UNIFIED DESIGN ALGORITHM FOR COMPLEX FIR AND IIR FILTERS

Worayot Lertniphonphun and James H. McClellan

Center for Signal and Image Processing
School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, GA 30332–0250, USA.
Email: worayot, mcclella@@ece.gatech.edu

ABSTRACT

In this paper, a general filter design norm is proposed with the intent of producing a unified design algorithm for all types of filters—FIR, IIR and 2-D FIR with complex specifications. The Chebyshev, least squares, and constrained least squares problems become special cases because this norm uses a convex combination of the 2-norm and the Chebyshev norm. The primary benefit of this new problem formulation is that a single efficient multiple exchange algorithm (similar to Remez) has been developed to cover all the different filter types for magnitude and phase approximation. In the new algorithm, a small subproblem is formed at each step and is solved with an iterative reweighted least squares technique which can handle the design of complex filters easily. Finally, the norm definition allows easy trade-offs between the relative importance of error energy and worst-case error.

1. INTRODUCTION

Filters with minimal Chebyshev error are often needed in communication and DSP applications. However, these filters tend to be vulnerable to out-of-band white noise because their stopband error is relatively high. Adams [1] discussed the issue thoroughly and suggested that the best filter should be designed by combining the minimax and least squares criteria as a balance between the two types of error. However, the combination problem was not solved directly; instead, a related problem, constrained least squares [1, 2], was used to do this combined optimization. Both the constrained problem and the Chebyshev problem still have challenging design questions, especially for complex filters, IIR filters, and multi-D filters.

Even though a constrained least-squares problem can be used to design filters with both small RMS error and maximal error, we must know the filter characteristics *a priori* to set up the constraint. On the other hand, this paper proposes a direct design procedure to optimize a combined 2-norm and Chebyshev norm. The combination turns out to be a norm that has some properties similar to the Chebyshev norm and some like the 2-norm. The combined norm forms a strictly convex unit ball which implies the uniqueness of the optimal solution without the Haar condition.

To solve the design problem, this research develops a multiple frequency exchange algorithm that is a generalization of the Remez algorithm [3]. The algorithm is similar to [4] for the complex Chebyshev problem where iterative reweighted least squares was used to solve the subproblem formed in each multiple exchange iteration. The new algorithm can handle the case of complex filter design and also IIR filter design. The fact that the new algorithm solves the design problem using a Remez-like procedure helps make the new algorithm as efficient as existing algorithms based on constrained least-squares. This algorithm is efficient enough to design 2-D filters with 400 free coefficients. For larger filters, the algorithm is limited by the procedure to solve the reweighted least squares subproblems because they require a very large amount of memory and computation on the order of $O(N^2)$.

This paper will formulate the new error norm problem, and provide a general algorithm that directly optimizes this norm and that works for FIR, IIR and 2-D FIR filters.

Prepared through collaborative participation in the Advanced Sensors Consortium sponsored by the U.S. Army Research Laboratory under Cooperative Agreement DAAL01-96-2-0001. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

2. PROBLEM STATEMENT

The design problem requires a finite number of filter coefficients: the feedforward coefficients, $b[n], n = 0, 1, \cdots, N$ and the feedback coefficients, $a[m], m = 0, 1, \cdots, M$. Note that filters are FIR when $M = 0$ and IIR, otherwise. In the causal IIR case, we can set $a[0] = 1$ to get an unique filter.

In this paper, filters are designed to approximate an ideal frequency response, $I(\omega)$, with an actual filter whose frequency response is

$$H(\omega) = \frac{B(\omega)}{A(\omega)}$$

where $X(\omega) = \sum_k x[k]e^{-j\omega k}$ is the discrete-time Fourier transform (DTFT) of $x[n]$.

The approximation is carried out by minimizing the norm of the weighted error, $E(\omega) = W(\omega)\,(I(\omega) - H(\omega))$. To achieve the design goal, this paper proposes a new norm, called the *combined norm,* that is a convex combination of the Chebyshev norm and the 2-norm with a weighting parameter $(0 \le \alpha \le 1)$:

$$\|E\|_\alpha^2 = \alpha\|E\|_\infty^2 + (1-\alpha)\|E\|_2^2 \quad (1)$$

where the Chebyshev norm is computed by

$$\|E\|_\infty = \max_\omega |E(\omega)|$$

and the 2-norm is computed by

$$\|E\|_2 = \sqrt{\frac{\int_\omega |E(\omega)|^2 d\omega}{\int_\omega d\omega}}.$$

In this paper, the 2-norm is normalized in order to have comparable weighting with the Chebyshev norm. The weight function, $W(\omega)$, which can be different for the 2-norm and Chebyshev norm, permits design flexibility for some special filters such as bandpass filters having a Chebyshev passband and a least-squares stopband. Note that the combined norm is formed as a convex combination of the squares of the two norms—the purpose of this is to have an error gradient that is linear.

3. COMBINED NORM

The combined norm introduced in this paper allows simultaneous control of both the maximal error and the RMS error. Depending on the choice of α, the combined norm minimization can exhibit properties similar to either the Chebyshev or least-squares solutions. A useful property for FIR filter design is that minimizing the combined norm is guaranteed to give an unique solution even without the Haar condition.

It is not difficult to show that the combined norm satisfies $\|x+y\|_\alpha \leq \|x\|_\alpha + \|y\|_\alpha$. Along with other easy to prove properties, the convexity property implies that the combined norm is actually a norm. Therefore, when the ideal function is bounded, we can claim that the optimal solution exists and is bounded.

For $\alpha < 1$, the norm becomes a strictly convex norm meaning that if $x \neq y$, $\|x\|_\alpha = \|y\|_\alpha = 1$, and $0 < t < 1$ then $\|tx + (1-t)y\|_\alpha < 1$. The strict convexity property implies that the optimal solution is unique. For $\alpha = 1$, the problem becomes a Chebyshev problem, in which case uniqueness holds due to the Haar condition for the complex exponential kernel (in the FIR case).

3.1. Nature of the Optimal Solution

Similar to the Chebyshev solution, the optimal solution of the combined norm problem has many extremal points where the error reaches its maximum and is equal to the Chebyshev error. However, the number of extremal points need not be greater than the number of design parameter as happens in the Chebyshev problem. This behavior of the optimal solution is essential to the development of the new design algorithm.

3.2. Equivalence of Combined Norm Minimization and Constrained Least Squares

The combined norm minimization and constrained least squares (CLS) are equivalent, even though the two optimization problems are formed differently. To show equivalence, let the filter H_n be the optimal solution for the combined norm minimization. Since H_n has the smallest combined norm, $\|E_n\|_\alpha^2 = \alpha\|E_n\|_\infty^2 + (1-\alpha)\|E_n\|_2^2$ is minimized. Denote the Chebyshev error with $\|E_n\|_\infty = \epsilon_n$, then the error $\|E_n\|_\alpha^2 - \alpha\epsilon_n^2 = (1-\alpha)\|E_n\|_2^2$ is minimized over all functions that have maximal error ϵ_n. Now consider solving the CLS problem

$$\min \|E\|_2 \qquad \text{s.t.} |E| \leq \epsilon_n .$$

The solution, H_c, will have the smallest error $\|E_c\|_2$. Using the uniqueness of the combined norm solution, we conclude that the solutions for the two optimizations are the same.

Note that other minimum norm problems such as

$$\min \left\{ \gamma\|E\|_\infty + (1-\gamma)\|E\|_2 \right\},$$

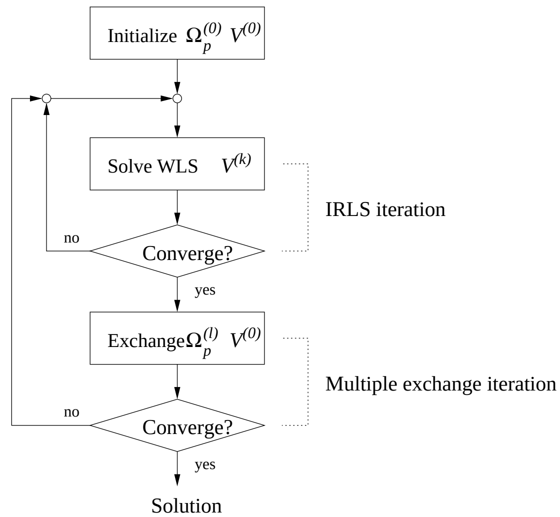


Figure 1: Block diagram for the new design algorithm.

are also equivalent to the combined norm problem. The proof is very similar to the one given above. The different formulations give considerable flexibility when selecting a design method matched to an application. The combined norm minimization tends to be more practical than CLS, because the problem does not require any *a priori* knowledge of the filter to set the constraint.

4. ALGORITHM

Chebyshev optimization can be done with a weighted least-squares algorithm such as Lawson's [5] iterative reweighted least squares (IRLS). Likewise, the combined norm problem can be solved by a similar reweighted least-squares iteration. However, the least-squares norm involves an integral that must be discretized in a numerical algorithm. If the entire domain is discretized into P points, the result is a matrix that is $P \times L$ where L is the number of filter coefficients. Usually P is chosen to be greater than $10L$, so the least-squares algorithm is very inefficient for large L. In order to have an efficient algorithm, we need to keep the matrix small and nearly square, so we will emulate the Remez algorithm which iteratively solves for the error on its extremal set. This approach to updating the weight on the small extremal set was first proposed in [4].

The block diagram for the new algorithm is shown in Fig. 1. It is similar to the Remez exchange, where there is an outer loop with an exchange procedure that finds the extremal subset and an inner loop with an IRLS procedure that computes the optimal filter coefficients on the restricted subset of extremal frequencies.

In order to guarantee convergence, the exchange rule for the extremal set must force the maximal error on the extremal set to be increasing at every step. The exchange procedure can be as simple as finding the set of local error maxima (as in [3, 4]). However, the convergence rate depends directly on the number of elements in the extremal set. Therefore, additional exchange rules (not discussed here) can be added to the procedure to accelerate the algorithm.

The more difficult procedure is to compute the filter coefficients. This is done by using the IRLS algorithm, because IRLS is not only efficient for a small grid set, but is also robust to the removal of any points that do not belong to the extremal set. The procedure for this subproblem starts by using the property that the Chebyshev problem is equivalent to a weighted least squares problem which further implies

$$\begin{aligned} \min \|E\|^2_{\alpha,\Omega_p} &= \min \alpha \|E\|^2_{\infty,\Omega_p} + (1-\alpha)\|E\|_2^2 \\ &= \min \alpha \|E\|^2_{V,\Omega_p} + (1-\alpha)\|E\|_2^2 \end{aligned} \tag{2}$$

where $\|E\|^2_{V,\Omega_p} = \sum_{\omega\in\Omega_p} |V^2(\omega)E^2(\omega)|$, Ω_p is the extremal set and V is the optimal weight. The optimal weight is computed by

$$V^{(k+1)} = \sqrt{\frac{V^{(k)^2}E^{(k)}}{\sum_{\Omega_p} V^{(k)^2}E^{(k)}}}$$

as for the Chebyshev problem.

The IRLS problem (2) is a weighted least squares problem that can be solved quite easily by solving for a zero of the gradient with respect to the design parameter. However, the term $\|E\|_2$ is still a norm on the continuous domain, so it must be solved on a fine grid. This eventually makes the algorithm inefficient. However, the term $\|E\|_2$ can be minimized efficiently by using the following variation. Consider

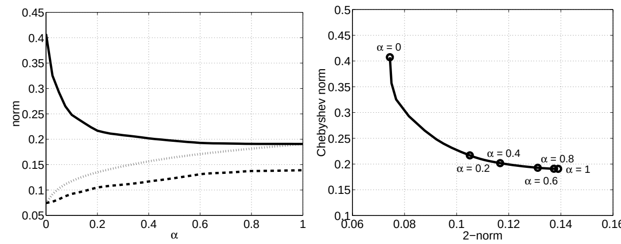
$$\begin{aligned} \|E\|_2^2 &= \|W(I-H)\|_2^2 \\ &= \|\mathbf{y} - \mathbf{X}\mathbf{h}\|^2 \\ &= \mathbf{h}^H\mathbf{X}^H\mathbf{X}\mathbf{h} - 2\Re\{\mathbf{y}^H\mathbf{X}\mathbf{h}\} + \mathbf{y}^H\mathbf{y} \\ &= \mathbf{h}^H\mathbf{A}\mathbf{h} - 2\Re\{\mathbf{b}^H\mathbf{h}\} + c \end{aligned} \tag{3}$$

where $\mathbf{X}$ is a matrix containing kernel values, $\mathbf{y}$ is a vector of the weighted ideal response, and $\mathbf{h}$ is a vector of the filter coefficients. Since $\mathbf{A}$, $\mathbf{b}$, c are fixed throughout the design, we can precompute their values, and the computation will be significantly reduced.

For FIR filter design, the system matrix $\mathbf{A}$ for the gradient equation is a Toeplitz matrix, so it can be solved efficiently by using the Levinson recursion. For IIR filter design, the algorithm has to be modified further as described in the next subsection.

4.1. IIR Filter Design

To use the new algorithm for IIR design, there are two approaches: find a local optimum by finding a solution with zero gradient, or find a suboptimal solution by linearizing the problem. Due to the limited space here, the details of the zero gradient approach will be omitted. Although the linearization method yields a design that is suboptimal, the solution has nearly the same error while the algorithm complexity is much less. The linearization is done by treating the problem $\min \|W(I - \frac{B}{A})\|_2$ as a linear problem $\min \|\frac{W}{|A|}(IA - B)\|_2$ where $|A|$ is fixed [7], so all of the weighted least squares subproblems can be solved as linear problems. However, the linearized IIR problem no longer possesses a gradient system matrix that is Toeplitz. Therefore, to make the algorithm efficient, the feedforward and feedback coefficients need to be computed separately. The feedforward coefficients are computed using Levinson's recursion, while the feedback coefficients are computed by using a generic numerical method such as the QR decomposition.



(a) (b)

Figure 2: Different error norms obtained when minimizing the combined norm, as a function of α. (a) Chebyshev norm (solid line), least-squares norm (dashed line), and the combined norm (dotted line) versus α. (b) The trade-off between the Chebyshev norm and least-squares norm.

4.2. Computation

The algorithms for both FIR and IIR design usually require less than ten exchange iterations, where each exchange iteration requires only a few tens of IRLS iterations. The IRLS subproblem may be solved efficiently by the Levinson recursion, $O(N^2)$ for computing the feedforward coefficient plus an additional $O(M^3)$, for solving the feedback coefficient. Since the algorithm is an $O(N^2)$ operator (assuming $N \gg M$), it will also be efficient for designing medium sized 2-D filters with this new combined norm definition.

5. EXAMPLES

This section shows some filters designed by the new algorithm for the case where the ideal specifications are: lowpass filter with a cutoff frequency at 0.4 (normalized frequency), a transition bandwidth of 0.1, a ratio of passband error to stopband error equal to 10, and a group delay of 13 samples.

Figure 2 shows how the three norms (Chebyshev, Least-squares and combined norm) depend on the parameter α for one case: an FIR filter of order 20. As expected, the Chebyshev norm decreases as α increases, while the least-squares norm increases. Figure 2 also shows the trade-off between the Chebyshev norm and the 2-norm versus α. This behavior is identical to the plot for the constrained least squares method [1, 2]. The combined-norm problem seems to be biased toward the Chebyshev norm, so a value of α between 0.2 and 0.4 is the best compromise for designing a filter with both small RMS error and maximal error. For the same filter specification, an IIR design is shown in Fig. 3. The filter has 16 zeros and 4 poles and was designed by the linearization approach with $\alpha = 0.5$. The IIR filter has a maximal error that is about 14 dB less than the order-20 FIR filter.

To compare the algorithm performance to some well-known methods, Table 1 summarizes the characteristics of several different filter designs. First, the least squares filter (i.e., $\alpha = 0$) has the lowest computation requirement. The Chebyshev filter ($\alpha = 1$) requires more computation because it contains the IRLS iteration. The computational requirement for the combined norm ($\alpha = 0.5$) is more than the Chebyshev problem because of the additional 2-norm computation during the IRLS iteration. Using the linearized IIR method, the IIR filter requires about the same amount of computation as the FIR filter, but the IIR filter has a much better fre-

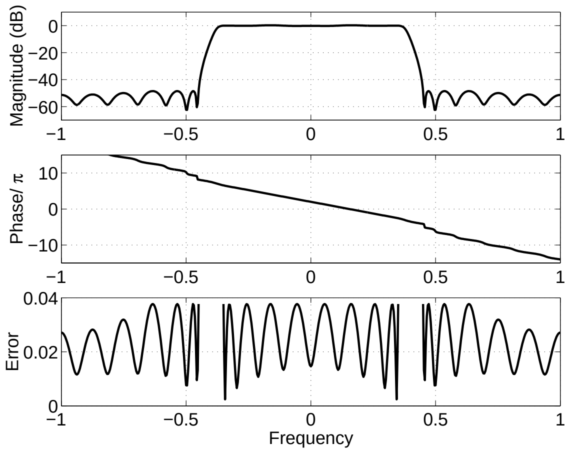


Figure 3: Frequency response of an IIR filter.

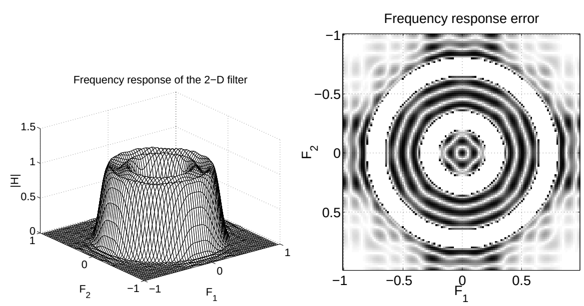


Figure 4: Frequency response of a 2-D filter. The gray-scale image makes it easy to see the equal-height errors in the pass bands.

Filter	#Mflops	max. error	RMS error
$N=51, M=0, \alpha=0$	3.20	0.0926	0.0139
$N=51, M=0, \alpha=1$	33.8	0.0380	0.0257
$N=51, M=0, \alpha=0.5$	56.0	0.0389	0.0234
$N=47, M=4, \alpha=0.5$	109	0.0070	0.0047
$N=43, M=8, \alpha=0.5$ suboptimal solution	178	0.0020	0.0014
$N=43, M=8, \alpha=0.5$ local optimal solution	645	0.0020	0.0014
linear-phase cremez [6]	3.46	0.0386	0.0259
nonlinear-phase cremez [6]	741	0.0383	0.0262
Lang's peak constrained [2]	11.9	0.0389	0.0234

Table 1: Comparison of filter designed. The filters are designed to be bandpass filter order 51 with stopband in $[0, 0.3] \cup [0.7, 1]$ and passband $[0.35, 0.65]$ with group delay of 30 samples. The weighting is equal for both stopbands and the passband. The filters were designed with different numbers of zeros, N, and different numbers of poles, M, and norm weighting parameters, α

quency response. For the locally optimal IIR solution, the computation is about three times greater than the linearized method but the norms are nearly identical—the optimized combined norm is slightly smaller (1.6819×10^{-3} versus 1.6821×10^{-3}). In addition, the frequency response of the solutions obtained by the new algorithm are equivalent to, or better than the those obtained by the other available algorithms for the FIR case [2, 6]. For the proposed algorithm, the computational requirements are sometimes greater because the new algorithm always deals with the complex case which inherently requires four times as much computation as an algorithm that is restricted to the real case.

The last part of this section shows the design for a 2-D bandpass filter when the passband is an annulus bounded by rings of radii 0.35 and 0.65 (in normalized frequency), the inner stopband is a circle with radius 0.2, and the outer stopband is a ring of radius greater than 0.8. Figure 4 shows the frequency response and error of a 19×19 2-D FIR filter. The design needs $O(10^{10})$ flops due to the large number of parameters being optimized.

6. CONCLUSIONS

A new filter design method was formulated based on the simultaneous minimization of a combined norm that is the weighted sum of the 2-norm and the Chebyshev norm. The combined norm problem is a norm, and it is also generalizes the Chebyshev problem, the least squares problem, and the peak constrained least squares problem, so it possess desirable properties for many applications. The new problem is solved by a multiple frequency exchange with subproblems solved by the iterative reweighted least squares method. The new algorithm is directly applicable to the design of complex filters, IIR filters, and 2-D filters. The amount of computation for the algorithm is $O(N^2)$ which is efficient for the design of high-order filters.

7. REFERENCES

[1] J. W. Adams, "FIR digital filters with least-squares stopbands subject to peak-gain constraints," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 376–388, Apr. 1991.

[2] M. C. Lang, "Multiple exchange algorithm for constrained design of FIR filters in the complex domain," *IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 1149–1152, 1999.

[3] J. H. McClellan, "A unified approach to the design of optimal FIR linear-phase digital filters," *IEEE Trans. Circuit Theory*, vol. CT20, pp. 697–701, Nov. 1973.

[4] C. Y. Tseng, "An efficient implementation of Lawson's algorithm with application to complex Chebyshev FIR filter design," *IEEE Trans. Circuits Syst. II*, vol. 42, pp. 245–260, Apr. 1995.

[5] C. L. Lawson "Contributions to the Theory of Linear Least Minimum Approximations," Ph. D. dissertation, Univ. Calif., Los Angeles, 1961.

[6] L. J. Karam and J. H. McClellan, "Chebyshev digital FIR filter design," *Signal Processing*, vol. 76, pp. 17–36, 1999.

[7] K. Steiglitz and L. E. McBride, "A technique for the identification of linear systems," *IEEE Trans. Automatic Control*, vol. 10, pp. 461–464, Oct. 1965.