} , $i=0, 1$ are their stopband attenuations, and τ is the system delay.

4 Design of IIR Wavelet Bases

4.1 Theory and Design Procedure

The theory of wavelets is closely related to that of multirate filter banks.^{19,20} It has been shown that discrete dyadic wavelets can be obtained from two-channel PR filter banks with added regularity condition. For biorthogonal dyadic wavelet bases, it has been proved that^{21,22} $H_0(z)$ and $G_0(z)$ should have K_0 (or \tilde{K}_0) zeros at $z=-1$ (the K -regularity condition). In addition, $H_1(z)$ and $G_1(z)$ should also have at least one zero at $z=1$. This is equivalent to saying that

$$\left. \frac{d^k H_0(z)}{dz^k} \right|_{z=\pi} = 0, \quad k=0, 1, \dots, K_0; \quad \left. \frac{d^k G_0(z)}{dz^k} \right|_{z=\pi} = 0, \quad k=0, \dots, \tilde{K}_0, \quad (27)$$

$$H_1(1)=G_1(1)=0. \quad (28)$$

Substituting Eqs. (8) and (9) into Eqs. (27) and (28), one gets a set of linear equations that must be satisfied. The problem is a constrained nonlinear optimization problem with linear constraints, which can be solved using the sub-routine NCONF in the IMSL library. If all the freedom is used to maximize the number of zeros at $z=-1$, then Eq.

Table 4 Coefficients of $\hat{\alpha}(n)$ in Example 5, $\hat{\alpha}(n)=-\hat{\alpha}(18-n)$, $n=10, \dots, 18$.

n	$\hat{\alpha}(n)$
0	$-5.736208133518101 \times 10^{-004}$
1	$2.091198148255894 \times 10^{-003}$
2	$-4.164267800479602 \times 10^{-003}$
3	$8.238824717706938 \times 10^{-003}$
4	$-1.424520240298507 \times 10^{-002}$
5	$2.396096715256554 \times 10^{-002}$
6	$-3.912511728316302 \times 10^{-002}$
7	$6.677497653020614 \times 10^{-002}$
8	$-1.315964380619456 \times 10^{-001}$
9	0

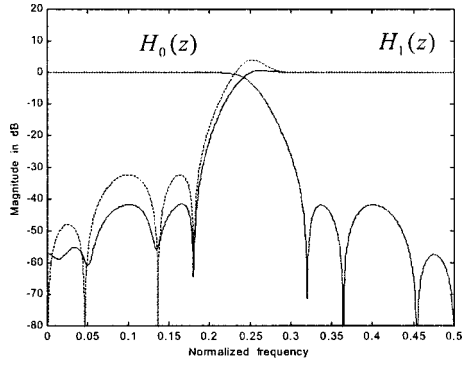


Fig. 6 Magnitude responses of $H_0(z)$ and $H_1(z)$ in example 3: proposed in this paper (solid line); proposed in Ref. 12 (dashed line).

(27) becomes a system of linear equations, which can be used to solve for the filter coefficients. Alternatively, if $p = 2$ in Eq. (21), the objective function becomes quadratic and the problem is recognized as a quadratic programming problem with equality constraints. Again, this can be solved numerically with relatively ease. In this section, we limit ourselves to a class of wavelet bases with $K_0 = 1$ using the methods introduced in Sec. 3. The advantage is that it is an analytic solution and is very easy to apply. More precisely, we only impose one zero at $z = -1$ for $H_0(z)$ and $G_0(z)$.

First, it is observed that $H_0(z)$ always has a zero at $z = -1$ if $\beta(z)$ is an all-pass function. Therefore, $H_0(z)$ satisfies the regularity condition with $K_0 = 1$. For the high-pass filter $H_1(z)$, a zero at $z = 1$ implies $\alpha(z)|_{z=1} = 1$. Since this constraint cannot be incorporated directly into the Parks-McClellan algorithm, it is imposed into $\alpha(z)$ by the following factorization:

$$\alpha(z) = (1 - z^{-1})\hat{\alpha}(z) + \frac{1}{2}(1 + z^{-1})z^{-L_d}, \quad (29)$$

where $\hat{\alpha}(z)$ is a type III linear-phase filter with order $M_{\text{even}} = M_{\text{odd}} - 1$, $L_d = M_{\text{even}}/2$. Equation (29) is obtained by observing that $\alpha(z)$ is symmetric so that its coefficients except the two around the center of symmetry can be written as a product of a type III linear-phase function and $(1 - z^{-1})$. The remaining two coefficients are multiple of $(1 - z^{-1})$.

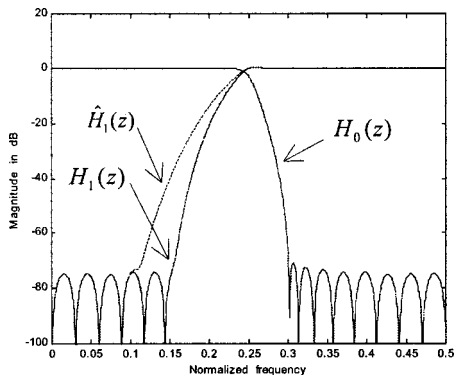


Fig. 7 Magnitude responses of example 4: $H_0(z)$ (solid line), optimal $H_1(z)$ (solid line); $\hat{H}_1(z)$ based on maximally flat $\alpha(z)$ proposed in Ref. 18 (dashed line).

Table 5 Frequency performance comparison with other methods.

	ω_{p_0}	ω_{s_0}	ω_{p_1}	ω_{s_1}	δ_{s_0}	δ_{s_1}	τ
Example 1	0.37π	0.63π	0.63π	0.37π	42	42	23
Example 2	0.4π	0.6π	0.6π	0.4π	52	52	39
Example 3	0.37π	0.63π	0.63π	0.37π	42	42	23
Example 4	0.4π	0.6π	0.6π	0.3π	71	75	45
Example	0.4π	0.6π	0.6π	0.4π	52	50	39
Ref. 12	0.3π	0.63π	0.63π	0.37π	42	32	17
Ref. 14	0.4π	0.6π	0.6π	0.4π	46	49	43
Ref. 18	0.4π	0.6π	0.6π	0.2π	71	71	45

Notes

ω_{p_0} : passband cutoff frequency of $H_0(z)$

ω_{s_0} : stopband cutoff frequency of $H_0(z)$

ω_{p_1} : passband cutoff frequency of $H_1(z)$

ω_{s_1} : stopband cutoff frequency of $H_1(z)$

δ_{s_0} : stopband attenuation of $H_0(z)$ (dB)

δ_{s_1} : stopband attenuation of $H_1(z)$ (dB)

τ : system delay (samples)

$+z^{-1})z^{-L_d}$. Using a scale factor of $1/2$, Eq. (29) guarantees that the required condition $\alpha(z)|_{z=1} = 1$ is satisfied. Since $\hat{\alpha}(z)$ is a type III linear-phase function, it can be written as

$$\hat{\alpha}(e^{j\omega}) = j \exp(-j\omega M_{\text{even}}/2) \sin \omega \tilde{P}(\cos \omega), \quad (30)$$

where

$$\tilde{P}(\cos \omega) = \sum_{k=0}^L a_k (\cos \omega)^k, \quad L = (M_{\text{even}} - 2)/2.$$

Substituting Eq. (30) into Eq. (29), one obtains

$$\begin{aligned} \alpha(e^{j\omega}) &= 2 \sin(\omega/2) \exp[j(\pi - \omega)/2] \\ &\quad \cdot j \exp(-j\omega M_{\text{even}}/2) \sin(\omega) \tilde{P}(\cos \omega) \\ &\quad + \cos(\omega/2) \exp(-j\omega/2) \exp(-j\omega M_{\text{even}}/2) \\ &= \exp(-j\omega M_{\text{odd}}/2) [-2 \sin(\omega/2) \sin(\omega) \tilde{P}(\cos \omega) \\ &\quad + \cos(\omega/2)]. \end{aligned} \quad (31)$$

Using again the expression of $H_0(e^{j\omega})$ in Eq. (18),

$$H_0(e^{j\omega}) = A(e^{j\omega}) \exp(-j2\omega N),$$

the error function $E(\omega)$ can be written as

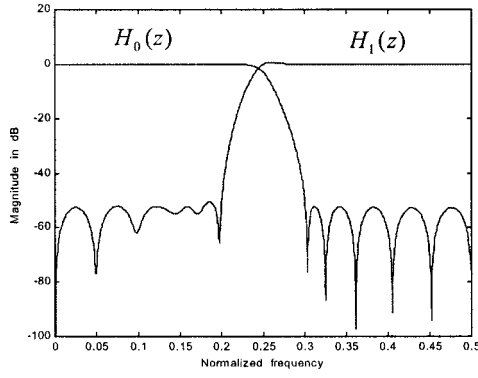


Fig. 8 Magnitude responses of $H_0(z)$ and $H_1(z)$ in Example 5.

$$\begin{aligned}
 E(\omega) &= \exp[-j\omega(2M+1)] \\
 &\quad - \alpha(e^{j2\omega})A(e^{j\omega})\exp(-j2\omega N) \\
 &= \exp[-j\omega(2M+1)] \\
 &\quad - \exp[-j\omega(M_{\text{odd}}+2N)]A(e^{j\omega}) \\
 &\quad \times \{\cos \omega - 2 \sin \omega \sin(2\omega) \tilde{P}[\cos(2\omega)]\} \\
 &= \exp[-j\omega(2M+1)] \{ [1 - A(e^{j\omega}) \cos \omega] \\
 &\quad + 2 \sin \omega \sin(2\omega) \tilde{P}[\cos(2\omega)] A(e^{j\omega}) \} \\
 &= \exp[-j\omega(2M+1)] \{ B(e^{j\omega}) \\
 &\quad + \tilde{P}[\cos(2\omega)] C(e^{j\omega}) \}, \quad 0 < \omega < \pi/2, \quad (32)
 \end{aligned}$$

where

$$\begin{aligned}
 B(e^{j\omega}) &= 1 - A(e^{j\omega}) \cos \omega \quad \text{and} \quad C(e^{j\omega}) \\
 &= 2 \sin \omega \sin(2\omega) A(e^{j\omega}). \quad (33)
 \end{aligned}$$

As $\tilde{P}[\cos(2\omega)]$ is a real-valued function, the minimum value of

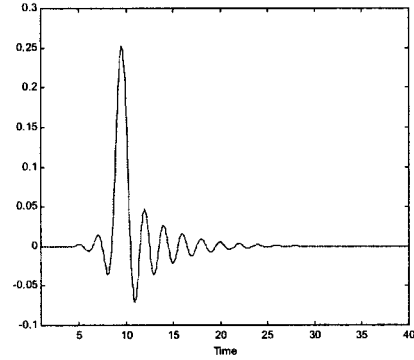
$$|E(\omega)| = |B(e^{j\omega}) + 2 \sin \omega \sin(2\omega) \tilde{P}[\cos(2\omega)] A(e^{j\omega})|$$

is attained when

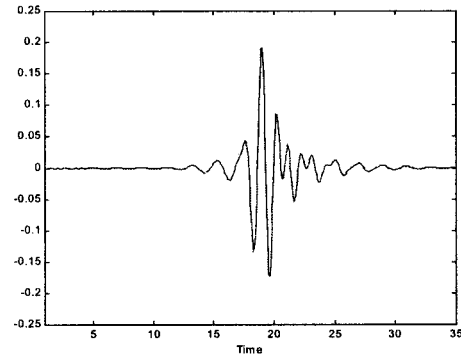
$$\begin{aligned}
 \tilde{P}[\cos(2\omega)] &= P_d[\cos(2\omega)] \\
 &= -\frac{\text{Re}[B^*(e^{j\omega})A(e^{j\omega})]}{2 \sin \omega \sin(2\omega) |A(e^{j\omega})|^2}, \\
 0 < \omega < \pi/2. \quad (34)
 \end{aligned}$$

The solution of $\alpha(z)$, $\hat{\alpha}_{\text{opt}}(z)$, is obtained by approximating $P_d[\cos(2\omega)]$ in Eq. (34) using different norm p . When $p = \infty$, the solution can readily be obtained by the Remez exchange algorithm, where efficient implementation in MATLAB is available. Writing $x = \cos(2\omega)$, the weighting function of Eq. (34) is

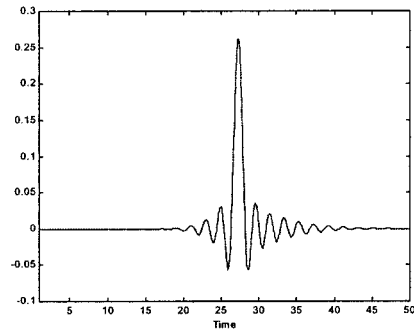
$$\begin{aligned}
 \tilde{W}_{\infty}(x) &= 2 \sin[0.5 \arccos(x)] \\
 &\quad \times \sin[\arccos(x)] |A\{\exp[j0.5 \arccos(x)]\}|^2. \quad (35)
 \end{aligned}$$



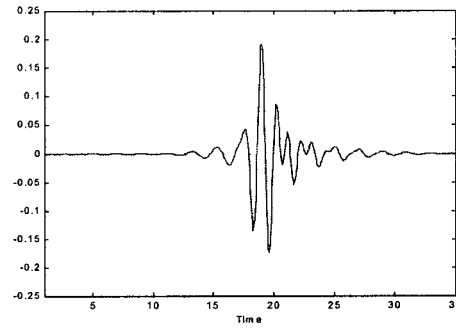
(a)



(b)



(c)



(d)

Fig. 9 Scaling and wavelet functions of Example 5: (a) analysis scaling function; (b) analysis wavelet function; (c) synthesis scaling function; (d) synthesis wavelet function.

4.2 Example 5

In this design example, an IIR 2-channel wavelet basis is designed using the method proposed in Sec. 4.1. The orders

of the allpass function $\beta(z)$ and the type III FIR linear-phase function $\hat{a}(z)$ are $N=5$ and $M_{\text{even}}=18$, respectively. The order of the polynomial $\tilde{P}(z)$, L , and M in Eq. (9) are, respectively, 8 and 14. The weighting function of Eq. (35) is used to achieve an equiripple response of $H_1(z)$. Figure 8 shows the frequency responses of the analysis low-pass and high-pass filters, $H_0(z)$ and $H_1(z)$. Compared with that in Example 2, there is no obvious degradation in the frequency response of $H_1(z)$ after the incorporation of the zero at $z=1$. The system delay is 39, which is also the same as that in Example 2. Figures 9(a) and 9(b) show, respectively, the analysis scaling and wavelet functions derived from this IIR filter bank. Their dual synthesis scaling and wavelet functions are shown in Fig. 9(c) and 9(d). We can see that they are very smooth and fairly symmetric due to the passband linear phase property. Coefficients of $\beta(z)$ and $\hat{a}(n)$ used in this design example are depicted in Tables 3 and 4, and frequency performance compared with other methods is depicted in Table 5.

5 Conclusion

A very simple algorithm for the design of two-channel causal stable IIR PR filter banks is presented. It is based on the structure previously proposed by Phoong et al.¹² By using a combination of all-pass and linear-phase FIR functions, the bumping problem found in the all-pass-based IIR PR filter bank¹² is significantly suppressed. In addition, flexible stopband attenuation of the high-pass filter can readily be obtained in exchange for a greater delay. The design problem is formulated as a polynomial approximation problem and is solved effectively using the Remez exchange algorithm. The extension of this method to the design of a class of IIR wavelet bases is also studied. Furthermore, as the proposed structure is robust to coefficient quantization and its design is extremely simple, it will be very useful in many applications, such as wavelet-based signal analysis and systems.^{23,24}

References

1. P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Englewood Cliffs, NJ (1992).
2. P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks and applications: a tutorial," *Proc. IEEE* **78**, 56–93 (Jan. 1990).
3. A. N. Akansu and R. A. Haddad, *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets*, Academic Press, Boston, (1992).
4. M. Vetterli, "A theory of multirate filter banks," *IEEE Trans. Acoust., Speech, Signal Process.* **35**, 356–372 (Mar. 1987).
5. S. Basu, C. H. Chiang, and H. N. Choi, "Wavelets and perfect reconstruction subband coding with causal stable IIR filters," *IEEE Trans. Circuits Syst.* **42**(1), 24–38 (Jan. 1995).
6. Z. Sijerčić and G. Agarwal, "Factorizations of IIR biorthogonal filter banks based on generalized Bezout identity," in *Proc. IEEE IS-CAS'96*, pp. 77–80 (1996).
7. M. M. Ekanayake and K. Premaratne, "Two-channel IIR QMF banks with approximately linear phase analysis and synthesis filters," *IEEE Trans. Signal Process.* **43**(10), 2313–2322 (1995).
8. T. A. Ramstad, "IIR filter banks for subband coding of images," in *Proc. IEEE ISCAS'88*, pp. 827–830 (1988).
9. S. K. Mitra, C. D. Creusere, and H. Babic, "A novel implementation of perfect reconstruction QMF banks using IIR filters for infinite length signals," in *Proc. IEEE ISCAS'92*, pp. 2312–2315 (May 1992).
10. C. Herley and M. Vetterli, "Wavelets and recursive filter banks," *IEEE Trans. Signal Process.* **41**, 2536–2556 (Aug. 1993).
11. M. Okuda, T. Fukuoka, M. Ikehara, and S. Takahashi, "Design of causal IIR filter banks satisfying perfect reconstruction," in *Proc. IEEE ISCAS'97*, Vol. 4, pp. 2433–2436 (June 1997).
12. S. M. Phoong, C. W. Kim, P. P. Vaidyanathan, and Rashid Ansari, "A new class of two-channel biorthogonal filter banks and wavelet bases," *IEEE Trans. Signal Process.* **43**(3), 649–664 (1995).
13. T. Q. Nguyen, T. I. Laakso, and R. D. Koilpillai, "Eigenfilter approach for the design of allpass filters approximating a given phase response," *IEEE Trans. Signal Process.* **42**(9), 2257–2263 (1994).
14. X. Zhang and T. Yoshikawa, "Design of two-channel stable IIR perfect reconstruction filter banks," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **E81-A**(8), 1592–1597 (1998).
15. D. B. H. Tay, "Design of causal stable IIR perfect reconstruction filter banks using transformation of variables," in *Proc. IEEE IS-CAS'97*, Vol. 4, pp. 2425–2428 (June 1997).
16. C. W. Kim and R. Ansari, "FIR/IIR exact reconstruction filter bank with applications to subband coding of images," in *Proc. Midwest CAS Symp.*, Vol. 1, pp. 227–230, Monterey, CA (May 1991).
17. J. S. Mao, S. C. Chan, and K. L. Ho, "Designing two-channel causal stable IIR PR filter banks and wavelet bases by model order reduction and constrained optimization," in *Proc. IEEE 42nd Midwest Symposium on Circuits and Systems*, New Mexico (August 1999).
18. H. Kiya, M. Yae, and M. Iwahashi, "A linear phase two-channel filter bank allowing perfect reconstruction," in *Proc. IEEE ISCAS'92*, Vol. 2, pp. 951–954 (May 1992).
19. I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.* **41**, 909–996 (Nov. 1988).
20. M. Vetterli and C. Herley, "Wavelets and filter banks: theory and design," *IEEE Trans. Signal Process.* **40**(9), 2207–2231 (1992).
21. S. C. Chan, Y. Luo, and K. L. Ho, "M-channel compactly supported biorthogonal cosine-modulated wavelets bases," *IEEE Trans. Signal Process.* **46**(4), 1142–1151 (1998).
22. A. Cohen, I. Daubechies, and J. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.* **45**, 485–560 (1992).
23. G. Fernández and T. Huntsberger, "Wavelet-based system for recognition and labeling of polyhedral junctions," *Opt. Eng.* **37**(1), 158–165 (1998).
24. F. Espinal, T. Huntsberger, B. D. Jawerth, and T. Kubota, "Wavelet-based fractal signature analysis for automatic target recognition," *Opt. Eng.* **37**(1), 166–174 (1998).



J. S. Mao received his BS degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 1989, and his MEng degree from the University of Electronic Science and Technology of China, Chengdu, China, in 1992. He is currently pursuing the PhD degree at the University of Hong Kong. From April 1992 to October 1995, he was a lecturer at the University of Electronic Science and Technology of China, Chengdu, China. He served as a research assistant with the University of Hong Kong from November 1995 to April 2000. His main research interests include digital signal processing, multirate filter bank systems, and speech coding.



S. C. Chan received his BSc and PhD degrees in electrical engineering from the University of Hong Kong in 1986 and 1992, respectively. From 1990 to 1994, he was a lecturer in the Department of Electronic Engineering, City University of Hong Kong. He joined the Department of Electrical and Electronic Engineering at the University of Hong Kong in 1994, where he is now an associate professor. His main research interests are in digital signal processing and their applications to communications and data compression. He had held visiting positions in Microsoft Research, Redmond, Washington, and Microsoft Research, Beijing, China, in 1998 and 1999, respectively.



K. L. Ho received the BSc (Eng) and the MPhil degrees in electrical engineering from the University of Hong Kong in 1971 and 1973, respectively, and the PhD degree from the University of London in 1977. In 1984, he joined the Department of Electrical and Electronic Engineering at the University of Hong Kong. His current research interests include signal processing and communications systems. He is a chartered engineer of the Engineering

Council, United Kingdom, and a member of the IEE, HKIE and IEEE.