Optimal Design of Composite Digital Filters Using Convex-Concave Procedure

Wu-Sheng Lu
Dept. of Electrical and Computer Engineering
University of Victoria
Victoria, BC, Canada V8W 2Y2
Email: wslu@ece.uvic.ca

Takao Hinamoto
Hiroshima Institute of Technology
Hiroshima, 731-5193, Japan
Email: hinamoto@ieee.org

***Abstract*—We study a class of composite digital filters, each has a dominating FIR component of order N and an extremely simple IIR component of order r with $r \ll N$ that are connected in parallel. We show that a constrained optimization setting known as convex-concave procedure (CCP) is naturally suited for the design of stable composite filters where the FIR and IIR components are jointly optimized in frequency-weighted minimax sense. A design algorithm based on successive CCP is presented and numerical examples are included to demonstrate that even with $r = 2$ composite filters offer substantial performance improvement relative to their FIR-alone counterparts.**

I. Introduction

Finite-impulse-response (FIR) and infinite-impulse-response (IIR) filters are the two major filter classes when it comes to the design of a digital system to filter discrete data. Many algorithms for effective design of these filters are available and the pros and cons of each filter class are well understood [1]–[4]. The focus of this paper is on a class of digital filters that appears to preserve desirable features from both classes while remedy their shortcomings. Specifically we are interested in a class of *composite filters*, each consists of a major FIR component of order N and a minor IIR component of order r with $r \ll N$ that are connected in parallel. The structure of such composite filters admits fast parallel implementation, and the stability issue for an IIR component of very low order becomes straightforward to deal with. We show that a constrained optimization setting known as convex-concave procedure [5][6], when applied in a successive manner, is naturally suited for the design of stable composite filters where the FIR and IIR components are jointly optimized in frequency-weighted minimax sense. We present technical details of such an algorithm as well as design examples to demonstrate that composite filters, even with a 2nd-order IIR component, can offer substantial performance improvement relative to their FIR-alone counterparts.

II. Composite Digital Filters

A. Structure

We consider a class of digital filters, called composite filters, characterized by $H(z) = H_{\text{FIR}}(z) + H_{\text{IIR}}(z)$ where $H_{\text{FIR}}(z)$ and $H_{\text{IIR}}(z)$ are FIR and IIR transfer functions, respectively. A composite filter has an obvious parallel structure (see Fig.1), we shall call them the FIR and IIR *components* of the filter. For simplicity, in the rest of the paper the FIR and IIR components of a composite filter are denoted by $H_{\text{F}}(z)$ and $H_{\text{I}}(z)$, respectively, thus we have

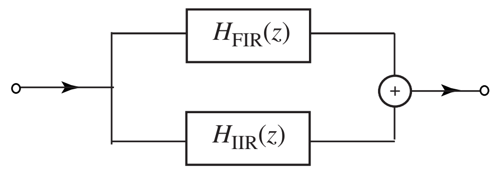

Fig. 1. Structure of a composite filter.

$$H(z) = H_{\text{F}}(z) + H_{\text{I}}(z) \tag{1}$$

where

$$H_{\text{F}}(z) = \sum_{i=0}^{N} c_i z^{-i} \text{ and } H_{\text{I}}(z) = \frac{\hat{b}(z)}{a(z)} = \frac{\sum_{i=0}^{r-1} \hat{b}_i z^{-i}}{1 + \sum_{i=1}^{r} a_i z^{-i}}$$

The order of $H(z)$ is defined as $(N,\ r)$ where N and r are the orders of its FIR and IIR components, respectively.

B. Why composite filters?

Our focus in this paper is on the design and performance evaluation of composite filters with extremely simple IIR components, for example a 2nd-order $H_{\text{I}}(z)$ i.e., $r = 2$. The primary reasons motivating our study are as follows:

(i) An adequately designed composite filter, even with an IIR component of very low order, can offer considerably better performance relative to its FIR counterpart, see Sec. 4 for details.

(ii) With the order of $H_{\text{I}}(z)$ as low as $r = 2$, the stability of a composite filter can be characterized exactly by linear constraints [3], and robust stability can easily be imposed as modified linear constraints [7]. For $H_{\text{I}}(z)$ with $r > 2$, convex constraints for stability are also available [8]–[10].

(iii) As seen in Fig. 1, composite filters naturally admit parallel implementations. Since the order of $H_{\text{F}}(z)$ is much lower than that of an FIR filter with practically the same approximate accuracy and the order of $H_{\text{I}}(z)$ is extremely low (e.g. $r = 2$), composite filters are of reduced complexity and admit simple two-channel fast implementations.

III. Minimax Design of Stable Composite Filters

A. Analysis and problem formulation

Given a desired frequency response $H_d(\omega)$ for $\omega \in [0,\ 2\pi]$, we seek to find a composite filter $H(z)$ of order $(N,\ r)$ such that the largest approximate error $|H(e^{j\omega}) - H_d(\omega)|$ over $[0,\ 2\pi]$ is minimized subject to the stability of $H_{\mathrm{I}}(z)$. For best results we optimize the FIR and IIR components *jointly*. To this end we write

$$H(z) = \frac{b(z)}{a(z)} = \frac{\sum_{i=0}^{n} b_i z^{-i}}{1 + \sum_{i=1}^{r} a_i z^{-i}} \tag{2}$$

where $n = N + r$. The stability of $H(z)$ is assured if and only if the zeros of $a(z)$ are inside the unit circle $\{z : |z| = 1\}$, and in this case $a(z)$ is said to be stable. The design problem at hand can now be formulated as

$$\underset{\boldsymbol{x}}{\text{minimize}} \max_{0 \le \omega \le 2\pi} \left| w(\omega) \left[\frac{b(\omega)}{a(\omega)} - H_d(\omega) \right] \right| \tag{3a}$$

$$\text{subject to:} \quad a(z) \text{ stable} \tag{3b}$$

where $\boldsymbol{x}$ collects the coefficients of $a(z)$ and $b(z)$, $w(\omega) \ge 0$ weights the importance of the approximation at frequency ω, hence is useful to facilitate the designer to emphasize frequency bands of interest while exclude certain do-not-care regions.

Once a solution of problem (3), $(a^*(z),\ b^*(z))$, is obtained, the optimal composite filter can be identified by standard long division of $b^*(z)$ over $a^*(z)$, namely

$$\frac{b^*(z)}{a^*(z)} = H_{\mathrm{F}}^*(z) + H_{\mathrm{I}}^*(z) \tag{4}$$

where $H_{\mathrm{F}}^*(z)$ is a polynomial of order $N = n - r$ from the division and $H_{\mathrm{I}}^*(z) = \hat{b}^*(z)/a^*(z)$ with $\hat{b}^*(z)$ the reminder polynomial of order $r - 1$.

Summarizing, the joint design of the FIR and IIR components of a composite filter can be accomplished by solving the IIR design problem in (3) followed by a polynomial division. We remark that with $n \gg r$ the IIR filter specified in (2) always has more zeros than poles. The benefit this type of IIR filters can bring about was pointed out by several researchers, see e.g. Sec. 2 of [8], although [8] only deals with least-squares designs and the design methodology presented there is entirely different from that of this paper, which employs a successive convex-concave procedure.

B. Convex-concave procedure (CCP)

The CCP [5][6] refers to a heuristic method that is found effective in dealing with a class of *nonconvex* problems of the form

$$\text{minimize} \quad f(\boldsymbol{x}) - g(\boldsymbol{x}) \tag{5a}$$

$$\text{subject to} \quad f_i(\boldsymbol{x}) \le g_i(\boldsymbol{x}) \quad \text{for } i = 1,\ 2,\ \ldots,\ m \tag{5b}$$

where $f(\boldsymbol{x})$, $g(\boldsymbol{x})$, $f_i(\boldsymbol{x})$, and $g_i(\boldsymbol{x})$ for $i = 1,\ 2,\ \ldots,\ m$ are convex. The basic CCP algorithm for (5) is an iterative procedure including two key steps (in the kth iteration):

(i) Convexification of the objective function and constraints by replacing $g(\boldsymbol{x})$ and $g_i(\boldsymbol{x})$, respectively, with their affine approximations

$$\hat{g}(\boldsymbol{x}, \boldsymbol{x}_k) = g(\boldsymbol{x}_k) + \nabla g(\boldsymbol{x}_k)^T (\boldsymbol{x} - \boldsymbol{x}_k) \tag{6a}$$

and

$$\hat{g}_i(\boldsymbol{x}, \boldsymbol{x}_k) = g_i(\boldsymbol{x}_k) + \nabla g_i(\boldsymbol{x}_k)^T (\boldsymbol{x} - \boldsymbol{x}_k) \quad \text{for } i = 1,\ 2,\ \ldots,\ m \tag{6b}$$

(ii) Solving the convex problem

$$\text{minimize} \quad f(\boldsymbol{x}) - \hat{g}(\boldsymbol{x}, \boldsymbol{x}_k) \tag{7a}$$

$$\text{subject to:} \quad f_i(\boldsymbol{x}) - \hat{g}_i(\boldsymbol{x}, \boldsymbol{x}_k) \le 0 \tag{7b}$$

Because of the convexity of all the functions involved in (5), it can be shown that the basic CCP is a descent algorithm [6] and the iterates $\{\boldsymbol{x}_k\}$ produced from (7) converge to the critical point of the original problem (5) [11]. The basic CCP requires a *feasible* initial point $\boldsymbol{x}_0$ to start the procedure. By introducing additional slack variables, a penalty CCP has been developed to accept nonfeasible initial points [6].

C. Fitting (3) to a CCP framework

We begin by converting (3) to an equivalent problem

$$\text{minimize} \quad \delta \tag{8a}$$

$$\text{subject to:} \quad \left| \frac{b_w(\omega)}{a(\omega)} - H_d^{(w)}(\omega) \right|^2 \le \delta, \ \ \omega \in [0, 2\pi] \tag{8b}$$

$$a(z) \text{ stable} \tag{8c}$$

where δ is an upper bound of the squared approximation error over the entire baseband, $b_w(\omega) = w(\omega) b(\omega)$ and $H_d^{(w)}(\omega) = w(\omega) H_d(\omega)$.

Next, we write constraint (8b) as

$$f(\boldsymbol{x}, \omega) \le g(\boldsymbol{x}, \omega) \quad \omega \in [0, 2\pi] \tag{9a}$$

where

$$\boldsymbol{x} = [a_1 \ \cdots a_r \ b_0 \ \cdots \ b_n]^T \tag{9b}$$

$$f(\boldsymbol{x}, \omega) = |b_w(\omega) - H_d^{(w)}(\omega) a(\omega)|^2 \tag{9c}$$

$$g(\boldsymbol{x}, \omega) = \delta |a(\omega)|^2 \tag{9d}$$

Straightforward algebraic manipulations show that, for a fixed frequency ω, $f(\boldsymbol{x}, \omega)$ is a convex quadratic function of $\boldsymbol{x}$ and that $g(\boldsymbol{x}, \omega)$ is a convex quadratic function of $\boldsymbol{x}$ provided that the upper bound δ is also *fixed*. In this case (9a) fits into the type of constraint in (5b). By replacing $g(\boldsymbol{x}, \omega)$ with

$$\hat{g}(\boldsymbol{x}, \boldsymbol{x}_k, \omega) = g(\boldsymbol{x}_k, \omega) + \nabla g(\boldsymbol{x}_k, \omega)^T (\boldsymbol{x} - \boldsymbol{x}_k) \tag{10}$$

constraint (9a) is relaxed to a convex one as

$$f(\boldsymbol{x}, \omega) - \hat{g}(\boldsymbol{x}, \boldsymbol{x}_k, \omega) \le 0 \quad \omega \in [0, 2\pi] \tag{11}$$

Note that the convexity of $g(\boldsymbol{x}, \omega)$ implies $g(\boldsymbol{x}, \omega) \ge \hat{g}(\boldsymbol{x}, \boldsymbol{x}_k, \omega)$, hence an $\boldsymbol{x}$ that satisfies (11) also satisfies (9a), thus a solution of the relaxed convex problem based on CCP always satisfies the constraints of the original problem (8). A problem arising from above analysis is the use of a *fixed* upper bound δ: if δ is set too small, no $\boldsymbol{x}$ satisfying (8b) (hence (9a)) would exist. In order to allow an infeasible initial point to start the CCP, a nonnegative slack variable s is introduced into (11) to further relax it to

$$f(\boldsymbol{x},\omega) - \hat{g}(\boldsymbol{x},\boldsymbol{x}_k,\omega) - s \leq 0 \quad \omega \in [0, 2\pi] \tag{12a}$$

$$s \geq 0 \tag{12b}$$

D. An interplay between s and δ

An important step that makes CCP a well suited design framework for the problem at hand is to establish a connection between the minimum slack variable s and tightest upper bound δ. To this end, we consider the convex problem

$$\begin{aligned} \text{minimize} \quad & s && \text{(13a)} \\ \text{subject to:} \quad & f(\boldsymbol{x},\omega) - \hat{g}(\boldsymbol{x},\boldsymbol{x}_k,\omega) - s \leq 0, \omega \in \Omega && \text{(13b)} \\ & s \geq 0 && \text{(13c)} \\ & a(z) \text{ stable} && \text{(13d)} \end{aligned}$$

where Ω denotes a set of K grid points placed uniformly over the frequency region of interest, and a fixed δ is involved via (13b). If we denote the solution of (13) by s^* and the solution of (8) by δ^*, then it follows from the analysis in Sec. 3.C that $s^* > 0$ if and only if $\delta < \delta^*$; and $s^* = 0$ if and only if $\delta \geq \delta^*$. In words, the solution of (13), i.e. s^*, is indicative of whether or not the value of δ that is set in $\hat{g}(\boldsymbol{x},\boldsymbol{x}_k,\omega)$ (see (10) and (9d)) is too small: this will be the case when s^* is strictly positive. Based on this, the design problem in (8) (hence (3)) can be solved by successively bisectioning a range of uncertainty (ROU) for δ, say $[\delta_L, \delta_U]$, in that the midpoint of the most updated ROU is used as the value of δ in (13) whose solution s^* is then evaluated to reduce the ROU by half: if $s^* > 0$, update the lower bound δ_L to be δ, otherwise update the upper bound δ_U to be δ. The procedure continues until the length of the ROU is less than a prescribed tolerance ε_1.

E. An algorithm for solving (3) by successive CCP

The input data of a design includes filter order (N, r); desired frequency response $H_d(\omega)$; frequency weight $w(\omega)$; initial point $\boldsymbol{x}_0$ that ensures the stability of $a(z)$ but otherwise arbitrary (e.g. any $\boldsymbol{x}_0$ with its first r components being zero as such an $\boldsymbol{x}_0$ simply corresponds to $a(z) \equiv 1$); tolerance ε_1 for ROU convergence and tolerance ε_2 for CPP convergence; initial ROU $[\delta_L, \delta_U]$ that contains δ^*, which can easily be identified as, e.g., $\delta_L = 0$ and δ_U being a sufficiently large scalar; and a $\tau > 0$ to keep the poles not too close to the unit circle $\{z : |z| = 1\}$ (see (14) below). The algorithm can now be outlined as follows.

Step 1: If $\delta_U - \delta_L < \varepsilon_1$, go to Step 5; else go to Step 2.
Step 2: Set $\delta = (\delta_L + \delta_U)/2$ and $k = 0$.
Step 3: Solve (13) iteratively until $\|\boldsymbol{x}_k - \boldsymbol{x}_{k+1}\| < \varepsilon_2$. Denote the solution as $(s^*,\ \boldsymbol{x}^*)$.
Step 4: If $s^* > 0$, set $\delta_L = \delta$, else set $\delta_U = \delta$; Go to Step 1.
Step 5: Set $\delta = (\delta_L + \delta_U)/2,\ k = 0$ and repeat Step 3. Output the solution as $(s^*, \boldsymbol{x}^*)$

We now conclude this section with a remark on stability constraint (13d). Convex stability constraints are available from the literature ([3], [7]–[10]). Especially if $r = 2$, necessary and sufficient stability conditions are characterized by three linear inequalities [3], and robust stability of a seond-order $a(z) = 1 + a_1 z^{-1} + a_2 z^{-2}$ is assured by constraints

$$a_2 \leq 1 - \tau,\ a_1 - a_2 \leq 1 - \tau,\ a_1 + a_2 \geq -1 + \tau \tag{14}$$

where parameter $\tau > 0$ controls the stability robustness in terms of how far the filter's poles are away from the unit circle $\{z : |z| = 1\}$.

IV. Design Examples

The algorithm described in Sec. 3 was applied to design a minimax lowpass and a minimax highpass composite filters. The primary purpose of these designs is to demonstrate that with the help of a second-order stable IIR parallel block, an FIR filter of much reduced length can offer a magnitude resonse that can only be achived by an FIR filter alone of considerably higher order. It will also demonstrate that the passband phase response of a composite filter can be made practically linear with a much reduced group delay relative to that of the FIR-alone counterpart. The algorithm was implemented using CVX [12].

A. Design of a minimax lowpass composite filter of order $(N, r) = (14, 2)$

The desired normalized passband and stopband edges were set to $\omega_p = 0.4\pi$ and $\omega_a = 0.45\pi$ respectively, and the desired passband group delay was set to $D = 8$. The weight $w(\omega)$ was set to 1, in the passband and stopband, and 0 elsewhere. A initial point $\boldsymbol{x}_0 = [\boldsymbol{a}_0; \boldsymbol{b}_0]$ was produced using $\boldsymbol{a}_0 = [0\ \ 0]^T$ and $\boldsymbol{b}_0$ being the impulse response of a lowpass linear phase FIR filter of order $n = N + r = 16$ with cutoff frequency $\omega_c = (\omega_p + \omega_a)/2 = 0.425\pi$ designed using the standard least-squares method. The initial ROU for δ was set to $[\delta_L, \delta_U] = [0.001,\ 0.01]$ and convergence tolerances were set to $\varepsilon_1 = \varepsilon_2 = 10^{-5}$. For stability robustness, parameter τ was set to 0.073.

With the input data given above, it took the algorithm ten iterations to reduce the ROU for δ to be less than 10^{-5}. The optimal δ was found to be $\delta^* = 1.9888 \times 10^{-3}$. The magnitude and passband phase responses of the composite filter obtained are shown in Fig. 2. For comparison, it requires a Parks-McClellan (PM) FIR filter of order 44 to produce a comparable magnitude response. However the group delay of the PM FIR filter is 22 while the passband group delay of the composite filter is only 8. See Table 1 for numerical details of the design and comparison. The coefficients of the optimized $H_{\mathrm{F}}^*(z)$ and $H_{\mathrm{I}}^*(z)$ are given in Table 2.

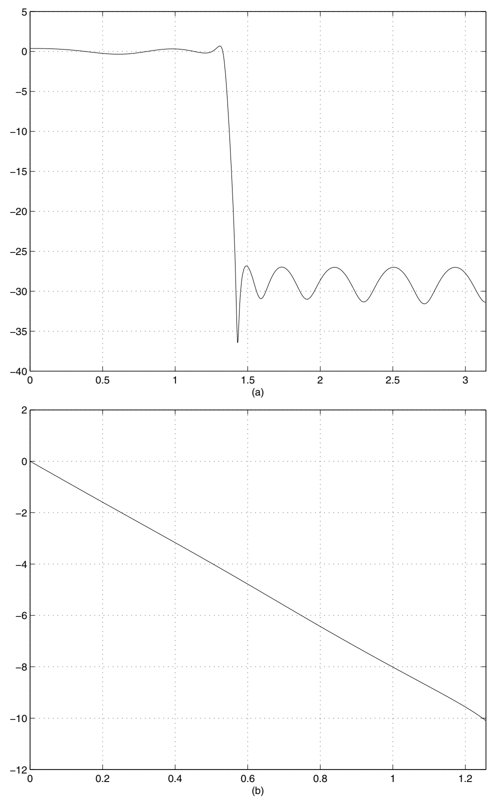

Fig. 2. (a) Magnitude and (b) Passband phase responses of the (14, 2) composite filter in Example A.

B. Design of a minimax highpass composite filter of order (14, 2)

We set $\omega_a = 0.55\pi, \omega_p = 0.6\pi$ and group delay $D = 8$. The weight $w(\omega)$ was set to 1 in the passband and stopband, and 0 elsewhere. An initial point $\boldsymbol{x}_0 = [\boldsymbol{a}_0; \boldsymbol{b}_0]$ was produced with $\boldsymbol{a}_0 = [0 \;\; 0]^T$ and $\boldsymbol{b}_0$ being the impulse response of a highpass linear-phase FIR filter of order $n = N + r = 16$ with cutoff frequency $\omega_c = (\omega_a + \omega_p)/2 = 0.575\pi$. The initial ROU for δ was set to $[\delta_L, \delta_U] = [0.001, \;\; 0.01]$, and convergence tolerances were set to $\varepsilon_1 = \varepsilon_2 = 10^{-5}$. The parameter τ for stability robustness was set to $\tau = 0.073$. The numerical results of the design as compared to a Parks-McClellan FIR filter of order 44 are shown in Table 1.

V. Conclusion

This paper investigates composite filters as an alternative class of digital filters for signal processing as it offers comparable approximation accuracy with much reduced group delay and implementation complexity relative to FIR-alone counterparts. Minimax design of stable composite filters are accomplished in a constrained optimization setting based on successive CCP, and the design examples are presented to support the proposed algorithm.

Table 1 Performance evaluation and comparison.

	Lowpass		Highpass	
	Composite	PM FIR	Composite	PM FIR
order	(14, 2)	44	(14, 2)	44
passband deviation	0.0446	0.0483	0.0446	0.0482
stopband attenuation	26.8348 dB	26.3304 dB	26.8470 dB	26.3409 dB
group delay	8	22	8	22
magnitude of poles	0.9628	—	0.9628	—

Table 2 Coefficients of the composite filter in Example A

$H_{\mathrm{F}}^*(z)$	$H_{\mathrm{I}}^*(z)$
$c_i, i = 0, \ldots, 14$	$a_i, i = 1, 2$
−0.038592705364126	−0.450010233740400
0.015978713746874	0.926999944641125
0.047388048611976	$\hat{b}_i, i = 0, 1$
0.025876214329427	0.003952528292775
−0.045614735818243	−0.049645518158357
−0.063824445494192	
0.067802041742507	
0.295596326680472	
0.426783865790648	
0.327553676393167	
0.081032227904421	
−0.092744852509694	
−0.077385272963324	
0.031527187309290	
0.074139792710036	

References

[1] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Prentice-Hall, 1975.

[2] T. W. Parks and C. S. Burrus, *Digital Filter Design*, Wiley, 1987.

[3] A. Antoniou, *Digital Filters: Analysis, Design, and Applications*, 2nd ed., McGraw-Hill, 1993.

[4] S. K. Mitra, *Digital Signal Processing: A Computer-Based Approach*, 4th ed., McGraw-Hill, 2010.

[5] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.

[6] T. Lipp and S. Boyd, "Variations and extensions of the convex-concave procedure," *Research Report*, Stanford University, Aug. 2014.

[7] W.-S. Lu and T. Hinamoto, "Optimal design of IIR digital filters with robust stability using conic-quadratic-programming updates," *IEEE Trans. Signal Processing*, vol. 51, no. 6, pp. 1581–1592, June 2003.

[8] M. C. Lang, "Least-squares design of IIR filters with prescribed magnitude and phase responses and a pole radius constraint," *IEEE Trans. Signal Processing*, vol. 48, no. 11, pp. 3109–3121, Nov. 2000.

[9] W.-S. Lu, "Design of stable minimax IIR digital filters using semidefinite programming," *Proc. ISCAS 2000*, vol. 1, pp. 355–358, Geneva, May 2000.

[10] B. Dumitrescu and R. Niemistö, "Multistage IIR filter design using convex stability domains defined by positive realness," *IEEE Trans. Signal Processing*, vol. 52, no. 4, pp. 962–974, April 2004.

[11] G. R. Lanckreit and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Advances in Neural Information Processing Systems*, pp. 1759–1767, 2009.

[12] CVX Research, http://cvx.com/cvx, August 2012.