

A SEMI-DEFINITE PROGRAMMING (SDP) METHOD FOR DESIGNING IIR SHARP CUT-OFF DIGITAL FILTERS USING FREQUENCY-RESPONSE MASKING

H. H. Chen, S. C. Chan and K. L. Ho

Department of Electrical and Electronic Engineering, The University of Hong Kong

ABSTRACT

This paper studies the design of frequency response masking (FRM) filters with infinite duration impulse response (IIR) model and masking sub-filters. They are useful in realizing sharp cutoff digital filters with low passband delays. The designs of the model and masking filters are carried out by means of semidefinite programming (SDP) and model order reduction. Design results show that low complexity FRM filters with low passband delay can be obtained.

1. INTRODUCTION

The implementation complexity and system delay of a linear-phase finite duration impulse response (FIR) digital filter with sharp cutoff are usually very large. This is mainly because the order of the filter N is inversely proportional to the width of the transition band. For example, to achieve a specification of $\omega_p = 0.4\pi$, $\omega_s = 0.42\pi$, and $\delta_p = \delta_s = 0.01$, N should be at least 185 according to the formula in [1]:

$$N \approx \frac{-20 \log_{10}(\sqrt{\delta_p \delta_s}) - 13}{14.6[(\omega_s - \omega_p)/(2\pi)]},$$
 where δ_p and δ_s are respectively the

maximum ripples in passband and stopband of the filter, ω_p and ω_s are respectively the cut-off frequencies of the filter. The delay is $N/2=92.5$ samples. This high implementation complexity and long system delay are undesirable in many applications.

One efficient method to reduce the implementation complexity of sharp cut-off FIR filters is to employ the Frequency-Response Masking (FRM) technique, which makes use of the transition bands of an up-sampled digital filter to realize the sharp transition band required. Jing and Fam [2] considered the design of narrow-band linear-phase lowpass filters composing of one model filter and one masking filter. This approach is suitable only for narrow-band filters. Later, Lim [3] proposed another FRM filter structure, which is not restricted to narrowband filters and supports arbitrary bandwidth. To further reduce the system delay of FRM filters, low-delay (LD) FIR subfilters were proposed to realize the model and masking filters in the FRM structure [4-6]. FRM filters employing an allpass-based (AP) IIR model filter and a linear-phase FIR masking filters were recently proposed in [7]. The model and masking filters are first designed separately and nonlinear optimization is then applied to search for the optimal solution. Allpass-based model filters significantly reduce the arithmetic complexity and to some extent the system delay. In this paper, a new semi-definite programming (SDP) method for designing IIR FRM filters is presented. IIR-based FRM filters are very efficient in low-delay and high stopband attenuation applications, where traditional allpass (AP)-based and LD FIR filters are limited by the delay they can achieve and the high

filter order. The masking and model filters are designed in turn using the SDP method and model reduction technique, which yields high quality filters and allows linear or convex quadratic constraints be imposed.

The rest of this paper is organized as follows: In section 2, the FRM approach is briefly introduced. The details of designing the IIR FRM filters using SDP are described in section 3; the corresponding examples are given in section 4. This paper ends with a conclusion in section 5.

2. FRM FILTERS

The transfer function of a FRM digital filter is given by:

$$H(z) = G(z^M)F_0(z) + G_c(z^M)F_1(z), \quad (2-1)$$

where $G(z)$ is the model filter and $G_c(z)$ is the complementary filter because their transfer functions satisfy $G_c(z) = z^{-\tau_G} - G(z)$, where τ_G is the group delay of $G(z)$. $G(z^M)$ and $G_c(z^M)$ are respectively the periodic filters of $G(z)$ and $G_c(z)$. M is some positive integer that determines the transition bandwidths of the subfilters. $F_0(z)$ and $F_1(z)$ are called the masking filters because they mask the model filters in the stopband of the overall filter. Basically, the masking filters extract several passbands of the model filters to form the passband of the overall filter. They also make use of the transition bands of the model or complementary model filters to realize the narrow transition band of the desired filter. There are two cases of utilizing the transition band of the masking and complementary filters: **Case 1**: the transition band is equal to one of the transition bands of $G(z^M)$, **Case 2**: the transition band of $H(z)$ is associated with one of the transition bands of $G_c(z^M)$. The relationship between the cut-off frequencies of the overall filter and subfilters can be found in [3].

From (2-1), it can be seen that the passband delay of $H(z)$ ($\tau = \tau_G \cdot M + \tau_F$) is mainly due to the passband delays of the model and complementary model filters when M is large. This is because their z variables are upsampled by a factor of M to achieve a very narrow transition band. Although the allpass-based model filter approach considered in [7] helps to reduce somewhat the system delay, the delay and stopband attenuation of an allpass-based (AP) filter are still tied up by the order of the AP filters. That is, for a given stopband attenuation, the order and hence the system delay of the filter are fixed. One possible approach is to employ low-delay FIR filters for the model filters as in [6]. However, although a low-delay FIR filter can achieve a very low system delay, its order will increase quite rapidly when the specifications get tighter. The high filter order not only increases the arithmetic complexity, but also increases pipelining and circuits delays during hardware implementation,

giving rise to a lower operating speed. This motivates us to consider the design of IIR-filter-based FRM filters. In general, both the masking and model filters can be IIR filters, and they should be designed jointly to minimize some distortion measure for $H(z)$. To avoid solving this difficult problem, we shall first design the masking filters, and then the model filters. Since the system delay is less sensitive to the passband delay of the masking filter, the latter is chosen as an AP-based IIR filter to reduce its arithmetic complexity, although a general IIR filter can still be used. In designing the model filter, a low-delay prototype FIR filter is first designed by compensating the distortion introduced by the masking filter. Then, model reduction (MR) is applied to this FIR filter to obtain its IIR counterpart. The performance of this IIR is quite satisfactory and further optimization usually does not give significant improvement. In the following section, the design of the AP-based masking filters and LD FIR prototype model filter will be discussed. Due to the decoupling of the design problem, both can be formulated as semi-definite programming (SDP), which supports least squares and minimax criterion, as well as additional linear or convex quadratic constraints. Due to page limitation, we only focus on the minimax design in this paper.

3. DESIGNING IIR FRM FILTERS WITH SDP

The design of the masking and model filters using semi-definite programming (SDP) will be described in turn in the following.

3.1. Designing AP-based IIR masking filters

The masking filter $F(z)$ is constructed from the sum of two allpass filters $A_0(z)$ and $A_1(z)$ that have the same phase response in the passband but a difference of π in the stopband:

$$F(z) = [A_0(z) + A_1(z)]/2.$$

To reduce arithmetic complexity, $A_0(z)$ is chosen as a pure delay. In other words, $A_1(z)$ should approximate the same delay in the passband, while providing an additional phase shift of π in the stopband. Here, we adopt the SDP approach in [10] for designing this allpass filter, because it can impose linear and convex quadratic constraints, together with pole radius constraints. Denote the allpass filter to be designed and the desired response by $A(z) = z^{-N}(\sum_{n=0}^N a_n z^n)/(\sum_{n=0}^N a_n z^{-n})$ and $\theta_d(\omega)$, respectively. The problem can be formulated as the following linear matrix inequalities (LMIs):

$$\min_{\mathbf{a}} \varepsilon \quad \text{subject to} \quad \mathbf{F}(\mathbf{a}, \varepsilon) \geq 0 \quad (3-1)$$

$$\mathbf{F}(\mathbf{a}, \varepsilon) = \text{diag}\{F_1(\mathbf{a}) + \varepsilon \cdot \mathbf{I}_2, F_2(\mathbf{a}) + \varepsilon \cdot \mathbf{I}_2, \dots, F_K(\mathbf{a}) + \varepsilon \cdot \mathbf{I}_2\},$$

$$F_k(\mathbf{a}) = \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix}, \quad k = 1, 2, \dots, K,$$

$$v_1 = \mathbf{a}^T \cdot (\delta \cdot \mathbf{c}(\omega_k) + W(\omega_k) \cdot \mathbf{s}(\omega_k)),$$

$$v_2 = \mathbf{a}^T \cdot (\delta \cdot \mathbf{c}(\omega_k) - W(\omega_k) \cdot \mathbf{s}(\omega_k)),$$

$$\mathbf{c}(\omega_k) = [1, \cos(\omega_k), \dots, \cos(N\omega_k)]^T,$$

$\mathbf{s}(\omega_k) = [0, \sin(\omega_k), \dots, \sin(N\omega_k)]^T$, $\mathbf{a}^T = [a_0, \dots, a_N]$, \mathbf{I}_2 is a (2×2) identity matrix, $W(\omega_k)$ is the weighting function at ω_k and δ is the tangent of the difference between the designed phase and the desired responses. ω_k , $k = 1, 2, \dots, K$, is a dense set of points in the passband and stopband, where the objective

function will be minimized. ε is called the feasibility variable which tells us that whether it is possible to satisfy the given ripple δ . Using a bi-section search, we can successively increase or decrease the ripple δ to obtain the final solution. Additional linear and convex quadratic constraints in the form

$$\mathbf{F}^{(l)}(\mathbf{a}) \geq \mathbf{0}, \quad l = 1, 2, \dots, N,$$

can be stacked along the diagonal entries of $\mathbf{F}(\mathbf{a}, \varepsilon)$ to form $\mathbf{F}'(\mathbf{a}, \varepsilon) \geq \mathbf{0}$, where $\mathbf{F}'(\mathbf{a}, \varepsilon) = \text{diag}\{\mathbf{F}(\mathbf{a}, \varepsilon), \mathbf{F}^{(1)}(\mathbf{x}), \dots, \mathbf{F}^{(N)}(\mathbf{x})\}$. (3-1) with $\mathbf{F}'(\mathbf{a}, \varepsilon) \geq \mathbf{0}$ is also a LMI, which can be solved for the constrained solution. Interested readers are referred to [12] for more details on SDP.

3.2. Design of the FIR/IIR model filter using SDP and MR

The transfer function of the periodic complementary model filter is $G_c(z^M) = z^{-M\tau_G} - G(z^M)$, where $G(z) = \sum_{n=0}^{N_G} g(n)z^{-n}$, and $g(n)$ is the impulse response. Substitute $z = e^{j\omega}$ into (2-1), one gets the frequency response:

$$H(e^{j\omega}) = G(e^{jM\omega})F_0(e^{j\omega}) + [e^{jM\tau_G\omega} - G(e^{jM\omega})]F_1(e^{j\omega}). \quad (3-2)$$

The FRM filter is designed by approximating the desired frequency response $H_d(\omega)$ over a set of disjoint intervals $\omega \in \Omega \subset [-\pi, \pi]$ using the minimax error criterion:

$$\min \delta \quad \text{subject to} \quad (3-3)$$

$$|H(e^{j\omega}) - H_d(\omega)| < \delta, \quad \omega \in \Omega \subset [-\pi, \pi]$$

Since $H(\omega) - H_d(\omega)$ is a linear function of model filter coefficients $g(n)$, given the masking filter, it can be reformulated as an ordinary low-delay FIR filter design problem using SDP [11] as follows

$$\min_{\mathbf{g}} \delta \quad \text{subject to} \quad (3-4)$$

$$\delta - \{\beta_R^2(\omega) + \beta_I^2(\omega)\} \geq 0, \quad \omega \in \Omega \subset [-\pi, \pi]$$

where

$$\beta_R = \mathbf{g}^T [c(\omega)F_{0R}(\omega) + s(\omega)F_{0I}(\omega) - c(\omega)F_{1R}(\omega) - s(\omega)F_{1I}(\omega)] + T_R(\omega) - H_{d_R}(\omega),$$

$$\beta_I = \mathbf{g}^T [c(\omega)F_{0I}(\omega) - s(\omega)F_{0R}(\omega) - c(\omega)F_{1I}(\omega) + s(\omega)F_{1R}(\omega)] + T_I(\omega) - H_{d_I}(\omega),$$

$$\mathbf{g} = [g(0) \quad g(1) \quad \dots \quad g(N_G)]^T$$

$$\mathbf{c}(\omega) = [1 \quad \cos(M\omega) \quad \dots \quad \cos(MN_G\omega)]^T.$$

$$\mathbf{s}(\omega) = [0 \quad \sin(M\omega) \quad \dots \quad \sin(MN_G\omega)]^T$$

$$F_{0R}(\omega) = \text{real}(F_0(\omega)), \quad F_{0I}(\omega) = \text{imag}(F_0(\omega))$$

$$F_{1R}(\omega) = \text{real}(F_1(\omega)), \quad F_{1I}(\omega) = \text{imag}(F_1(\omega))$$

$$T_R(\omega) = \text{real}(e^{-jM\tau_G\omega} \cdot F_1(\omega)), \quad T_I(\omega) = \text{imag}(e^{-jM\tau_G\omega} \cdot F_1(\omega))$$

$$H_{d_R}(\omega) = \text{real}(H_d(\omega)), \quad H_{d_I}(\omega) = \text{imag}(H_d(\omega))$$

Real(x) and imag(x) stand respectively for the real and image parts of x. The constraint in (3-4) is equivalent to [11]

$$\mathbf{\Lambda}(\mathbf{h}, \omega) = \begin{bmatrix} \delta & \beta_R(\omega) & \beta_I(\omega) \\ \beta_R(\omega) & 1 & 0 \\ \beta_I(\omega) & 0 & 1 \end{bmatrix} \geq 0, \quad (3-5)$$

where $\omega \in \Omega \subset [-\pi, \pi]$. Digitizing the frequency variable ω in (3-5) into a dense set of frequencies $\{\omega_i, 1 \leq i \leq m\}$ in the interested band, and writing it in a more compact form gives

$$\Lambda(\mathbf{h}) \succeq 0 \quad (3-6)$$

where $\Lambda(\mathbf{h}) = \text{diag}\{\Lambda(\mathbf{h}, \omega_1), \dots, \Lambda(\mathbf{h}, \omega_m)\}$

Defining the augmented variable $\mathbf{x} = [\delta \ \mathbf{g}^T]^T$, (3-6) can be formulated into the following standard SDP problem:

$$\begin{aligned} & \text{minimize } \mathbf{c}^T \mathbf{x} \\ & \text{subject to } \Lambda(\mathbf{x}) \succeq 0, \mathbf{c} = [1, 0, \dots, 0]^T \end{aligned} \quad (3-7)$$

Additional linear and convex quadratic constraints can also be incorporated as mentioned previously. From this prototype LD FIR filter, model order reduction techniques such as the ERA algorithm [13] can be applied to obtain the required IIR model filter.

4. EXAMPLES

We now present some design examples to illustrate 1) low-delay and high quality FRM filters can be realized by IIR masking and model filters, and 2) the flexibility (say in imposing linear constraints on the filter) and good performance of the SDP design method.

Example 1: A lowpass filter is designed to satisfy the following specifications: $\omega_p = 0.4\pi$, $\omega_s = 0.42\pi$, $\delta_{\max} = 0.2\text{dB}$, $\delta_{\min} = 50\text{ dB}$, passband group delay = 46. To satisfy the specification, the following parameters of the FRM filter is chosen: $M=6$; orders of the two AP-based masking filters are 9 and 11, respectively; delay and order of the FIR prototype model filter are 6 and 52, respectively. Figure 4-1 shows the design results, where we can see that the maximal ripple in the passband is 0.2dB and the minimum stopband attenuation is 52.5dB. It has 75 coefficients totally. To meet the same specifications, the order of a conventional direct-form FIR filter is 200. If the model filter is realized using an AP-based filter, instead of an IIR filter proposed here, it needs 77 coefficients to meet the magnitude specification. However, its passband delay can only be lowered to 76. As mentioned earlier, both the passband delay and the stopband attenuation increase with the order of the AP-based filter. Therefore, it is structurally impossible for the AP-based approach [7] in some cases to realize FRM filters with simultaneously very low passband delay and high stopband attenuation.

Example 2: As mentioned earlier and in the previous example, we notice that the price for achieving a low passband delay is the high order of the FIR model filter. This prevents intensive pipelining to be used to achieve high-speed operation. To shorten the filter order, the technique of model reduction is employed to convert this high order FIR filter to a much lower order IIR filter with similar frequency characteristics. This usually leads to slight degradation in performances of the resulting filters. By model reducing the LD FIR prototype in example 1, we get an 15-order IIR model filter and the frequency response of the resulting FRM filter is shown in Figure 4-2.

Example 3: This example is used to demonstrate the flexibility of the SDP method in imposing linear constraints such as zeros at specific locations (say $\omega = \pi$), which is desirable in realizing digital interpolator. In particular, example 1 is redesigned with a third order zero at $\omega = \pi$ using the SDP method. Due to page limitation, the detailed derivation is omitted. Basically, this requirement is set up as linear equality constraints, which can either be stacked along the existing LMI, or being employed to

eliminate some of the design variables. From (2-1), we can see that the required zeros will be imposed if the same linear constraints on the two masking filters are imposed. Figure 4-3 shows the FRM filter so obtained. From the pole-zero plot, not shown here due to space limitation, the required zeros constraints are properly imposed.

Using the proposed SDP method, no initial guess or nonlinear optimization as in other approach is required. The design sub-programs in section 3 are all convex, meaning that unique global minimum is guaranteed (although their combination might not). We have also compared the SDP approach considered here with the nonlinear optimization method proposed in [7], and found that they are comparable to each other. For example, we get 53.4dB in the stopband of the overall filter for $M=4$ in example 2 of [7]. The stopband attenuation in [7] is 49.4dB after nonlinear optimization but the phase response is better than our approach. Details are omitted here due to page limits.

There are some differences with Lu's approach [6] in designing this kind of filters. The model filters and masking filters are all FIR filters and are designed all together. But it needs approximation to formulate the designing problem into LMIs. Besides, the initial design is needed in that method.

5. CONCLUSION

A new SDP method for designing IIR FRM filters is presented. IIR-based FRM filters is very efficient in low-delay and high stopband attenuation applications, where traditional AP-based and LD FIR filters are limited by the delay they can achieve and the high filter order. The masking and model filters are designed in turn using the SDP method and model reduction technique, which yields high quality filters and allows linear or convex quadratic constraints be imposed.

6. REFERENCES

- [1] J. F. Kaiser, "Nonrecursive digital filter design using the I_0 -sinc window function," *Proc. IEEE Int. Symp. Circuits Syst.*, 1974, San Francisco, pp. 20-23, 1974.
- [2] Z. Jing and A. T. Fam, "A new structure for narrow transition band, lowpass digital filter design," *IEEE Trans. Acoust., Speech, Signal Process.* ASSP-32, pp. 362-370, 1984.
- [3] Y. C. Lim, "Frequency-Response masking approach for the synthesis of sharp linear phase digital filters," *IEEE Trans. Circuits System.* CAS-33, pp. 357-364, 1986.
- [4] C. K. Chen and J. H. Lee, "Design of Sharp-Cutoff FIR Digital filters with Prescribed Constant Group Delay," *IEEE Trans. Circuits and Systems II: Analog and Digital signal Processing*, vol. 43, no. 1, pp. 1-13, Jan. 1996.
- [5] L. Svensson and H. Johansson, "Frequency-Response Masking FIR filters with short delay," *IEEE ISCAS 2002*, vol. 3, pp. 233-236, 26-29 May 2002.
- [6] W. S. Lu and T. Himamoto, "Optimal design of frequency-response masking filters using semidefinite programming," *IEEE Trans. On Circuits and System I: fundamental theory and applications*, vol. 50, no. 4, pp. 557-568, April, 2003.
- [7] H. Johansson and L. Wanhammar, "High-speed recursive digital filters based on frequency-response masking approach," *IEEE. Trans. Circuits and System II: Analog and Digital signal Processing*, vol. 47, no. 1, pp. 48-61, Jan. 2000.
- [8] T. Saramaki and A. T. Fam, "Subfilter approach for designing efficient FIR filters," *Proc. IEEE Int. Symp. Circuits Syst., Espoo, Finland, 1988*, pp. 2903-2915, 1988.

[9] T. Saramaki, Y. Neuvo, and S. K. Mitra, "Design of computationally efficient interpolated FIR filters," *IEEE Trans. Circuits Syst. CAS-35*, pp. 70-88, 1988.

[10] Carson K. S. Pun and S. C. Chan, "The minimax design of digital all-pass filters with prescribed pole radius constraint using semidefinite programming (SDP)," *IEEE ICASSP '03*, vol. 6, pp. VI_413 – VI_416, April 6-10, 2003.

[11] W.-S Lu, "Design of nonlinear-phase FIR digital filters: A semi-definite programming approach", *ISCAS'99*, vol. III, pp. 263-266,

Orlando, FL, May 1999.

[12] H. Wolkowicz, R. Saigal and L. Vandenberg, "Handbook of Semidefinite Programming: theory, algorithms, and applications", *Kluwer Academic Publishers*, 2000.

[13] J. N. Juang and R. S. Pappa, "An eigensystem realization algorithm for modal parameter identification and model reduction," *Journal of Guidance, Control, and Dynamics*, vol. 8, no. 5, September-October 1985, pp. 620-627.

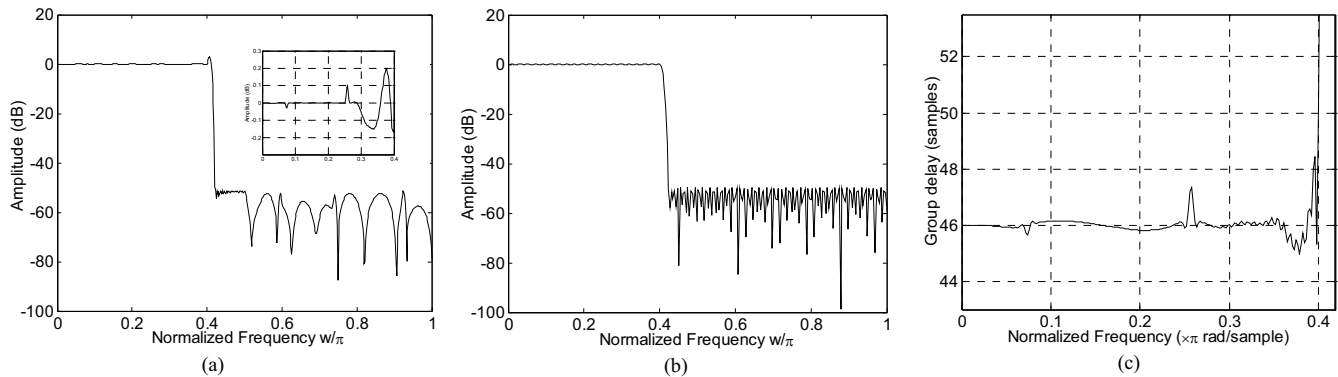


Figure 4-1. Design results of example 1: the magnitude response of (a) the overall filter $H(z)$ (passband details in small figure), (b) the conventional direct-form FIR filter ($N=200$), (c) the group delay of the overall filter.

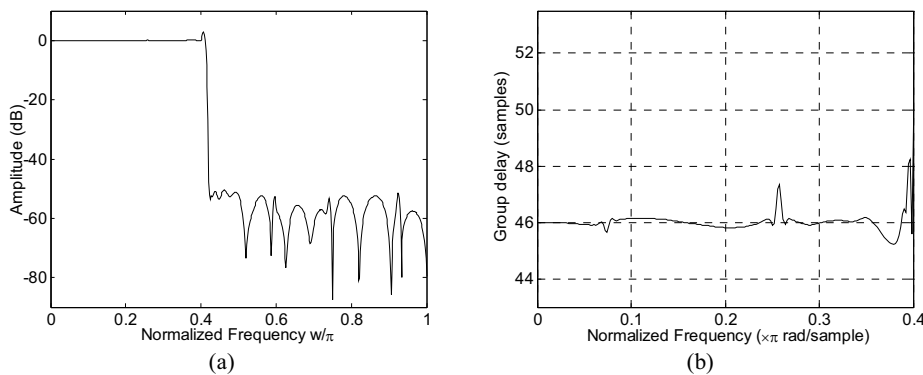


Figure 4-2. Design results of the FRM filter of which the model filters are model reduced form those of example 1: (a) The magnitude response of the overall filter. (b) The group delay of the overall filter.

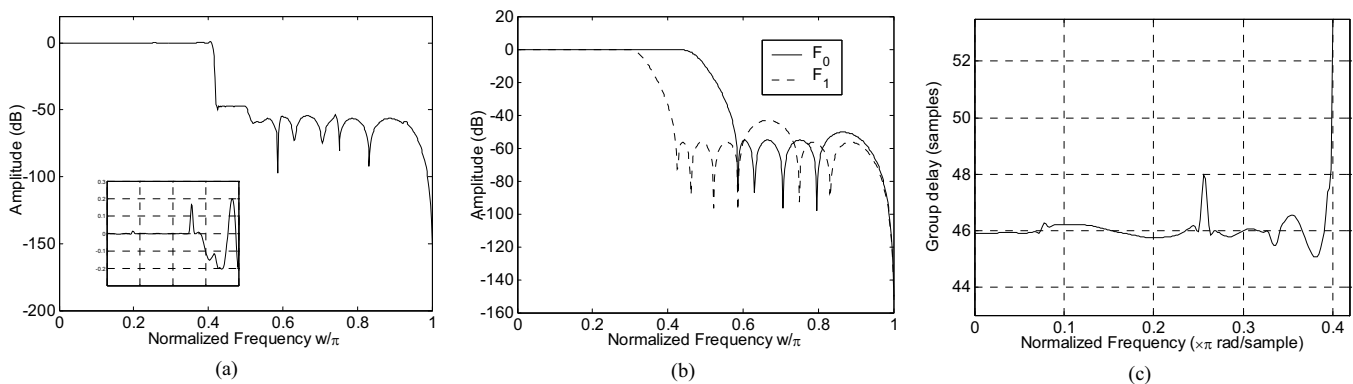


Figure 4-3. Design results of example 3: magnitude response of (a) the overall filter with three zeros at $\omega = \pi$ (passband details in small figure), (b) the masking filters, (c) the group delay of the overall filter.