Minimax Design of Two-Dimensional FIR Digital Filters by Using an Interior-Point Algorithm

Ju-Hong Lee   and   Charng-Kann Chen

Department of Electrical Engineering
National Taiwan University
Taipei, Taiwan, R. O. C.

Abstract

This paper considers the design of two dimensional (2-D) linear phase FIR digital filters optimal in the minimax sense. A design method based on an affine-scaling variant of Karmarkar's linear programming algorithm is presented. In the design process, we first formulate the design problem into a linear programming form. To avoid the huge computation load and storage space required by using a standard simplex algorithm, we present a Karmarkar's algorithm based method to solve the design problem. In each iteration of the proposed method, we only need to calculate a weighted least square (WLS) solution. Nearly optimal solutions can be obtained after several iterations. Design examples and comparison are presented to show the effectiveness of the proposed method.

I.   Introduction

Two-dimensional (2-D) FIR digital filters have been received much attention owing to their applications in the areas, such as image processing, radar, and geophysical data processing. The reasons for their versatility include the guaranteed stability of FIR filters, the ability to achieve an exactly linear phase response, and relative low coefficient sensitivity. However, the minimax design of 2-D linear phase FIR filters is not an easy task. It is due to many existing theoretical and practical difficulties in the design of 2-D FIR optimal filters.

In recent years, many techniques for designing 2-D linear phase FIR filters optimal in the minimax sense have been reported in the literature [1]-[8]. Notable among them for producing an equiripple design is the McClellan transformation method [1]. Although this method is simple and fast, it is only suboptimal in the minimax sense and can not closely approximate all kinds of magnitude responses. In [3],[4], the Remez exchange algorithm is extended for designing of 2-D FIR filters, several Remez exchange type procedures were developed. However, the basis functions do not satisfy the Haar condition and the alternation theorem can not be applied to the 2-D case. Hence, algorithms of this kind suffer from heavy computation load and its convergence to a correct solution has not been demonstrated when used to design optimal filters [5]. The weighted least square (WLS) technique has been shown to be an efficient method for designing 1-D FIR filters optimal in the minimax sense [6],[9]. Motivated by the success in the 1-D case, several authors extend the WLS approach to the 2-D case by slightly modifying the Lawson's algorithm [10]. Unfortunately, such extension results in a very slow algorithm, the iteration number is very large even when the filter size is small [6], [7].

Basically, the minimax design involves the evaluation of the resulting frequency response at a great amount of frequency samples in the considered frequency bands. Evaluating at each sample produces a linear equation which describes the relationship between the design parameters and the desired magnitude response at that frequency point. To ensure the performance of the designed filter, the number of the sample points must be much larger than the number of the independent parameters. Thus, the design problem involves how to find the minimax solution of a large overdetermined linear system. General speaking, most of the above design techniques concern such a problem but use different approaches to solve it. For a 2-D minimax design problem, one way to obtain the optimal solution is using the linear programming (LP) technique [2]. Conventionally, the linear programming is in conjunction with the simplex method. However, directly using the simplex method to solve the LP problem is very inefficient.

Recently, Karmarkar [10] proposed an algorithm for solving LP problems. It has been shown that the algorithm and its variants [11] can efficiently solve high dimensional LP problems. Based on the concept of projected gradient, an LP problem can be solved efficiently. In this paper, based on a simple variant of Karmarkar's algorithm presented in [11], we derive a WLS design procedure for solving the minimax 2-D linear phase FIR design problem. The nearly optimal solution can be obtained after several iterations. Compared with existing methods, the proposed WLS method has shown a great improvement both on the Chebyshev error and the iteration number.

This paper is organized as follows. Section II formulates the design problem of 2-D linear phase FIR filters. In Section III, a relative simple ("affine-scaling") variant of Karmarkar's algorithm is presented. Based on this simple algorithm, we propose a WLS design method for 2-D FIR filters in Section IV. Several design examples for illustration and comparison are presented in Section V. Finally, we conclude this paper in Section VI.

II.   Problem Formulation

The frequency response of a causal 2-D FIR digital filter with impulse response sequence $h(n_1, n_2)$, $n_1 = 0, 1, \cdots, N_1 - 1$, $n_2 = 0, 1, \cdots, N_2 - 1$ can be expressed as

$$H(\omega_1, \omega_2) = \sum_{n1=0}^{N_1-1} \sum_{n2=0}^{N_2-1} h(n_1, n_2) e^{-j\omega_1 n_1} e^{-j\omega_2 n_2}. \qquad (1)$$

To achieve a linear phase response, the impulse response $h(n_1, n_2)$ must be symmetric about $(\frac{N_1-1}{2}, \frac{N_2-1}{2})$, i.e., $h(n_1, n_2) = h(N_1 - 1 - n_1, N_2 - 1 - n_2)$. Then, we can express the magnitude response of a 2-D linear phase FIR filter as

$$H_M(\omega_1, \omega_2) = \sum \sum a(n_1, n_2) \cos(\omega_1 n_1 + \omega_2 n_2), \qquad (2)$$

where the coefficient $a(n_1, n_2)$ is related to the impulse response of the filter, and the region of support of $a(n_1, n_2)$ is about one half of that of $h(n_1, n_2)$. If there is additional symmetry in $H_M(\omega_1, \omega_2)$, for example, quadrantal or octagonal symmetry, the number of independent $a(n_1, n_2)$ can be further reduced [8]. In general, the magnitude response $H_M(\omega_1, \omega_2)$ can be written as

$$H_M(\omega_1, \omega_2) = \sum_{i=0}^{N-1} a_i \phi_i(\omega_1, \omega_2), \qquad (3)$$

where N is the number of the independent parameters in $a(n_1, n_2)$, a_i is a parameter in the set of $a(n_1, n_2)$ for a particular (n_1^i, n_2^i), and $\phi_i(\omega_1, \omega_2)$ is the corresponding basis function. Let $\omega = (\omega_1, \omega_2)$ be a particular sample point in a compact approximating frequency set K of the 2-D frequency plane, and $H_d(\omega)$ be the desired magnitude response at that point. Taking L samples on a dense grid of frequency uniformly distributed over K, namely, $\omega_1, \cdots, \omega_L$, the design problem is to solve the following overdetermined linear equations in the minimax sense

$$\mathbf{Ua} = \mathbf{H_d} \qquad (4)$$

where

$$\mathbf{U} = \begin{pmatrix} \phi_0(\omega_1) & \phi_1(\omega_1) & \cdots & \phi_{N-1}(\omega_1) \\ \phi_0(\omega_2) & \phi_1(\omega_2) & \cdots & \phi_{N-1}(\omega_2) \\ \vdots & \ddots & \ddots & \vdots \\ \phi_0(\omega_L) & \phi_1(\omega_L) & \cdots & \phi_{N-1}(\omega_L) \end{pmatrix},$$

$$\mathbf{H_d} = (H_d(\omega_1), H_d(\omega_2), \cdots, H_d(\omega_L))^T,$$

$$\mathbf{a} = (a_0, a_1, \cdots, a_{N-1})^T.$$

Solving the minimax solution of (4) is equivalent to solving the following constrained minimization problem:

$$\begin{aligned} \text{Minimize} \quad & \delta \\ \text{Subject to} \quad & |\mathbf{Ua} - \mathbf{H_d}| \leq \delta \mathbf{1}, \end{aligned} \qquad (5)$$

where $\mathbf{1}$ denotes an $L \times 1$ vector whose elements are 1. Define the following matrices

$$\begin{aligned} \mathbf{b} &= [\mathbf{0}^T \ 1]^T, \\ \mathbf{w} &= [-\mathbf{a}^T \ -\delta]^T, \\ \mathbf{c} &= [-\mathbf{H_d}^T \ \mathbf{H_d}^T]^T, \\ \mathbf{A} &= \begin{bmatrix} \mathbf{U}^T & -\mathbf{U}^T \\ \mathbf{1}^T & \mathbf{1}^T \end{bmatrix}, \end{aligned} \qquad (6)$$

where $\mathbf{0}$ denotes an $N \times 1$ zero vector. The optimization problem in (5) can be rewritten as

$$\begin{aligned} \text{Maximize} \quad & \mathbf{b}^T \mathbf{w} \\ \text{Subject to} \quad & \mathbf{A}^T \mathbf{w} \leq \mathbf{c}. \end{aligned} \qquad (7)$$

(7) is a dual form LP problem. Basically, it can be solved optimally by using a standard simplex LP package. However, due to the constraint matrix $\mathbf{A}$ is quite large (with size $(N+1) \times 2L$, and L is about 10^3), directly using the simplex algorithm always needs large storage space and converges very slowly. Next, we describe how to use a variant of Karmarkar's algorithm to efficiently solve (7).

III. A Modified Karmarkar's LP Algorithm

In this section, we briefly describe the modified Karmarkar's LP algorithm proposed in [11]. The algorithm deals with the primal form problem as follows

$$\begin{aligned} \text{Minimize} \quad & \mathbf{c}^T \mathbf{x} \\ \text{Subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}. \end{aligned} \qquad (8)$$

Note that (7) and (8) are the primal-dual pair LP problem. The basic three ideas of the modified Karmarkar's algorithm for solving (8) are: (i). The algorithm operates from an interior point of the feasible region toward to the optimal solution rather than from the vertex points of the feasible region. (ii). The algorithm moves the current interior point in a direction that decreases the cost function at the fastest rate, while keeps the resulting new interior point satisfy the constraint $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$. The projected gradient concept can be used. (iii). A scheme is applied to transform the feasible region to place the current trial solution near to its center, thereby giving a substantial reduction in the objective function when implementing (ii).

Next, we describe how to use the above three ideas to solve (8). Assume that we have a current trial solution $\mathbf{x}$ which satisfies $\mathbf{A}\mathbf{x} = \mathbf{b}$, and $\mathbf{x} > 0$. Using the idea (iii), we first transform the feasible region to place the current trial solution near to its center. Define $\mathbf{D} = \text{diag}(\mathbf{x})$, $\tilde{\mathbf{A}} = \mathbf{AD}$, and $\tilde{\mathbf{c}} = \mathbf{Dc}$, it is obvious that $\tilde{\mathbf{x}} = \mathbf{D}^{-1}\mathbf{x} = \mathbf{1}$ is a feasible solution of the following affine scaling form of (8)

$$\begin{aligned} \text{Minimize} \quad & \tilde{\mathbf{c}}^T \tilde{\mathbf{x}} \\ \text{Subject to} \quad & \tilde{\mathbf{A}} \tilde{\mathbf{x}} = \mathbf{b}, \quad \tilde{\mathbf{x}} \geq \mathbf{0}, \end{aligned} \qquad (9)$$

and $\tilde{\mathbf{x}}$ is also at the center of the feasible region of (9).

With the feasible solution $\tilde{\mathbf{x}}$ of (9), next step is to improve this solution. To decrease the objective function while satisfying the constraints, we can move $\tilde{\mathbf{x}}$ in the direction

$$\mathbf{d} = -(\mathbf{I} - \tilde{\mathbf{A}}(\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T)^{-1}\tilde{\mathbf{A}})\tilde{\mathbf{c}}. \qquad (10)$$

Here, the projected gradient concept is used to find the descent direction. Then, we determine the step size β such that $\tilde{\mathbf{c}}^T(\tilde{\mathbf{x}} + \beta\mathbf{d})$ is smallest and keep $\tilde{\mathbf{x}} + \beta\mathbf{d} > 0$. If v is the absolute value of the smallest element of $\mathbf{d}$, β can be chosen as $0.99/v$. Finally, we calculate $\mathbf{x} = \mathbf{D}(\tilde{\mathbf{x}} + \beta\mathbf{d})$ as the trial solution for the next iteration.

Based on the above description, an algorithm which solves (8) can be summarized as follows:

Algorithm 1: algorithm for solving (8)

Initialization: $\mathbf{x}$: Any solution such that $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{x} > 0$,
Main loop:

1. $\mathbf{D} = \text{diag}(\mathbf{x})$
2. $\mathbf{w} = (\mathbf{AD}^2\mathbf{A}^T)^{-1}\mathbf{AD}^2\mathbf{c}$
3. If a preset stopping criterion is satisfied, stop this algorithm.
4. $\mathbf{d} = -\mathbf{D}^2(\mathbf{c} - \mathbf{A}^T\mathbf{w})$
5. β = absolute value of $0.99/\min(\mathbf{D}^{-1}\mathbf{d})$
6. $\mathbf{x} = \mathbf{x} + \beta\mathbf{d}$

End loop.

The WLS solution is calculated in step 1 and step 2. Step 4 determines the descent direction. Step 5 decides the step size and we update the current trial solution at step 6. The main computation load of Algorithm 1 is the calculation of the WLS solution. If the constraint matrix $\mathbf{A}$ is sparse or possesses some special structure, then it is possible to further reduce the computation load in calculating the WLS solution. Although Algorithm 1 is an iterative procedure for solving the primal problem, it also solves the dual problem. As the primal variable $\mathbf{x}$ converges to the optimal solution of (8), the variable $\mathbf{w}$ also optimally solves (7) [11]. In other words, we can also use Algorithm 1 to solve (7). These observations lead to the development of an efficient algorithm for designing the optimal 2-D linear phase FIR filters.

IV. The Proposed WLS Method

In this section, we derive a WLS procedure to solve the 2-D linear phase FIR design problem shown in (7). This procedure is obtained by substituting the matrices defined in (6) into Algorithm 1. The detail of the proposed design procedure is described step by step as follows:

Minimax design procedure of 2-D linear phase FIR filters

Step 1: Let $\mathbf{x}$ be composed of two $L \times 1$ vectors $\mathbf{x_1}$, $\mathbf{x_2}$, i.e., $\mathbf{x} = [\mathbf{x_1}^T \ \mathbf{x_1}^T]^T$. The simplest way for choosing an arbitrary $\mathbf{x}$ such that $\mathbf{Ax} = \mathbf{b}$, $\mathbf{x} > 0$ is to let $\mathbf{x_1} = \mathbf{x_2} = \frac{1}{2L}\mathbf{1}$.

Step 2: Calculate the WLS solution $\mathbf{w} = (\mathbf{AD^2A^T})^{-1}\mathbf{AD^2c}$. This can be done by using the following procedures:

$\quad$ (2.1). $\mathbf{f} = \text{diag}(\mathbf{x_1})\mathbf{x_1} - \text{diag}(\mathbf{x_2})\mathbf{x_2}$
$\quad$ (2.2). $\mathbf{W} = \text{diag}(\text{diag}(\mathbf{x_1})\mathbf{x_1} + \text{diag}(\mathbf{x_2})\mathbf{x_2})$
$\quad$ (2.3). $\mathbf{u} = \mathbf{U^Tf}$
$\quad$ (2.4). $w = -(\mathbf{1^TW1})^{-1}$
$\quad$ (2.5). $h_d = \mathbf{f^TH_d}$
$\quad$ (2.6). $\mathbf{a} = (\mathbf{U^TWU} + \mathbf{u}w\mathbf{u^T})^{-1}(\mathbf{U^TWH_d} + \mathbf{u}wh_d)$
$\quad$ (2.7). $\mathbf{Ae} = \mathbf{H_d} - \mathbf{Ua}$
$\quad$ (2.8). $\delta = -w\mathbf{f^TAe}$.

Note that $\mathbf{w} = [-\mathbf{a}^T \ -\delta]^T$.

Step 3: Note that $\max(|\mathbf{Ae}|)$ denotes the Chebyshev error of the current trial solution $\mathbf{a}$. This quantity is nearly equal to δ when the algorithm converges to the optimal $\mathbf{a}$. Thus, If $|\frac{\delta}{\max(|\mathbf{Ae}|)}| \geq 0.999$, then we stop the iterative process. Otherwise, go to the next step.

Step 4: Calculate the descent direction $\mathbf{d} = -\mathbf{D^2}(\mathbf{c} - \mathbf{A^Tw})$. If $\mathbf{d}$ is composed of two $L \times 1$ vectors $\mathbf{d_1}$, $\mathbf{d_2}$, i.e., $\mathbf{d} = [\mathbf{d_1}^T \ \mathbf{d_2}^T]^T$. Then

$\quad$ (4.1). $\mathbf{d_1} = -\text{diag}(\text{diag}(\mathbf{x_1})\mathbf{x_1})(\delta\mathbf{1} - \mathbf{Ae})$
$\quad$ (4.2). $\mathbf{d_2} = -\text{diag}(\text{diag}(\mathbf{x_2})\mathbf{x_2})(\delta\mathbf{1} + \mathbf{Ae})$

Step 5: Determine the step size β as follows

$$\beta = \left| \frac{0.99}{\min(\min(\text{diag}(\mathbf{x_1})^{-1}\mathbf{d_1}), \min(\text{diag}(\mathbf{x_2})^{-1}\mathbf{d_2}))} \right|$$

Step 6: Update the current trial solution $\mathbf{x}$ by setting

$\quad$ (6.1). $\mathbf{x_1} = \mathbf{x_1} + \beta\mathbf{d_1}$
$\quad$ (6.2). $\mathbf{x_2} = \mathbf{x_2} + \beta\mathbf{d_2}$

and go to *step 2*.

V. Design Examples

In this section, we present several computer simulations of designing 2-D linear phase FIR digital filters using the proposed technique. All the simulations for the design examples are performed on an 117-MHz 80486 personal computer using an MATLAB programming language.

Example 1
In this example, we design a 27×27 circularly symmetric bandpass filter with two passband edges 0.4π and 0.6π, two stopband edges 0.2π and 0.8π. The algorithm converges after 22 iterations with a Chebyshev error 0.0033. The magnitude responses in linear and dB scale are shown in Fig. 1(a) and Fig. 1(b), respectively. From

Fig. 1(b), we can see that this design is nearly equiripple. The same example has been designed by a least square approach [13] which results in a solution with a peak error 0.031.

Example 2
In the second example, we compare the performance of our proposed algorithm and the design algorithm in [7]. The main difference between these two methods is the strategy for selecting the least square error weighting function. As in the Example 1 of [7], we design a circularly symmetric lowpass filter with passband edge 0.4π and stopband edge 0.6π. The design results with different filter size are shown in Table I. It can be seen that the proposed algorithm significantly outperforms that of [7].

VI. Conclusion

This paper has presented an efficient technique for designing 2-D linear phase FIR filters optimal in the minimax sense. Based on a simple variant of Karmarkar's linear programming algorithm, we can easily design a nearly optimal FIR filter by calculating several weighted least square solutions. The proposed method does not have the degeneracy problem existing in exchange ascent algorithms. Computer simulations show that the proposed design technique can produce better results in fewer iterations than existing WLS methods.

References

[1] J. H. McClellan, "The design of two-dimensional digital filter by transformations," in Proc. 7th Ann. Princeton Conf. Inform. Sci. Syst., pp. 247-251, 1973.

[2] J. V. Hu and L. R. Rabiner, "Design techniques for two-dimensional digital filters," IEEE Trans. Audio Electroacoust., vol. AU-20, pp. 249-257, Oct. 1972.

[3] Y. Kamp and J. P. Thiran, "Chebyshev approximation for two-dimensional nonrecursive digital filters," IEEE Trans. Circuits Syst., vol. CAS-22, pp.208-218, Mar. 1975.

[4] D. B. Harris and R. M. Mersereau, "A comparison of algorithms for minimax design of two-dimensional linear phase FIR digital filters," IEEE Trans. Acoust., Speech, Signal Processing, ASSP-25, pp.492-500, Dec. 1977.

[5] J. S. Lim, Two-Dimensional Signal and Image Processing. Englewood Cliffs, NJ: Prentice-Hall, 1990.

[6] V. R. Algazi, M. Suk, and C. S. Rim, "Design of almost minimax FIR filters in one and two dimensions by WLS techniques," IEEE Trans. Circuits Syst., vol. CAS-33, pp.590-596, Jun. 1986.

[7] C.-Y. Chi and S.-L. Chiou, "A new self-initiated WLS approximation method for the design of two-dimensional equiripple FIR digital filters," Proc. IEEE 1992 ISCAS., pp. 1436-1439, San Diego, 1992.

[8] C. Charalambous, "The performance of an algorithm for minimax design of two-dimensions linear phase FIR digital filters," IEEE Trans. Circuits Syst., vol. CAS-33, pp.590-596, Jun. 1986.

[9] Y. C. Lim, J.-H. Lee, C.-K. Chen and R. H. Yang, "A weighted least squares algorithm for quasi-equiripple FIR and IIR digital filter design," IEEE Trans. Signal Processing., vol. ASSP-40, pp. 551-558, Mar. 1992.

[10] C. L. Lawson, "Contributions to the theory of linear least maximum approximations," Ph. D. dissertation, U.C.L.A., 1961.

[11] N. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica*, vol. 4, pp. 373-395, 1984.

[12] R. J. Vanderbei, M. S. Meketon, and B. A. Freeman, "A modification of Karmarkar's linear programming algorithm," *Algorithmica*, vol. 1, pp. 395-407, 1986.

[13] S.-C. Pei and J.-J. Shyu, "2-D FIR eigenfilters: a least-squares approach," *IEEE Trans. Circuits Syst.*, vol. CAS-37, pp.24-34, Jan. 1990.

(a)

(b)
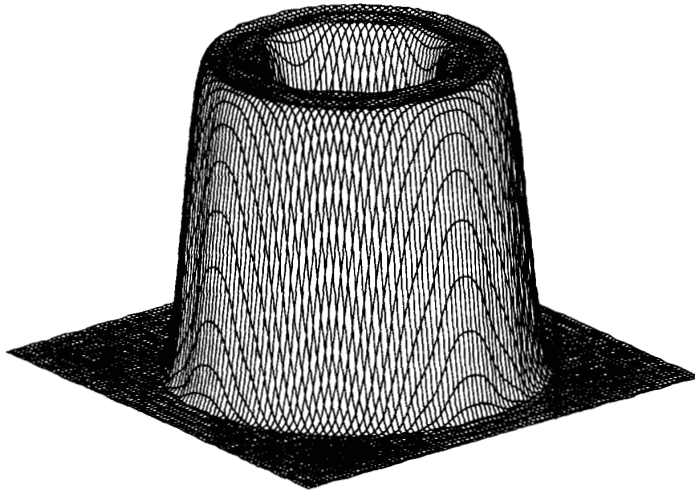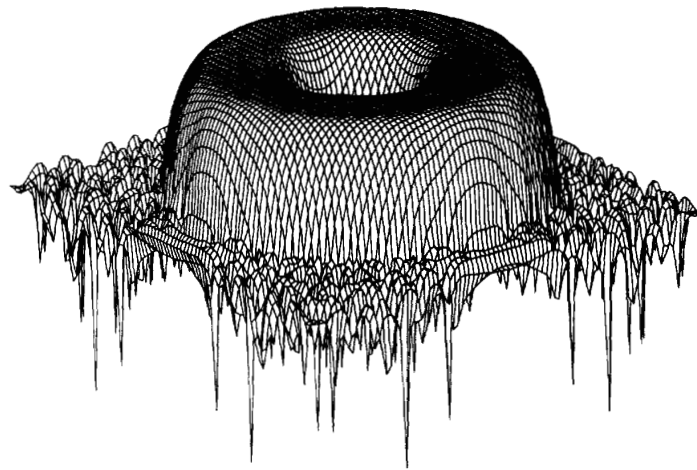
Fig. 1

TABLE I

Filter size	Algorithm	Maximum error	Iteration number
7 × 7	Proposed method	0.1270	16
	Chi & Chiou[7]	0.1281	51
9 × 9	Proposed method	0.1141	15
	Chi & Chiou[7]	0.1151	36
11 × 11	Proposed method	0.0558	19
	Chi & Chiou[7]	0.0577	174