Multidim Syst Sign Process (2013) 24:417–434
DOI 10.1007/s11045-011-0169-9

Efficient 2-D based algorithms for WLS designs of 2-D FIR filters with arbitrary weighting functions

Ruijie Zhao · Xiaoping Lai

Received: 24 August 2011 / Revised: 27 November 2011 / Accepted: 1 December 2011 /
Published online: 10 December 2011

Abstract The impulse response coefficients of a two-dimensional (2-D) finite impulse response (FIR) filter naturally constitute a matrix. It has been shown by several researchers that, two-dimension (2-D) based algorithms that retain the natural matrix form of the 2-D filter's coefficients are computationally much more efficient than the conventional one-dimension (1-D) based algorithms that rearrange the coefficient matrix into a vector. In this paper, two 2-D based algorithms are presented for the weighted least squares (WLS) design of quadrantally symmetric 2-D FIR filters with arbitrary weighting functions. Both algorithms are based on matrix iterative techniques with guaranteed convergence, and they solve the WLS design problems accurately and efficiently. The convergence rate, solution accuracy and design time of these proposed algorithms are demonstrated and compared with existing algorithms through two design examples.

Keywords 2-D FIR filter · 2-D based algorithm · Linear operator · Weighted least squares design

1 Introduction

Two-dimensional (2-D) digital filters have been widely applied in image processing, sonar and radar signal processing, geophysical signal processing and so on (Lim 1990; Lu and Antoniou 1992). The minimax and weighted least squares (WLS) are two frequently used criterions for optimal designs of 2-D finite impulse response (FIR) filters. Usually, a WLS design obtains smaller energy of the magnitude error and consumes less time than a minimax

R. Zhao (✉)
School of Mechanical, Electrical and Information Engineering, Shandong University at Weihai, Weihai 264209, China
e-mail: zhao_rj@163.com

X. Lai
Institute of Information and Control, Hangzhou Dianzi University, Hangzhou 310018, China
e-mail: laixp@hdu.edu.cn

design, although it may result in larger maximum magnitude error than the minimax design. In addition, a WLS design may obtain solutions close to the minimax filter by choosing a proper weighting function or by using an iterative reweighting technique. Therefore, efficient and numerically stable WLS algorithms are considerably important for further research of 2-D FIR filter design.

High computational complexity is a major problem encountered in optimal designs of 2-D FIR filters. Unlike the conventional one-dimension (1-D) based design algorithms that rearrange the filter's coefficient matrix into a vector, two-dimension (2-D) based algorithms retain the 2-D filter's coefficient matrix in its natural form, leading to a considerable reduction in computational complexity and memory space. However, up to the authors' knowledge, most of the design algorithms for 2-D FIR filters advanced so far, such as Gislason et al. (1993), Lu (2002), Tzeng (2007), Lai and Cheng (2007), Shyu et al. (2011), Lu and Hinamoto (2011), are 1-D based. This has greatly limited the capability of fast designing 2-D FIR filters, especially with high orders.

Several 2-D based algorithms for 2-D FIR filter designs have been reported in the literature, see, e.g., Zhu et al. (1997), Hsieh et al. (1997), Gu and Aravena (1994), Aravena and Gu (1996), Zhao and Lai (2010, 2011). In Zhu et al. (1997), an analytical solution to the unweighted LS design of 2-D FIR filters has been obtained. But the unweighted LS solution usually has heavy magnitude overshoot near the cutoff frequency of the frequency response. The iterative 2-D based WLS algorithm proposed by Hsieh et al. (1997) for the minimax design of 2-D FIR filters is efficient but may not converge. Gu and Aravena (1994) developed two iterative algorithms for the WLS design of 2-D FIR filters. But the weight function can only be 0- and 1-valued. In Aravena and Gu (1996), they presented other two WLS algorithms, a fix point algorithm and a conjugate gradient algorithm, where the weighting functions can be arbitrarily nonnegative valued. Their algorithms are much more efficient than the conventional 1-D based algorithms, but the design time is still long. In addition, only equally spaced frequency grid can be used in their algorithms. Recently, the authors developed a 2-D based algorithm in Zhao and Lai (2011) for the WLS design of quadrantally symmetric FIR filters based on a matrix iterative technique. The algorithm converges very fast, but the weighting function is two-valued.

In this paper, we will investigate the WLS design of 2-D FIR filters with arbitrary weighting functions and present two matrix iterative algorithms for the design problem. Some preliminary results of the first algorithm were presented in Zhao and Lai (2010). Simulation results in Zhao and Lai (2010) show that the algorithm is more efficient and flexible than those of Aravena and Gu (1996) in general, especially for rectangular filters. However, we find that the iteration number of the algorithm increases dramatically with the filter order and transition band width for several types of 2-D filters, such as elliptic and fan filters. In this paper, we introduce an appropriate scalar parameter into the algorithm to reduce the spectral radius of corresponding iterative operator, resulting in an improved algorithm with a faster convergence rate. In addition, we observe from simulations that the improved algorithm returns more accurate solutions. Two design examples are presented to demonstrate the effectiveness of the algorithms. Comparisons with existing algorithms in terms of iteration number, design time and design accuracy are provided.

This paper is organized as follows. In Sect. 2, the WLS design problem of 2-D quadrantally symmetric FIR filters is formulated as an optimization problem with a cost function expressed in terms of the filter's coefficient matrix, and a theorem for the uniqueness condition of the optimal solution to the design problem is established. The two efficient 2-D based algorithms for the WLS design problem are presented in Sect. 3. An iterative procedure for the largest eigenvalue of a nonnegative, compact, and self-adjoint operator defined on a

finite-dimensional Hilbert space is also developed in Sect. 3. In Sect. 4, two design examples are provided to demonstrate the properties of the presented algorithms and comparisons are made with existing algorithms. Finally, the conclusions are drawn in Sect. 5.

2 WLS design problem of 2-D quadrantally symmetric FIR filters

Consider an $N_1 \times N_2$ 2-D FIR filter with real impulse response $\{h(n_1, n_2), n_1 = 0, 1, \ldots, N_1 - 1;\ n_2 = 0, 1, \ldots, N_2 - 1\}$, whose frequency response can be expressed as

$$H(e^{j\omega_1}, e^{j\omega_2}) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} h(n_1, n_2) e^{-j(n_1\omega_1 + n_2\omega_2)}, \tag{1}$$

where ω_1 and ω_2 are the horizontal and vertical frequencies, respectively. For convenience, the frequency response (1) is often expressed in terms of its magnitude and phase responses as

$$H(e^{j\omega_1}, e^{j\omega_2}) = G(\omega_1, \omega_2) e^{j\varphi(\omega_1, \omega_2)},$$

where $G(\omega_1, \omega_2)$ is the magnitude response and $\varphi(\omega_1, \omega_2)$ is the phase response, both of which are real functions of ω_1 and ω_2. It is well known that if the impulse response $h(n_1, n_2)$ satisfies the restriction: $h(n_1, n_2) = h(n_1, N_2 - 1 - n_2) = h(N_1 - 1 - n_1, n_2)$, then the phase response $\varphi(\omega_1, \omega_2)$ is a linear function of ω_1 and ω_2, and the magnitude response $G(\omega_1, \omega_2)$ is symmetric with respect to both of the ω_1 and ω_2 axes. Such 2-D filters are referred to as quadrantally symmetric filters.

We can write the magnitude response of a quadrantally symmetric filter in a vector-matrix form as

$$G(\omega_1, \omega_2, \mathbf{A}) = \phi^T(\omega_1, N_1) \mathbf{A} \phi(\omega_2, N_2) \tag{2}$$

where the superscript "T" denotes the transpose operation, $\phi(\omega, N)$ is a vector given by

$$\phi(\omega, N) = \begin{cases} \left[\frac{1}{\sqrt{2}}, \cos(\omega), \cos(2\omega), \ldots, \cos\left(\frac{N-1}{2}\omega\right)\right]^T, & \text{for odd } N, \\ \left[\cos\left(\frac{\omega}{2}\right), \cos\left(\frac{3\omega}{2}\right), \ldots, \cos\left(\frac{N-1}{2}\omega\right)\right]^T, & \text{for even } N, \end{cases} \tag{3}$$

and $\mathbf{A} = (a_{ij})$ is a real coefficient matrix related to the impulse response $h(n_1, n_2)$. It follows from (2) and (3) that the dimension of matrix $\mathbf{A}$ is $L_1 \times L_2$, where

$$L_k = \begin{cases} N_k/2, & \text{for even } N_k \\ (N_k + 1)/2, & \text{for odd } N_k \end{cases};\ k = 1, 2.$$

The entries of $\mathbf{A}$ are completely determined by the impulse response $\{h(n_1, n_2)\}$, and vice versa. For example, if both N_1 and N_2 are even, the entries of $\mathbf{A}$, a_{ij}'s, are related to $h(n_1, n_2)$ by

$$a_{ij} = 4h(L_1 - i, L_2 - j), \quad i = 1, 2, \ldots, L_1;\ j = 1, 2, \ldots, L_2.$$

Due to the quadrantal symmetry, we only need to consider the approximation of the desired frequency response on the quarter-plane $\Omega = \{(\omega_1, \omega_2) | 0 \le \omega_1 \le \pi; 0 \le \omega_2 \le \pi\}$ for the design of quadrantally symmetric filters. We discretize Ω by an $M_1 \times M_2$ rectangular frequency grid defined by $\Pi = \{(\omega_{1i}, \omega_{2j}) | i = 1, 2, \ldots, M_1;\ j = 1, 2, \ldots, M_2\}$, where

$0 \le \omega_{1i} \le \pi, 0 \le \omega_{2j} \le \pi, \omega_{1i} \ne \omega_{1k}$ for $i \ne k$, and $\omega_{2j} \ne \omega_{2l}$ for $j \ne l$. It should be pointed out that we don't assume uniform frequency grid Π. This implies that ω_{1i}, ω_{2j} can take arbitrary values satisfying $0 \le \omega_{1i}, \omega_{2j} \le \pi$. Let the desired magnitude response be denoted by $D(\omega_1, \omega_2)$, a real-valued function of ω_1 and ω_2. Then, the sum of the weighted square errors between the desired and designed magnitude responses is formulated as

$$E(\mathbf{A}) = \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} W(\omega_{1i}, \omega_{2j}) |G(\omega_{1i}, \omega_{2j}, \mathbf{A}) - D(\omega_{1i}, \omega_{2j})|^2, \tag{4}$$

where $W(\omega_1, \omega_2)$ is a nonnegative weighting function. Without loss of generality, we assume that $\max_{i,j}\{W(\omega_{1i}, \omega_{2j})\} = 1$.

In order to represent $E(\mathbf{A})$ in a matrix form, we define the following matrices and vectors:

$$\begin{aligned}
\mathbf{D} &= (d_{ij}),\ d_{ij} = D(\omega_{1i}, \omega_{2j});\ (i = 1, 2, \ldots, M_1;\ j = 1, 2, \ldots, M_2) \\
\mathbf{W} &= (w_{ij}),\ w_{ij} = W(\omega_{1i}, \omega_{2j});\ \ (i = 1, 2, \ldots, M_1;\ j = 1, 2, \ldots, M_2) \\
\mathbf{W}^{(\frac{1}{2})} &= (\sqrt{w_{ij}});\ (i = 1, 2, \ldots, M_1;\ j = 1, 2, \ldots, M_2) \\
\mathbf{p}_i &= \varphi(\omega_{1i}, N_1),\ i = 1, 2, \ldots, M_1;\ \mathbf{q}_j = \varphi(\omega_{2j}, N_2);\ j = 1, 2, \ldots, M_2 \\
\mathbf{P} &= [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_{M_1}];\ \mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_{M_2}].
\end{aligned}$$

Then, Eq. (4) can be rewritten as

$$E(\mathbf{A}) = \|(\mathbf{P}^T\mathbf{A}\mathbf{Q} - \mathbf{D}) \circ \mathbf{W}^{(\frac{1}{2})}\|_F^2, \tag{5}$$

where the notation $\|*\|_F$ and "$\circ$" denotes the Frobenius norm of a matrix and the Hadamard matrix product, respectively.

The WLS design of the 2-D quadrantally symmetric filter with coefficient matrix $\mathbf{A}$ is to minimize the sum of the weighted square errors, $E(\mathbf{A})$, with respect to $\mathbf{A}$, i.e.,

$$\min_{\mathbf{A}} E(\mathbf{A}). \tag{6}$$

Theorem 1 *The cost function $E(\mathbf{A})$ in (6) is a convex function. Further, it is strictly convex if and only if*

$$\|(\mathbf{P}^T\mathbf{X}\mathbf{Q}) \circ \mathbf{W}^{(\frac{1}{2})}\|_F \ne 0 \tag{7}$$

for all $\mathbf{X} \in \mathbb{R}^{L_1 \times L_2}$ with $\mathbf{X} \ne \mathbf{0}$, where $\mathbb{R}^{L_1 \times L_2}$ denotes the collective of all $L_1 \times L_2$ real matrices.

Theorem 1 was originally given in Zhao and Lai (2010) without proof. We now give a rigorous proof of the theorem as follows.

Proof In order to show the convexity of $E(\mathbf{A})$, it is only required to prove that the following inequality holds for any matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{L_1 \times L_2}$ and any scalar $\lambda \in (0, 1)$.

$$[\lambda E(\mathbf{A}) + (1 - \lambda)E(\mathbf{B})] - E(\lambda\mathbf{A} + (1 - \lambda)\mathbf{B}) \ge 0. \tag{8}$$

Using (4) and noting that $G(\omega_{1i}, \omega_{2j}, \mathbf{X}) = \mathbf{p}_i^T \mathbf{X} \mathbf{q}_j$ for any matrix $\mathbf{X} \in \mathbb{R}^{L_1 \times L_2}$, we have

$$
\begin{aligned}
&[\lambda E(\mathbf{A}) + (1-\lambda) E(\mathbf{B})] - E(\lambda \mathbf{A} + (1-\lambda)\mathbf{B}) \\
&= \lambda \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w_{ij} |\mathbf{p}_i^T \mathbf{A} \mathbf{q}_j - d_{ij}|^2 + (1-\lambda) \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w_{ij} |\mathbf{p}_i^T \mathbf{B} \mathbf{q}_j - d_{ij}|^2 \\
&\quad - \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w_{ij} |\lambda \mathbf{p}_i^T \mathbf{A} \mathbf{q}_j + (1-\lambda) \mathbf{p}_i^T \mathbf{B} \mathbf{q}_j - d_{ij}|^2 \\
&= \lambda \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w_{ij} |\mathbf{p}_i^T \mathbf{A} \mathbf{q}_j - d_{ij}|^2 + (1-\lambda) \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w_{ij} |\mathbf{p}_i^T \mathbf{B} \mathbf{q}_j - d_{ij}|^2 \\
&\quad - \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w_{ij} |\lambda (\mathbf{p}_i^T \mathbf{A} \mathbf{q}_j - d_{ij}) + (1-\lambda)(\mathbf{p}_i^T \mathbf{B} \mathbf{q}_j - d_{ij})|^2 \\
&= \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} \lambda w_{ij} |\mathbf{p}_i^T \mathbf{A} \mathbf{q}_j - d_{ij}|^2 + \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} (1-\lambda) w_{ij} |\mathbf{p}_i^T \mathbf{B} \mathbf{q}_j - d_{ij}|^2 \\
&\quad - \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w_{ij} [\lambda^2 (\mathbf{p}_i^T \mathbf{A} \mathbf{q}_j - d_{ij})^2 + (1-\lambda)^2 (\mathbf{p}_i^T \mathbf{B} \mathbf{q}_j - d_{ij})^2 \\
&\qquad + 2\lambda(1-\lambda)(\mathbf{p}_i^T \mathbf{A} \mathbf{q}_j - d_{ij})(\mathbf{p}_i^T \mathbf{B} \mathbf{q}_j - d_{ij})] \\
&= \lambda(1-\lambda) \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w_{ij} |\mathbf{p}_i^T (\mathbf{A} - \mathbf{B}) \mathbf{q}_j|^2 \\
&= \lambda(1-\lambda) \| [\mathbf{P}^T (\mathbf{A} - \mathbf{B}) \mathbf{Q}] \circ \mathbf{W}^{(\frac{1}{2})} \|_F^2 \geq 0.
\end{aligned} \tag{9}
$$

In addition, we can see from (9) that for any matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{L_1 \times L_2}$ with $\mathbf{A} \neq \mathbf{B}$, the inequality (8) becomes strict if and only if $\| (\mathbf{P}^T \mathbf{X} \mathbf{Q}) \circ \mathbf{W}^{(\frac{1}{2})} \|_F \neq 0$ for all nonzero matrix $\mathbf{X} \in \mathbb{R}^{L_1 \times L_2}$. This completes the proof. □

The strict convexity of the cost function $E(\mathbf{A})$ guarantees the uniqueness of the optimal solution to problem (6). Throughout the paper, we assume that $E(\mathbf{A})$ is strictly convex, i.e., the condition (7) holds. The optimality condition of the optimization problem (6) is $dE(\mathbf{A})/d\mathbf{A} = \mathbf{0}$, or, from Zhao and Lai (2010),

$$
\mathbf{P}[(\mathbf{P}^T \mathbf{A} \mathbf{Q}) \circ \mathbf{W}] \mathbf{Q}^T = \mathbf{P}(\mathbf{D} \circ \mathbf{W}) \mathbf{Q}^T. \tag{10}
$$

For the unweighted case, i.e., $w_{ij} = 1$ for all $i = 1, 2, \ldots, M_1$ and $j = 1, 2, \ldots, M_2$, the coefficient matrix $\mathbf{A}$ to be determined can be obtained analytically by

$$
\mathbf{A} = (\mathbf{P}\mathbf{P}^T)^{-1} \mathbf{P} \mathbf{D} \mathbf{Q}^T (\mathbf{Q}\mathbf{Q}^T)^{-1} \tag{11}
$$

provided that both $\mathbf{P}\mathbf{P}^T$ and $\mathbf{Q}\mathbf{Q}^T$ are nonsingular. It is not difficult to verify that the function spaces spanned by the base

$$
\{ 1/\sqrt{2}, \cos(\omega), \cos(2\omega), \ldots, \cos((N-1)\omega/2) \}, \quad \text{for odd } N,
$$

or

$$
\{ \cos(\omega/2), \cos(3\omega/2), \ldots, \cos((N-1)\omega/2) \}, \quad \text{for even } N,
$$

satisfy the Haar condition, which means that the matrices $\mathbf{P}$ and $\mathbf{Q}$ are of full ranks if $M_1 \geq L_1$ and $M_2 \geq L_2$. In practical design, the density of the frequency grid should be set to $M_1 \geq 2N_1$ and $M_2 \geq 2N_2$. Therefore the matrices $\mathbf{PP}^T$ and $\mathbf{QQ}^T$ are positive and nonsingular.

3 Efficient 2-D based algorithms

In this section, two efficient 2-D based algorithms are presented for the WLS design problem (6) of quadrantally symmetric 2-D FIR filters with arbitrary nonnegative weighting functions. Both algorithms are based on the optimality condition (10) and solve for the coefficient matrix $\mathbf{A}$ by iterative methods. The first algorithm is a basic one, some preliminary results of which were presented in Zhao and Lai (2010), and the second algorithm is an improved version of the basic algorithm.

3.1 Matrix iterative algorithm I

By replacing the matrix $\mathbf{W}$ in the left hand side of (10) by $(\mathbf{U} + \mathbf{W} - \mathbf{U})$, where $\mathbf{U}$ is a unit-entry matrix with the same size as $\mathbf{W}$, we get

$$\mathbf{P}[(\mathbf{P}^T\mathbf{AQ}) \circ (\mathbf{U} + \mathbf{W} - \mathbf{U})]\mathbf{Q}^T = \mathbf{P}(\mathbf{D} \circ \mathbf{W})\mathbf{Q}^T, \tag{12}$$

or,

$$(\mathbf{PP}^T)\mathbf{A}(\mathbf{QQ}^T) + \mathbf{P}[(\mathbf{P}^T\mathbf{AQ}) \circ (\mathbf{W} - \mathbf{U})]\mathbf{Q}^T = \mathbf{P}(\mathbf{D} \circ \mathbf{W})\mathbf{Q}^T,$$

which motivates the following iterative matrix equation for the solution matrix of the optimality condition (10).

$$\mathbf{A}_{n+1} = (\mathbf{PP}^T)^{-1}\mathbf{P}[\mathbf{D} \circ \mathbf{W} + (\mathbf{P}^T\mathbf{A}_n\mathbf{Q}) \circ (\mathbf{U} - \mathbf{W})]\mathbf{Q}^T(\mathbf{QQ}^T)^{-1}. \tag{13}$$

In order to establish the convergence of the above iterative equation, we introduce two bounded linear operators T_1 and T_2 on the matrix space $\mathbb{R}^{L_1 \times L_2}$, defined respectively by

$$\begin{aligned} T_1(\mathbf{X}) &= (\mathbf{PP}^T)^{-1}\mathbf{P}[(\mathbf{P}^T\mathbf{XQ}) \circ \mathbf{W}]\mathbf{Q}^T(\mathbf{QQ}^T)^{-1}, \\ T_2(\mathbf{X}) &= (\mathbf{PP}^T)^{-1}\mathbf{P}[(\mathbf{P}^T\mathbf{XQ}) \circ (\mathbf{U} - \mathbf{W})]\mathbf{Q}^T(\mathbf{QQ}^T)^{-1}, \end{aligned}$$

satisfying the relation $T_2 = T_0 - T_1$, where T_0 represents the identity operator on $\mathbb{R}^{L_1 \times L_2}$. Then, we can rewrite the iterative equation (13) in a compact form as

$$\mathbf{A}_{n+1} = \hat{\mathbf{D}} + T_2(\mathbf{A}_n), \tag{14}$$

where

$$\hat{\mathbf{D}} = (\mathbf{PP}^T)^{-1}\mathbf{P}(\mathbf{D} \circ \mathbf{W})\mathbf{Q}^T(\mathbf{QQ}^T)^{-1}. \tag{15}$$

It has been verified in Zhao and Lai (2010) that both T_1 and T_2 are compact self-adjoint operators, and thus all of their eigenvalues are real numbers. Moreover, because the operators T_1 and T_2 are defined on the finite-dimensional space $\mathbb{R}^{L_1 \times L_2}$, the numbers of eigenvalues of T_1 and T_2 are both finite, and their spectra, denoted by $\sigma(T_1)$ and $\sigma(T_2)$, are sets of all eigenvalues of T_1 and T_2 respectively. Then, a sufficient condition for the convergence of iterative equation (14) can be described by

$$r_\sigma(T_2) = \max\{|\lambda|, \lambda \in \sigma(T_2)\} < 1, \tag{16}$$

where $r_\sigma(*)$ represents the spectrum radius of an operator.

Further, we define the following two quantities for a bounded self-adjoint linear operator T defined on a Hilbert space $\mathbf{H}$:

$$m(T) = \inf_{\langle \mathbf{x},\mathbf{x}\rangle=1} \langle T(\mathbf{x}), \mathbf{x}\rangle, \quad \mathbf{x} \in \mathbf{H}, \tag{17}$$

$$M(T) = \sup_{\langle \mathbf{x},\mathbf{x}\rangle=1} \langle T(\mathbf{x}), \mathbf{x}\rangle, \quad \mathbf{x} \in \mathbf{H}, \tag{18}$$

where the notation $\langle *, * \rangle$ represents an inner product defined on the Hilbert space $\mathbf{H}$. Then, from the basic linear operator theories (see,e.g., Gohberg et al. 2003), $m(T_i)$ and $M(T_i)$ are the smallest and largest eigenvalues of T_i $(i = 1, 2)$.

The following lemma is a key lemma for the convergence of the iterative equation (14).

Lemma 1 $0 < m(T_1) \le M(T_1) \le 1$, *and* $0 \le m(T_2) \le M(T_2) < 1$.

Proof Firstly, we define an inner product on the matrix space $\mathbb{R}^{L_1 \times L_2}$ by

$$\langle \mathbf{X}, \mathbf{Y}\rangle_1 = \mathrm{tr}\{(\mathbf{P}^T\mathbf{X}\mathbf{Q})^T(\mathbf{P}^T\mathbf{Y}\mathbf{Q})\}, \tag{19}$$

where tr$\{*\}$ represents the trace of a matrix. Noting that matrices $\mathbf{P}$ and $\mathbf{Q}$ are of full ranks, it is not difficult to verify that (19) is indeed an inner product of $\mathbf{X}$ and $\mathbf{Y}$.

Then, for any nonzero matrix $\mathbf{X} \neq \mathbf{0} \in \mathbb{R}^{L_1 \times L_2}$, we have

$$\begin{aligned}
\langle T_1(\mathbf{X}), \mathbf{X}\rangle_1 &= \mathrm{tr}\{[\mathbf{P}^T T_1(\mathbf{X})\mathbf{Q}]^T(\mathbf{P}^T\mathbf{X}\mathbf{Q})\} \\
&= \mathrm{tr}\{(\mathbf{P}^T\mathbf{X}\mathbf{Q})^T\mathbf{P}^T T_1(\mathbf{X})\mathbf{Q}\} \\
&= \mathrm{tr}\{(\mathbf{P}^T\mathbf{X}\mathbf{Q})^T\mathbf{P}^T(\mathbf{P}\mathbf{P}^T)^{-1}\mathbf{P}[(\mathbf{P}^T\mathbf{X}\mathbf{Q}) \circ \mathbf{W}]\mathbf{Q}^T(\mathbf{Q}\mathbf{Q}^T)^{-1}\mathbf{Q}\} \\
&= \mathrm{tr}\{\mathbf{Q}^T(\mathbf{Q}\mathbf{Q}^T)^{-1}\mathbf{Q}(\mathbf{Q}^T\mathbf{X}\mathbf{P})\mathbf{P}^T(\mathbf{P}\mathbf{P}^T)^{-1}\mathbf{P}[(\mathbf{P}^T\mathbf{X}\mathbf{Q}) \circ \mathbf{W}]\} \\
&= \mathrm{tr}\{(\mathbf{P}^T\mathbf{X}\mathbf{Q})^T[(\mathbf{P}^T\mathbf{X}\mathbf{Q}) \circ \mathbf{W}]\} \\
&= \|(\mathbf{P}^T\mathbf{X}\mathbf{Q}) \circ \mathbf{W}^{(\frac{1}{2})}\|_F^2
\end{aligned}$$

It follows from definition (17) that $m(T_1) \ge 0$. If $m(T_1) = 0$, zero would be an eigenvalue of T_1, i.e., $0 \in \sigma(T_1)$. Then there is a $\mathbf{X} \neq 0$ such that $T_1(\mathbf{X}) = 0 \cdot \mathbf{X} = 0$ and thus $\langle T_1(\mathbf{X}), \mathbf{X}\rangle_1 = 0$, which is contradictory to condition (7). Consequently, $m(T_1) > 0$.

Similarly, it can be obtained that

$$\langle T_2(\mathbf{X}), \mathbf{X}\rangle_1 = \mathrm{tr}\{(\mathbf{P}^T\mathbf{X}\mathbf{Q})^T[(\mathbf{P}^T\mathbf{X}\mathbf{Q}) \circ (\mathbf{U} - \mathbf{W})]\} \ge 0, \tag{20}$$

noting that all entries of matrix $\mathbf{U} - \mathbf{W}$ are nonnegative. Equation (20) means $m(T_2) \ge 0$.

From $T_1 = T_0 - T_2$, we have

$$\begin{aligned}
M(T_1) &= \sup_{\langle \mathbf{X},\mathbf{X}\rangle_1=1} \langle T_1(\mathbf{X}), \mathbf{X}\rangle_1 \\
&= \sup_{\langle \mathbf{X},\mathbf{X}\rangle_1=1} \langle \mathbf{X} - T_2(\mathbf{X}), \mathbf{X}\rangle_1 \\
&= 1 - \inf_{\langle \mathbf{X},\mathbf{X}\rangle_1=1} \langle T_2(\mathbf{X}), \mathbf{X}\rangle_1 \\
&= 1 - m(T_2) \le 1.
\end{aligned}$$

In the same way, we can obtain

$$M(T_2) = 1 - m(T_1) < 1,$$

which completes the proof. □

Lemma 1 shows that T_1 is a positive operator with $r_\sigma(T_1) \leq 1$, and T_2 is a nonnegative operator with $r_\sigma(T_2) < 1$. It follows immediately that:

Theorem 2 *Given an initial matrix* $\mathbf{A}_0$, *let* $\{\mathbf{A}_n\}$ *be the matrix sequence generated by the iterative equation (14), then the sequence* $\{\mathbf{A}_n\}$ *converges.*

One of the good choices for the initial matrix $\mathbf{A}_0$ is $\hat{\mathbf{D}}$ given in (15). Then the matrix iterative algorithm I can be described as follows:

Matrix Iterative Algorithm I
Step 1: Given the frequency grid Π, the filter lengths N_1, N_2, and an error tolerance $\varepsilon > 0$, construct the desired magnitude matrix $\mathbf{D}$ and the weighted matrix $\mathbf{W}$. Compute the matrices $\mathbf{P}$ and $\mathbf{Q}$. Let $\hat{\mathbf{P}} = (\mathbf{P}\mathbf{P}^T)^{-1}\mathbf{P}$, $\hat{\mathbf{Q}} = (\mathbf{Q}\mathbf{Q}^T)^{-1}\mathbf{Q}$. Compute the matrix $\hat{\mathbf{D}}$ by (15). Let $\mathbf{A}_0 = \hat{\mathbf{D}}$, $\hat{\mathbf{W}} = \mathbf{U} - \mathbf{W}$ and $n = 0$.
Step 2: Let $\mathbf{A}_{n+1} = \hat{\mathbf{D}} + \hat{\mathbf{P}}[(\mathbf{P}^T\mathbf{A}_n\mathbf{Q}) \circ \hat{\mathbf{W}}]\hat{\mathbf{Q}}$. If $\| \mathbf{A}_{n+1} - \mathbf{A}_n \|_F < \varepsilon$, terminate the algorithm. Otherwise, let $n = n + 1$ and repeat this step.

3.2 Matrix iterative algorithm II

Compared with existing algorithms, the matrix iterative algorithm I achieves great improvement in design time (Zhao and Lai 2010). However, from a large number of design examples we observe that the iteration number of the matrix iterative algorithm I fast increases with the filter size and the transition band width for certain types of 2-D filters, e.g., elliptic filters and fan filters.

From the iterative equation (14), it can be seen that the convergence rate of the matrix iterative algorithm I is mainly determined by the spectral radius of T_2, i.e., $r_\sigma(T_2)$. The smaller the $r_\sigma(T_2)$ is, the faster the matrix iterative algorithm I converges. This motivates us to seek other iterative operators possessing smaller spectral radius than $r_\sigma(T_2)$.

To this end, we introduce a positive real parameter $\mu > 0$ into the matrix equation (12) as

$$\mathbf{P}[(\mathbf{P}^T\mathbf{A}\mathbf{Q}) \circ (\mathbf{W} - \frac{1}{\mu}\mathbf{U} + \frac{1}{\mu}\mathbf{U})]\mathbf{Q}^T = \mathbf{P}(\mathbf{D} \circ \mathbf{W})\mathbf{Q}^T.$$

Then we have

$$\mathbf{A} = (\mathbf{P}\mathbf{P}^T)^{-1}\mathbf{P}[\mu\mathbf{D} \circ \mathbf{W} + (\mathbf{P}^T\mathbf{A}\mathbf{Q}) \circ (\mathbf{U} - \mu\mathbf{W})]\mathbf{Q}^T(\mathbf{Q}\mathbf{Q}^T)^{-1},$$

which suggests the following iterative equation for the solution of Eq. (10)

$$\mathbf{A}_{n+1} = (\mathbf{P}\mathbf{P}^T)^{-1}\mathbf{P}[\mu\mathbf{D} \circ \mathbf{W} + (\mathbf{P}^T\mathbf{A}_n\mathbf{Q}) \circ (\mathbf{U} - \mu\mathbf{W})]\mathbf{Q}^T(\mathbf{Q}\mathbf{Q}^T)^{-1}. \tag{21}$$

Define a linear operator T_3 on the matrix space $\mathbb{R}^{L_1 \times L_2}$ by

$$T_3(\mathbf{X}) = (\mathbf{P}\mathbf{P}^T)^{-1}\mathbf{P}[(\mathbf{P}^T\mathbf{X}\mathbf{Q}) \circ (\mathbf{U} - \mu\mathbf{W})]\mathbf{Q}^T(\mathbf{Q}\mathbf{Q}^T)^{-1}.$$

It is not difficult to verify that the operators T_3 and T_1 satisfy $T_3 = T_0 - \mu T_1$. Hence, T_3 is a bounded, compact and self-adjoint operator. In addition, the iterative equation (21) can be rewritten as

$$\mathbf{A}_{n+1} = \mu\hat{\mathbf{D}} + T_3(\mathbf{A}_n), \tag{22}$$

with the same $\hat{\mathbf{D}}$ as in (15).

Clearly, for the special case of $\mu = 1$, we have $T_3 = T_2$. Now, we want to seek an optimal value μ^* of the parameter $\mu > 0$, at which the spectral radius of T_3, $r_\sigma(T_3)$, achieves its

minimal value. Obviously, the minimal value of $r_\sigma(T_3)$, i.e. the value of $r_\sigma(T_3)$ at μ^* is definitely not larger than $r_\sigma(T_2)$, which is equal to the value of $r_\sigma(T_3)$ at $\mu = 1$. In practice, the minimal value of $r_\sigma(T_3)$ is always less than $r_\sigma(T_2)$. Then, once the optimal μ is found and applied to (22), the convergence rate of the iterative equation (22) will be undoubtedly faster than the matrix iterative algorithm I.

Theorem 3 *The spectral radius of operator T_3 attains its minimal value at*

$$\mu = \frac{2}{m(T_1) + M(T_1)}.$$

Proof Firstly, according to definition (17), we have

$$\begin{aligned} m(T_3) &= \inf_{\langle \mathbf{X}, \mathbf{X} \rangle_1 = 1} \langle T_3(\mathbf{X}), \mathbf{X} \rangle_1 \\ &= \inf_{\langle \mathbf{X}, \mathbf{X} \rangle_1 = 1} \langle \mathbf{X} - \mu T_1(\mathbf{X}), \mathbf{X} \rangle_1 \\ &= 1 - \mu \sup_{\langle \mathbf{X}, \mathbf{X} \rangle_1 = 1} \langle T_1(\mathbf{X}), \mathbf{X} \rangle_1 \\ &= 1 - \mu M(T_1) \end{aligned}$$

In a similar way, we can obtain

$$M(T_3) = 1 - \mu m(T_1).$$

The above results indicate that $1 - \mu M(T_1)$ and $1 - \mu m(T_1)$ are the smallest and largest eigenvalues of T_3, respectively. Consequently, the spectral radius of operator T_3 is a function of the parameter $\mu > 0$, as given by

$$r_\sigma(T_3, \mu) = \max\{|1 - \mu M(T_1)|, |1 - \mu m(T_1)|\}, \quad \mu > 0.$$

Noting that $0 < m(T_1) \leq M(T_1) \leq 1$, it can be easily obtained that

$$\begin{aligned} r_\sigma(T_3, \mu) &= \begin{cases} \max\{1 - \mu M(T_1), 1 - \mu m(T_1)\} & 0 < \mu \leq \frac{1}{M(T_1)}, \\ \max\{\mu M(T_1) - 1, 1 - \mu m(T_1)\} & \frac{1}{M(T_1)} < \mu \leq \frac{1}{m(T_1)}, \\ \max\{\mu M(T_1) - 1, \mu m(T_1) - 1\} & \frac{1}{m(T_1)} < \mu \end{cases} \\ &= \begin{cases} 1 - \mu m(T_1), & 0 < \mu \leq \frac{1}{M(T_1)}, \\ 1 - \mu m(T_1), & \frac{1}{M(T_1)} < \mu \leq \frac{2}{m(T_1)+M(T_1)}, \\ \mu M(T_1) - 1, & \frac{2}{m(T_1)+M(T_1)} \leq \mu \leq \frac{1}{m(T_1)}, \\ \mu M(T_1) - 1, & \frac{1}{m(T_1)} < \mu \end{cases} \\ &= \begin{cases} 1 - \mu m(T_1), & 0 < \mu \leq \frac{2}{m(T_1)+M(T_1)}, \\ \mu M(T_1) - 1, & \frac{2}{m(T_1)+M(T_1)} \leq \mu. \end{cases} \end{aligned}$$

Obviously, $1 - \mu m(T_1)$ and $\mu M(T_1) - 1$ are respectively monotonically decreasing and increasing functions of μ, which implies that $r_\sigma(T_3, \mu)$ attains its minimal value at

$$\mu = \frac{2}{m(T_1) + M(T_1)}.$$

The proof is complete. □

In order to compute the optimal μ that makes $r_\sigma(T_3)$ attain its minimal value, we need to compute $m(T_1)$ and $M(T_1)$ first. From the proof of Lemma 1 we know that $m(T_1) = 1 - M(T_2)$. Then, if we have obtained the largest eigenvalues of T_1 and T_2, i.e., $M(T_1)$ and $M(T_2)$, the optimal μ can be computed by

$$\mu = \frac{2}{1 - M(T_2) + M(T_1)}.$$

To this end, we present in the following a procedure to iteratively compute the largest eigenvalue of a nonnegative, compact and self-adjoint operator T defined on a finite dimensional Hilbert space $\mathbf{H}$.

Procedure 1
Step 1: Given an initial point x_0 in the space $\mathbf{H}$, set $k = 1$.
Step 2: Let $y_k = T(x_{k-1})$.
Step 3: Let

$$x_k = \frac{y_k}{\sqrt{\langle y_k, y_k \rangle}}.$$

Step 4: Let $\gamma_k = \langle x_k, T(x_k) \rangle$, $k = k + 1$, and return Step 2.

It can be proved that the limit of the sequence $\{\gamma_k\}$ generated by the above procedure is just the largest eigenvalue of T, i.e., $\lim_{k\to\infty} \gamma_k = M(T)$. To this end, we need the following theorem (Gohberg et al. 2003).

Theorem 4 *Suppose T is a compact self-adjoint operator on Hilbert space $\mathbf{H}$. There exist an orthonormal system $u_1, u_2, \ldots$ of eigenvectors of T with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots$ such that for all $x \in H$,*

$$T(x) = \sum_k \lambda_k \langle x, u_k \rangle u_k.$$

Because the Hilbert space $\mathbf{H}$ is finite-dimensional and T is nonnegative, T has a finite number of eigenvalues and all its eigenvalues are nonnegative. We assume T has s eigenvalues described by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s \geq 0$ with corresponding eigenvectors $u_1, u_2, \ldots, u_s$ satisfying Theorem 4. Then, y_1 obtained in Step 2 can be written as

$$y_1 = \sum_{i=1}^{s} \lambda_i \alpha_i u_i,$$

where $\alpha_i = \langle x_0, u_i \rangle$, $i = 1, 2, \ldots, s$, according to Theorem 4. Consequently, x_1 in Step 3 is given by

$$x_1 = \frac{y_1}{\sqrt{\langle y_1, y_1 \rangle}} = \frac{\sum_{i=1}^{s} \lambda_i \alpha_i u_i}{\sqrt{\langle \sum_{i=1}^{s} \lambda_i \alpha_i u_i, \sum_{i=1}^{s} \lambda_i \alpha_i u_i \rangle}}.$$

Considering the orthonormality of the eigenvectors $u_1, u_2, \ldots, u_s$, i.e.,

$$\langle u_i, u_j \rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases},$$

we have

$$x_1 = \frac{\sum_{i=1}^{s} \lambda_i \alpha_i u_i}{\sqrt{\sum_{i=1}^{s} \lambda_i^2 \alpha_i^2}}.$$

In view of $T(u_i) = \lambda_i u_i$ for $i = 1, 2, \ldots, s$, we further have

$$x_k = \frac{\sum_{i=1}^{s} \lambda_i^k \alpha_i u_i}{\sqrt{\sum_{i=1}^{s} \lambda_i^{2k} \alpha_i^2}}.$$

Therefore, γ_k in Step 4 is given by

$$\gamma_k = \langle x_k, T(x_k) \rangle = \frac{\langle \sum_{i=1}^{s} \lambda_i^k \alpha_i u_i, \sum_{i=1}^{s} \lambda_i^{k+1} \alpha_i u_i \rangle}{\sum_{i=1}^{s} \lambda_i^{2k} \alpha_i^2} = \frac{\sum_{i=1}^{s} \lambda_i^{2k+1} \alpha_i^2}{\sum_{i=1}^{s} \lambda_i^{2k} \alpha_i^2}.$$

Dividing simultaneously by $\lambda_1^{2k} \alpha_1^2$ the numerator and denominator of the right hand side of the above equation, we have

$$\gamma_k = \frac{\lambda_1 + \sum_{i=2}^{s} \lambda_i (\frac{\lambda_i}{\lambda_1})^{2k} (\frac{\alpha_i}{\alpha_1})^2}{1 + \sum_{i=2}^{s} (\frac{\lambda_i}{\lambda_1})^{2k} (\frac{\alpha_i}{\alpha_1})^2}.$$

It then follows from $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s \geq 0$ that $\lim_{k \to \infty} \gamma_k = \lambda_1 = M(T)$.

Combining the iterative equation (22) with Procedure 1, we finally obtain the following improved iterative algorithm:

Matrix Iterative Algorithm II

Step 1: Given the frequency grid Π, the filter lengths N_1, N_2, and two error tolerances $\varepsilon_1 > 0, \varepsilon_2 > 0$, construct the desired magnitude matrix $\mathbf{D}$ and the weighted matrix $\mathbf{W}$. Compute matrices $\mathbf{P}$ and $\mathbf{Q}$. Let $\hat{\mathbf{P}} = (\mathbf{P}\mathbf{P}^T)^{-1}\mathbf{P}$, $\hat{\mathbf{Q}} = (\mathbf{Q}\mathbf{Q}^T)^{-1}\mathbf{Q}$. Compute matrix $\hat{\mathbf{D}}$ by (15). Let $\mathbf{A}_0 = \hat{\mathbf{D}}$ and $n = 0$.

Step 2: Given an integer $K > 0$, let $\mathbf{X}_0 = \hat{\mathbf{D}}$, $\tilde{\mathbf{X}}_0 = \hat{\mathbf{D}}$, $\tilde{\mathbf{W}} = \mathbf{U} - \mathbf{W}$ and $k = 1$. Compute

$$\mathbf{Y}_k = T_1(\mathbf{X}_k) = \hat{\mathbf{P}}[(\mathbf{P}^T \mathbf{X}_{k-1} \mathbf{Q}) \circ \mathbf{W}]\hat{\mathbf{Q}}, \quad \mathbf{X}_k = \frac{\mathbf{Y}_k}{\sqrt{\langle \mathbf{Y}_k, \mathbf{Y}_k \rangle_1}},$$

$$\tilde{\mathbf{Y}}_k = T_2(\tilde{\mathbf{X}}_k) = \hat{\mathbf{P}}[(\mathbf{P}^T \tilde{\mathbf{X}}_{k-1} \mathbf{Q}) \circ \tilde{\mathbf{W}}]\hat{\mathbf{Q}}, \quad \tilde{\mathbf{X}}_k = \frac{\tilde{\mathbf{Y}}_k}{\sqrt{\langle \tilde{\mathbf{Y}}_k, \tilde{\mathbf{Y}}_k \rangle_1}}.$$

Step 3: If $\| \mathbf{X}_k - \mathbf{X}_{k-1} \|_F + \| \tilde{\mathbf{X}}_k - \tilde{\mathbf{X}}_{k-1} \|_F < \varepsilon_1$ or $k > K$, let

$$\mu = \frac{2}{1 - \langle \tilde{\mathbf{X}}_k, T_2(\tilde{\mathbf{X}}_k) \rangle_2 + \langle \mathbf{X}_k, T_1(\mathbf{X}_k) \rangle_1}, \quad \hat{\mathbf{W}} = \mathbf{U} - \mu \mathbf{W},$$

and go to Step 4. Otherwise let $k = k + 1$ and repeat Step 3.

Step 4: Let $\mathbf{A}_{n+1} = \mu \hat{\mathbf{D}} + \hat{\mathbf{P}}[(\mathbf{P}^T \mathbf{A}_n \mathbf{Q}) \circ \hat{\mathbf{W}}]\hat{\mathbf{Q}}$. If $\| \mathbf{A}_{n+1} - \mathbf{A}_n \|_F < \varepsilon_2$, terminate the algorithm. Otherwise, let $n = n + 1$ and repeat this step.

In the above algorithm, the optimal μ described by Theorem 3 is obtained through Steps 2 and 3 (correspond to Procedure 1). It is easy to verify that the amount of computation required in Step 2 of each iteration is about 2 times of that required in Step 4. If the number of iterations required in Step 2 is too large, the efficiency of the entire algorithm will decrease. Thus, an integer K, which is generally set as 15, is used to limit the largest number of iterations in Step 2.

Table 1 Comparison of Algorithms I and II with the CG algorithm for filters of different sizes in Example 1

Filter size	Algorithm I			Algorithm II			CG Algorithm		
	N_{it}	T_{cpu} (s)	R_I	N_{it}	T_{cpu} (s)	R_{II}	N_{it}	T_{cpu} (s)	R_c
19×21	39	0.036	1.051×10^{-7}	24 (13)	0.045	1.774×10^{-8}	12	0.27	4.278×10^{-10}
41×41	118	0.53	6.571×10^{-6}	70 (9)	0.33	1.476×10^{-6}	20	1.58	1.066×10^{-8}
55×56	295	2.1	9.408×10^{-5}	179 (9)	1.5	2.550×10^{-5}	36	3.6	2.285×10^{-7}
71×81	1,225	18.45	4.081×10^{-3}	883 (6)	12.8	1.317×10^{-3}	75	18.48	1.857×10^{-6}

4 Simulations and comparisons

In this section, we show properties of the two proposed algorithms and compare them with existing algorithms through two design examples. In all algorithms, we use the same frequency grid Π with $M_1 = 4N_1$ and $M_2 = 4N_2$. The matrix iterative algorithm I has been shown in Zhao and Lai (2010) to be very fast for designing rectangular filters. Thus in this paper, we consider 2-D filters of other types, e.g., elliptic and fan filters.

Example 1 Design of elliptic FIR filters with the same specifications as in Zhao and Lai (2010). The passband Ω_p and stopband Ω_s are described as follows:

$$\Omega_p = \left\{ (\omega_1, \omega_2) | \frac{\omega_1^2}{(0.4\pi)^2} + \frac{\omega_2^2}{(0.5\pi)^2} \le 1 \right\},$$

$$\Omega_s = \left\{ (\omega_1, \omega_2) | \frac{\omega_1^2}{(0.45\pi)^2} + \frac{\omega_2^2}{(0.55\pi)^2} \ge 1, \omega_1 \le 1, \omega_2 \le 1 \right\}.$$

The weights on Ω_p, Ω_s, and the transition band are 1, 0.25, and 0, respectively.

We conduct the design with the matrix iterative algorithm I, the matrix iterative algorithm II, and the conjugate gradient (CG) algorithm in Aravena and Gu (1996). In Algorithm I and the CG algorithm, the error tolerance is taken to be $\varepsilon = 10^{-5}$. In Algorithm II, the two error tolerances are taken to be $\varepsilon_1 = 10^{-3}$ and $\varepsilon_2 = 10^{-5}$, respectively. Filters with different sizes have been designed, and Table 1 lists the numbers of iterations N_{it}, the design time T_{cpu} and the relative errors R_I, R_{II} and R_c defined respectively by

$$R_I = \frac{E_I - E}{E}, \quad R_{II} = \frac{E_{II} - E}{E}, \quad R_c = \frac{E_c - E}{E}, \tag{23}$$

where E_I, E_{II}, E_c and E represent the cost function values obtained by Algorithm I, Algorithm II, the CG algorithm and the 1-D WLS method in Gislason et al. (1993), respectively. For Algorithm II, the iteration numbers needed to obtain the optimal μ in Step 2 are in the bracket just behind the iteration numbers required in Step 4. Table 2 lists the maximum passband errors (E_p) and minimum stopband attenuations (A_s) obtained by the three algorithms.

It can be found from Table 1 that Algorithms I and II need less design time although more iterations than the CG algorithm, and that Algorithm II consumes less design time, requires fewer iterations, and obtains more accurate solutions than Algorithm I. From Table 2, we find that the resulting filters obtained by the three algorithms have almost the same maximum passband errors and minimum stopband attenuations. Figure 1 shows the magnitude response of filter with size 19×21.

Table 2 Maximum passband errors and minimum stopband attenuations obtained by Algorithms I and II and the CG algorithm for filters of different sizes in Example 1

Filter size	Algorithm I		Algorithm II		CG Algorithm	
	E_p	A_s (dB)	E_p	A_s (dB)	E_p	A_s (dB)
19×21	2.039×10^{-1}	8.498	2.039×10^{-1}	8.497	2.039×10^{-1}	8.497
41×41	8.801×10^{-2}	15.093	8.801×10^{-2}	15.093	8.801×10^{-2}	15.092
55×56	4.965×10^{-2}	19.554	4.965×10^{-2}	19.556	4.964×10^{-2}	19.555
71×81	2.885×10^{-2}	24.837	2.884×10^{-2}	24.836	2.883×10^{-2}	24.835

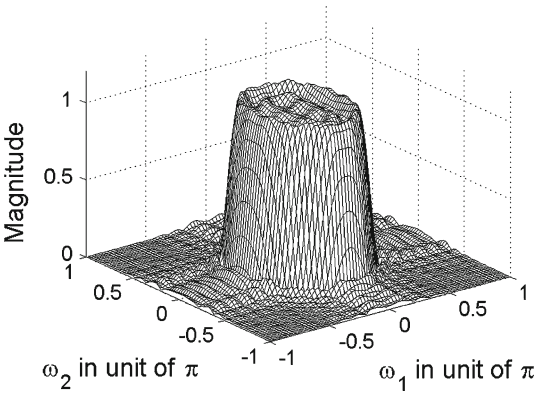

Fig. 1 Magnitude response of the 19×21 elliptic filter in Example 1

It should be pointed out that only uniform spaced frequency grid can be used in the CG algorithm, while Algorithms I and II do not have this limitation. This makes the presented algorithms more flexible. For example, the frequency grid may always include the frequencies at the passband and stopband edges as the grid points in our algorithms, no matter what the grid density is taken as. For a uniform grid, however, the band edges may be excluded from the grid if the grid density is not properly chosen. This may result in large error at the band edges.

Example 2 Design of $N \times N$ fan FIR filters with passband Ω_p and stopband Ω_s in the first quadrant

$$\Omega_p = \{(\omega_1, \omega_2) \mid 0 \le \omega_1 \le \omega_2 \tan\beta, 0 \le \omega_2 \le \pi\},$$
$$\Omega_s = \{(\omega_1, \omega_2) \mid \varpi \le \omega_1 \le \pi, 0 \le \omega_2 \le \tfrac{(\omega_1 - \varpi)}{\tan\beta}\},$$

as shown in Fig. 2, where $\beta = \pi/6$ and ϖ is a parameter used to control the width of the transition band. The weights on the passband, stopband and transition band are taken to be 1, 0.5 and 0, respectively.

In this example, we use the matrix iterative algorithms I and II to design the fan filters. At first, we design filters of different sizes with the same transition band width ($\varpi = 0.06\pi$). The design results are listed in Table 3. From the table, we see that both algorithms can design the filters very efficiently. Although the iteration numbers required by both algorithms increase with filter size, Algorithms II requires much fewer iterations than Algorithm I because the iterative operator of Algorithm II has smaller spectral radius than that of Algorithm I. For

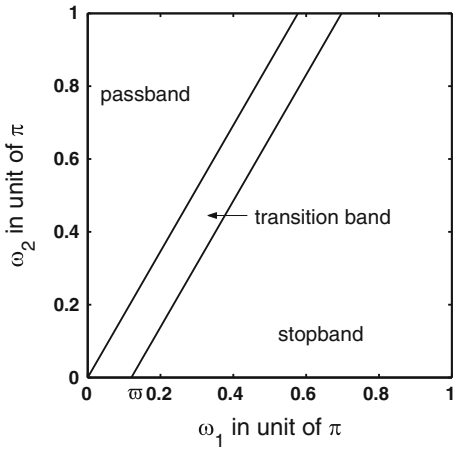

Fig. 2 Frequency bands of the fan filter in Example 2

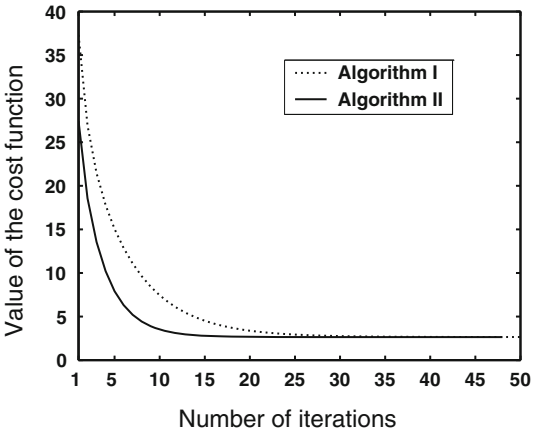

Fig. 3 Convergence of Algorithms I and II for the 41×41 filter

Table 3 Comparison between Algorithms I and II for filters of different sizes with $\varpi = 0.06\pi$ in Example 2

Filter size	Algorithm I			Algorithm II		
	N_{it}	T_{cpu} (s)	R_I	N_{it}	T_{cpu} (s)	R_{II}
21×21	27	0.014	3.494×10^{-8}	17(7)	0.019	6.003×10^{-9}
41×41	83	0.20	2.068×10^{-6}	48(7)	0.17	4.276×10^{-7}
61×61	312	2.8	1.284×10^{-4}	200(6)	1.9	4.587×10^{-5}
71×71	622	7.8	9.826×10^{-4}	402(5)	5.4	3.371×10^{-4}

low-order filters, the two algorithms have similar efficiency. But for high-order filters, Algorithm II is more efficient than Algorithm I. Figure 3 illustrates the convergence rates of both algorithms for the 41×41 filter. Evidently, Algorithm II converges faster than Algorithms I.

Then, we design filters of the same size 41×41 but with increasing transition band widths. Table 4 summaries the design results. It is interesting that the influence of the transition band width on the convergence is similar to that of the filter order. That is, the increase of the

Table 4 Comparison of Algorithms I and II for the 41×41 filters with different transition band widths in Example 2

Values of ϖ	Algorithm I			Algorithm II		
	N_{it}	T_{cpu} (s)	R_I	N_{it}	T_{cpu} (s)	R_{II}
$\varpi = 0.04\pi$	37	0.16	1.526×10^{-7}	22(5)	0.11	4.428×10^{-8}
$\varpi = 0.08\pi$	203	0.63	2.536×10^{-5}	115(11)	0.41	6.391×10^{-6}
$\varpi = 0.1\pi$	583	1.4	4.216×10^{-4}	375(7)	0.78	1.408×10^{-4}
$\varpi = 0.12\pi$	1520	4.7	5.088×10^{-3}	963(9)	3.0	1.556×10^{-3}

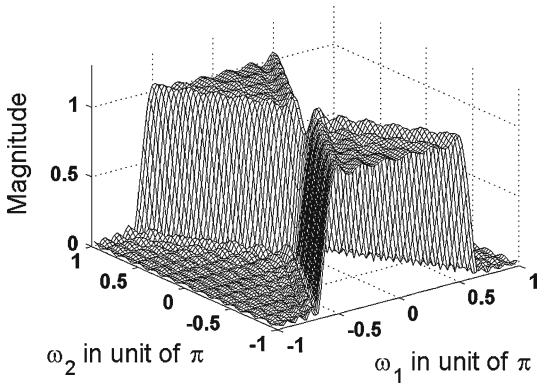

Fig. 4 Magnitude response of the 41×41 fan filter with $\varpi = 0.06\pi$ in Example 2

transition band width results in corresponding increase of the iteration number. A reason for this observation is, the increase of the transition band leads to increase of the spectral radius of the iterative operators of Algorithms I and II, and thus leads to decrease of the convergence rates of the algorithms.

When the transition band is narrow, the efficiencies of the two algorithms are close to each other. But with the increase of the transition band width, the efficiency superiority of Algorithm II over Algorithm II becomes obvious. Figure 4 shows the magnitude response of the 41×41 filter with $\varpi = 0.06\pi$.

As a comment, an empirical value of μ instead of its optimal value can also work for Algorithm II. If we have a good estimation of the optimal μ, we needn't compute its optimal value, and then the total computation time could be reduced further. To this end, we examine the dependence of the optimal μ on the filter size and transition band width. In order to obtain a value of μ closer to its optimal value, the error tolerance ε_1 in Algorithm II is now set to 10^{-4}, and the iteration number limit set by the integer K in Step 2 is removed. We first fix the transition band width by $\varpi = 0.06\pi$. The change of the optimal μ with the filter size N is shown in Fig. 5. Then we fix the filter size by 41×41, the values of the optimal μ versus ϖ ranging from 0.01π to 0.12π in step of 0.01π are shown in Fig. 6. We can see from the two figures that the optimal value of μ increases with the filter size and transition band width.

In order to show the dependence of the iteration number on the parameter μ, we show in Fig. 7 the change of the iteration number required by Algorithm II for the filter with $N = 41, 45$ and $\varpi = 0.05\pi, 0.06\pi$ when the optimal μ is replaced by some preset μ values ranging from 1 to 1.9. With the increase of μ, the iteration number decreases gradually at first and then increases fast when μ is close to 2.

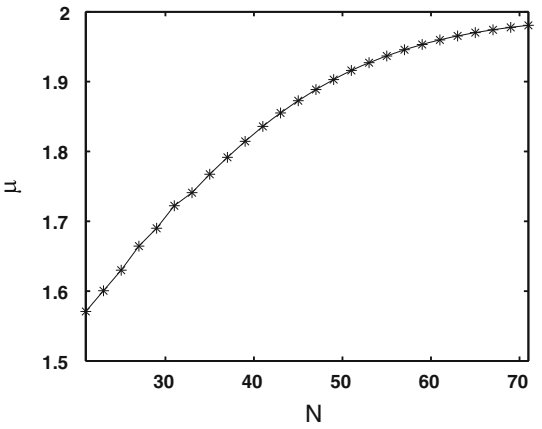

Fig. 5 Value of μ versus the filter size N

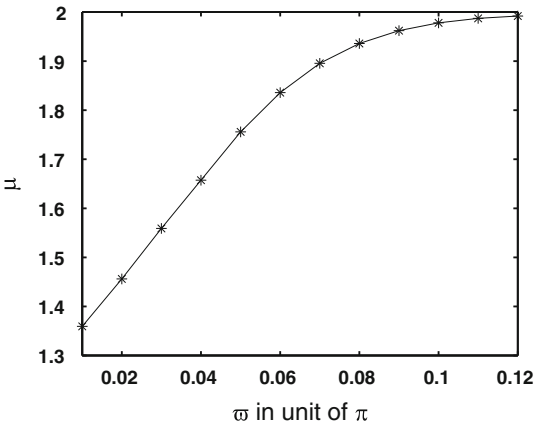

Fig. 6 Value of μ versus ϖ

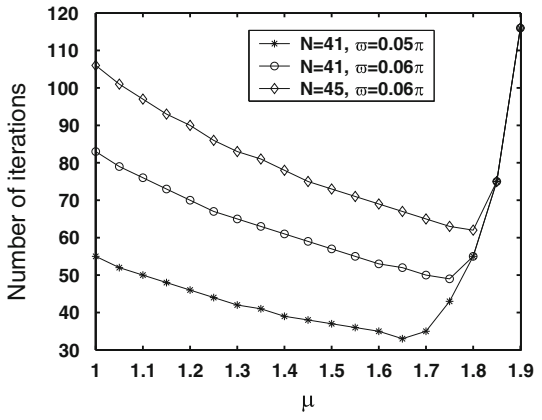

Fig. 7 Dependence of the iteration number on μ

We find from a large number of simulations that the optimal μ satisfies $1 < \mu < 2$. For most 2-D filters of different shapes except for the rectangular filters, for which Algorithm I is efficient enough, 1.7 is a very good value of μ to replace its optimal value.

5 Conclusions

Two efficient 2-D based algorithms have been presented in this paper for the WLS design of 2-D FIR filters with arbitrary weighting functions. Both algorithms have been derived from the optimal condition of the design problem. Efficiencies and design accuracies of these algorithms have been demonstrated and compared with existing algorithms through designing 2-D filters of different shapes, different orders and with different transition band widths. Results show that Algorithm II needs fewer iterations than Algorithm I and obtains more accurate solutions. Comparisons also show that Algorithm II consumes less design time than the conjugate gradient algorithm of Aravena and Gu (1996).

Acknowledgments This work was supported in part by the National Nature Science Foundation of China under Grants 61175001 and 60974102, in part by the National Basic Research Program of China under Grants 2012CB821200 and 2009CB320600, and in part by the Shandong Provincial Nature Science Foundation of China under Grant ZR2010FQ016.

References

Aravena, J. L., & Gu, G. (1996). Weighted least mean square design of 2-D FIR digital filters: The general case. *IEEE Transactions on Signal Processing, 44*(10), 2568–2578.

Gislason, E., Johansen, M., Ersboll, B. K., & Jacobsen, S. K. (1993). Three different criteria for the design of two-dimensional zero phase FIR digital filters. *IEEE Transactions on Signal Processing, 41*(10), 3070–3074.

Gohberg, I., Goldberg, S., & Kaashoek, M. A. (2003). *Basic classes of linear operators*. Basel: Birkhauser.

Gu, G., & Aravena, J. L. (1994). Weighted least mean square design of 2-D FIR digital filters. *IEEE Transactions on Signal Processing, 42*(11), 3178–3187.

Hsieh, C.-H., Kuo, C.-M., Jou, Y.-D, & Han, Y.-L (1997). Design of two-dimensional FIR digital filters by a two-dimensional WLS technique. *IEEE Transactions on Circuits and Systems-II, 44*(5), 348–358.

Lai, X. P., & Cheng, Y. (2007). A sequential constrained least-square approach to minimax design of 2-D FIR filters. *IEEE Transactions on Circuits and Systems-II, 54*(11), 994–998.

Lim, J. S. (1990). *Two-dimensional signal and image processing*. NJ: Prentice-Hall.

Lu, W. S., & Antoniou, A. (1992). *Two-dimensional digital filters*. New York: Marcel Dekker.

Lu, W. S. (2002). A unified approach for the design of 2-D digital filters via semidefinite programming. *IEEE Transactions on Circuits and Systems-I, 49*(6), 814–826.

Lu, W. S., & Hinamoto, T. (2011). Two-dimensional digital filters with sparse coefficients. *Multidimensional Systems and Signal Processing, 22*(1), 173–189.

Shyu, J. J., Pei, S. C., & Huang, Y. D. (2011). An iterative approach for minimax design of multidimensional quadrature mirror filters. *Signal Processing, 91*(8), 1730–1740.

Tzeng, T. S. (2007). Design of 2-D FIR digital filters with specified magnitude and group delay responses by GA approach. *Signal Processing, 87*(9), 2036–2044.

Zhao, R. J., & Lai, X. P. (2010). On the WLS design of 2-D quadrantally symmetric FIR filters with arbitrary weighting functions. In *The 3rd International symposium on systems and control in aeronautics and astronautics*, pp. 697–701.

Zhao, R. J., & Lai, X. P. (2011). A fast matrix iterative technique for the WLS design of 2-D quadrantally symmetic FIR filters. *Multidimensional Systems and Signal Processing, 22*(4), 303–317.

Zhu, W. P., Ahmad, O. M., & Swamy, M. N. S. (1997). A closed-form solution to the least-square design problem of 2-D linear-phase FIR filters. *IEEE Transactions on Circuits and Systems-II, 44*(12), 1032–1039.

Author Biographies

Ruijie Zhao was born in Hebei, China, on March 7, 1978. He received his B.S. degree from Shandong University at Weihai, Weihai, China in 2001 and M.S. degree from Shandong University, Jinan, China, in 2004. He has been with Shandong University at Weihai, Weihai, China, since 2004, where he became a Lecturer in 2006. His main research interests include digital filter design, optimization algorithm, and approximation of functions.

Xiaoping Lai was born in Jiangxi, China, on August 24, 1965. He received the B.S., M.S., and Ph.D. degree from Shandong University, Jinan, China, in 1985, 1988, and 2000, respectively, and became a Professor in 2001. He was with Shandong University at Weihai from 1988 to 2008, and has been with Hangzhou Dianzi University since Jan. 2008. His main research interests include digital filter design, optimization, and artificial neural networks.