計測自動制御学会東北支部 第225回研究集会 (2005.11.11)
資料番号 225-9

線形位相特性を持つ2次元状態空間ディジタルフィルタのGAによる設計

GA-Based Design of 2-D State Space Digital Filters with Linear Phase

○南 相哲*，阿部 正英*，川又 政征*

○Sang-Churl Nam*, Masahide Abe*, Masayuki kawamata*

*東北大学大学院工学研究科電子工学専攻

*Department of Electronic Engineering, School of Engineering, Tohoku University

キーワード： 2次元ディジタルフィルタ (2-D digital filters)，状態空間モデル (state space model)，遺伝的アルゴリズム (genetic algorithm)，線形位相特性 (linear phase characteristic)，群遅延 (group delay)

連絡先： 〒980-8579 仙台市青葉区荒巻字青葉6-6-05 東北大学 大学院 工学研究科 電子工学専攻 川又研究室
南 相哲，Tel.: (022)795-7095，Fax.: (022)263-9169，E-mail: nam@mk.ecei.tohoku.ac.jp

1. Introduction

Two of the important applications of 2-D digital filters are currently found in image processing and seismic signal processing. It has been shown that the phase response as well as the magnitude response of a 2-D digital filter is important and usually a zero phase or a linear phase response (corresponding to a constant group delay) is preferable in many cases [1,2]. For FIR filters, phase linearity is easily enforced analytically through coefficient symmetry. In the case of IIR filters with causal numerator and denominator, however, phase linearity can not be implemented. Thus, minimization of phase nonlinearity must be incorporated into the optimization procedure.

In this paper, we discuss a GA-based design method of 2-D state space digital filters which approximate not only a specified magnitude response but also a constant group delay. In our proposed method, we restrict ourselves to a class of 2-D separable denominator digital filters. This is because the stability test of the 2-D separable denominator digital filters is much easier than that of the general (non-separable denominator) 2-D digital filters. On the other hand, the 2-D separable denominator digital filters have a disadvantage that their performance does not excel general 2-D IIR digital filters in approximation of spatial-domain and frequency-domain specification. However, any 2-D spatial and frequency domain specifications can be approximated with 2-D separable denominator digital filters as nearly accurately as with general 2-D IIR filters.

The organization of this paper is as follows: Section 2 introduces Roesser's local state space model to represent 2-D separable denominator digital filters and describes the formulation of the design problem. Section 3 discusses some basic concepts of genetic algorithm and genetic operators used in our proposed method. After that, the coding problem of the filter coefficients and the calculation of the fitness value are explained. Section 4 shows the necessary steps for searching the optimal solution from the design problem. Section 5 explains the stability check of 2-D state space digital filters. Section 6 presents a numerical example to demonstrate the effectiveness of the proposed method. Finally, the concluding remarks are given in Section 7.

2. Statement of the Problem

2.1 2-D Separable Denominator Digital Filters

A linear, shift-invariant, and causal 2-D separable denominator recursive digital filter can be described with the following tranfer function:

$$H(z_1,z_2) = \frac{N(z_1,z_2)}{D_1(z_1)D_2(z_2)} = \frac{\sum_{k_1=0}^{N_1}\sum_{k_2=0}^{N_2} a(k_1,k_2) z_1^{-k_1} z_2^{-k_2}}{(1-\sum_{k_1}^{N_1} b_1^{(k_1)} z_1^{-k_1})(1-\sum_{k_2}^{N_2} b_2^{(k_2)} z_2^{-k_2})} \quad (1)$$

where $D_1(z_1)$ and $D_2(z_2)$ are polynomials of degrees N_1 in z_1^{-1} and N_2 in z_2^{-1}, respectively. The 2-D digital filter with the transfer function given by Eq. (1) can be represented by Roesser's local state-space model [3] as follows:

$$\begin{bmatrix} \boldsymbol{x}_h(n_1+1,n_2) \\ \boldsymbol{x}_v(n_1,n_2+1) \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{0} & \boldsymbol{A}_4 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_h(n_1,n_2) \\ \boldsymbol{x}_v(n_1,n_2) \end{bmatrix} + \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} u(n_1,n_2) \quad (2)$$

$$y(n_1,n_2) = \begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_h(n_1,n_2) \\ \boldsymbol{x}_v(n_1,n_2) \end{bmatrix} + du(n_1,n_2) \quad (3)$$

where $u(i,j)$ is the input, $y(i,j)$ is the output, $\boldsymbol{x}_h(i,j)$ is a horizontal state vector of order N_1, $\boldsymbol{x}_v(i,j)$ is a vertical state vector of order N_2, and matrices $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_4$, $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{c}_1$, $\boldsymbol{c}_2$, and scalar d are real coefficient matrices with appropriate dimensions.

By applying the 2-D z–transform to Eqs. (2) and (3) while assuming zero initial conditions ($\boldsymbol{x}_h(0,n_2) = \boldsymbol{x}_v(n_1,0) = 0$ for $n_1,n_2 \leqq 0$), we obtain the following transfer function:

$$H(z_1,z_2) = \boldsymbol{c}\left[z_1\boldsymbol{I}_{N_1} \oplus z_2\boldsymbol{I}_{N_2} - \boldsymbol{A}\right]^{-1}\boldsymbol{b} + d \quad (4)$$

where $\oplus$ denotes the direct sum of matrices, and $\boldsymbol{I}_{N_1}$ and $\boldsymbol{I}_{N_2}$ are $N_1 \times N_1$ and $N_2 \times N_2$ identity matrices, respectively. By letting $z_1 = e^{j\omega_1}$ and $z_2 = e^{j\omega_2}$ in Eq. (4), the frequency response $H(e^{j\omega_1},e^{j\omega_2})$ of the filter is expressed as

$$H(e^{j\omega_1},e^{j\omega_2}) = H(\omega_1,\omega_2)e^{j\phi(\omega_1,\omega_2)} \quad (5)$$

where

$$H(\omega_1,\omega_2) = |H(e^{j\omega_1},e^{j\omega_2})| \quad (6)$$

$$\phi(\omega_1,\omega_2) = \arg H(e^{j\omega_1},e^{j\omega_2}) \quad (7)$$

are the magnitude response and the phase response of the filter, respectively. The group delay functions are defined as

$$\tau_k = -\frac{\partial\phi(\omega_1,\omega_2)}{\partial\omega_k} \quad (k=1,2). \quad (8)$$

A 2-D digital filter should have a linear or almost linear phase response to minimize phase distortions in 2-D digital filtering. This is equivalent to having a constant or almost constant group delay.

2.2 Formulation of the Design Problem

The design problem is to find a set of coefficient matrices $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_4$, $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ and scalar d such that the function $H(e^{j\omega_1}, e^{j\omega_2})$ simultaneously approximates both the given magnitude and a prescribed group delay in the passband.

Let E_m be an error function between the desired and the resulting magnitude response. Let E_{τ_1} and E_{τ_2} be also error functions between the desired and the resulting group delays with respect to ω_1 and ω_2, respectively. Then the design problem can be formulated as follows:

[**Formulation**] Search a set of coefficient matrices $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_4$, $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ and scalar d which simultaneously minimize

$$E_m(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_4, \boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{c}_1, \boldsymbol{c}_2, d) \tag{9}$$

and

$$\begin{aligned} &E_{\tau_1}(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_4, \boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{c}_1, \boldsymbol{c}_2, d) \\ &+E_{\tau_2}(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_4, \boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{c}_1, \boldsymbol{c}_2, d) \end{aligned} \tag{10}$$

subject to the constraint that the resultant filter is stable, where

$$E_m = \left\{ \sum_m \sum_n |H_d(\omega_{1m}, \omega_{2n}) - H(\omega_{1m}, \omega_{2n})|^2 \right\}^{1/2} \tag{11}$$

$$E_{\tau_1} = \left\{ \sum\sum_{\text{passband}} |\tau_{d_1}(\omega_{1n}, \omega_{2n}) - \tau_1(\omega_{1m}, \omega_{2n})|^2 \right\}^{1/2} \tag{12}$$

$$E_{\tau_2} = \left\{ \sum\sum_{\text{passband}} |\tau_{d_2}(\omega_{1n}, \omega_{2n}) - \tau_2(\omega_{1m}, \omega_{2n})|^2 \right\}^{1/2} \tag{13}$$

where m and n are sample points in the 2-D frequency plane, H_d and H are the given magnitude response and the magnitude response of the resulting filter, respectively, and $\tau_{d_i}(i = 1, 2)$ and $\tau_i(i = 1, 2)$ are the given group delay and the group delay of the resulting filter with respect to ω_1 and ω_2, respectively.

3. Application of GA to the Design Problem

3.1 Genetic Algorithm

Genetic algorithms (GAs) have become a popular and powerful search algorithm, which is based on ideas borrowed from the theories of natural selection [4]. It is very useful to find the global optimum solution because of requiring no calculation of gradient and being not susceptible to local optimum problems with multi-modal error surface. In our proposed method, to apply the GA to the design problem, all elements of the filter coefficients are encoded into the Gray codes of B bits, and these are concatenated together to form a *chromosome*. After evaluating the fitness with respect to every chromosome, the most suitable solutions in population are likely to survive and to be transmitted to the new generation by using *selection*, *crossover* and *mutation*, which are called genetic operators. The better chromosomes are selected according to fitness value to reproduce the next generation, whereas the poorer chromosomes are lost. This selection of the better chromosomes decreases the mean fitness of population but does not generate any new chromosome. The selected chromosomes are carried out through a crossover procedure. The crossover takes two chromosomes (called parent1 and parent2) and generates two new chromosomes (called offspring1 and offspring2) [4]. The crossover procedure used in this paper is Uniform Crossover [5], which randomly generates the crossover mask with the same bit length as the chromosome, and exchanges a pair of chromosomes according to the crossover mask with crossover rate P_c. The crossover procedure can be illustrated as follows:

Parent 1 :	1 1 1 1 1 1 1 1
Parent 2 :	0 0 0 0 0 0 0 0
Mask :	1 0 0 0 1 0 1 0
Offspring 1 :	1 0 0 0 1 0 1 1
Offspring 2 :	0 1 1 1 0 1 0 0.

Note that even though selection and crossover effectively search the optimal solution in the design procedure, they may lose some potentially useful genetic material. Therefore, after the crossover, mutation procedure is introduced to protect a loss of useful genetic material from premature convergence. The mutation operation alters the gene from "0" to "1" or from "1" to "0" with mutation rate P_m [4].

Unlike traditional optimization methods, the GA can escape from local minimum value through the mutation procedure, even after it converges into local minimum value [4]. In our design procedure, we use an elitist strategy where one chromosome or a few of the best chromosomes are copied into the next generation without mutation and crossover procedure. The elitist strategy may increase the speed of domination of a population by a super chromosome, but on balance it appears to improve the performance of the GA [6]. The cycle of evolution is repeated until a desired termination criterion is satisfied. This termination criterion is set by the number of computational runs, or the amount of variation of the best solution between different generations, or a pre-defined value of fitness.

3.2 Coding and Fitness Function

As for the coding, all coefficients a_{ij}, b_i, c_j, d of the matrices $\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{c}$ and scalar d are encoded into the Gray codes of B bits inside the interval $(-1, 1)$. To decode the Gray codes into real values, the Gray codes are converted to the binary codes. Next to that, all binary codes are calculated to the corresponding real value according to the following equations:

$$a_{ij}^{(k)} = \frac{(-1)^{\alpha_{ij,B}}}{2^{B-1}} \sum_{b=0}^{B-1} \alpha_{ij,b} 2^{b-1} \tag{14}$$

$$b_{i}^{(k)} = \frac{(-1)^{\beta_{i,B}}}{2^{B-1}} \sum_{b=0}^{B-1} \beta_{i,b} 2^{b-1} \tag{15}$$

$$c_{j}^{(k)} = \frac{(-1)^{\gamma_{j,B}}}{2^{B-1}} \sum_{b=0}^{B-1} \gamma_{j,b} 2^{b-1} \tag{16}$$

$$d = \frac{(-1)^{\delta}}{2^{B-1}} \sum_{b=0}^{B-1} \delta 2^{b-1} \tag{17}$$

where $i = 1, 2, \cdots, m+n$, $j = 1, 2, \cdots, m+n$ and $k = 1, 2, \cdots, N$; N is a population size.

In order to obtain the fitness values for each chromosome, we first calculate the objective function values by using the following equations:

$$g_1^{(k)} = \begin{cases} E_m^{(k)}, & \text{if the filter is stable} \\ E_s, & \text{if the filter is unstable} \end{cases} \tag{18}$$

$$g_2^{(k)} = \begin{cases} E_{\tau_1}^{(k)} + E_{\tau_2}^{(k)}, & \text{if the filter is stable} \\ E_s, & \text{if the filter is unstable} \end{cases} \tag{19}$$

where the parameter E_s is a predefined big value, which is assigned to unstable chromosomes (filters).

Since each chromosome individually goes through the same evaluating exercise, the range of this value varies from one chromosome to another. To maintain uniformity of the range of this value, it is necessary for the objective values to be scaled. Rank based fitness scaling procedure [7] is used in our design problem as follows:

$$f^{(k)} = \text{ranking}(-g^{(k)}) \tag{20}$$

where $k = 1, 2, \cdots, N$; N is a population size, and ranking $(\cdot)$ represents rank based fitness scaling procedure. This rank based fitness scaling approach can help to avoid the premature convergence caused by "super chromosomes" that have an unusually high fitness ratio. In Eqs. (18) and (19), we assign a big value (E_s) to unstable chromosomes. The reason for this is that although the unstable chromosomes are used as genetic information in the next generation, their chances of survival will decrease by the selection scheme. As a consequence, it is concluded that unstable chromosomes will gradually decrease in number with generation renewal and disappear before convergence in the design procedure. By using rank based fitness scaling procedure, the fitness value with respect to $g_1^{(k)}$ and $g_2^{(k)}$ is obtained. To design a digital filter approximating both the magnitude and the group delay, we use the following cost function:

$$f^{(k)} = \alpha f_1^{(k)} + (1-\alpha) f_2^{(k)} \quad (0 \leqq \alpha \leqq 1). \tag{21}$$

where the positive constant α represents the important factor and is a measure of the significance of each objective in the optimization process.

4. Design Procedure based on a GA

There are several tuning parameters to be specified before executing the GA. These parameters are as follows:

N	population size
P_c	probability of crossover
P_m	probability of mutation
C_{tc}	termination condition.

The design algorithm using the GA is summarized as follows.

Step 0: (Initialization) Choose all GA parameters described in the above and generate N chromosomes randomly using the Gray codes.

Step 1: (Stability test and Evaluation) Check the stability of each chromosome and evaluate its fitness using Eq. (21).

Step 2: (Selection) Select new N chromosomes randomly by *Roulette wheel selection* approach based on each fitness.

Step 3: (Crossover) Execute the crossover operation on $N/2$ chromosome pairs selected in Step 2 with the probability P_c.

Step 4: (Mutation) Apply the mutation operation on N chromosomes generated in Step 3 with the probability P_m.

Step 5: Stop the design procedure if the value of the best solution does not improve in a prescribed termination constant C_{tc}-th generation. If not, repeat Steps $1-4$ with a new population generated in Step 4.

5. Stability Issue

The stability of 2-D separable denominator state-space digital filters is checked by using the following two conditions:

Theorem 1 A 2-D separable denominator state space digital filter is stable if and only if for its state transition matrix $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{0} & \boldsymbol{A}_4 \end{bmatrix}$,

$$\begin{aligned} |\lambda_i(\boldsymbol{A}_1)| < 1, \quad i = 1, 2, \cdots, m \\ |\lambda_j(\boldsymbol{A}_4)| < 1, \quad j = 1, 2, \cdots, n \end{aligned} \tag{22}$$

where $\lambda_i(\boldsymbol{A}_1)$ and $\lambda_j(\boldsymbol{A}_4)$ denote the eigenvalues of state transition matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_4$, respectively.

When we check the stability of randomly generated chromosome inside the interval $(-1, 1)$ using Theorem 1, it is turned out that more than a half of chromosomes are unstable. As a result of this fact, a stable 2-D digital filter with small approximation error can rarely be obtained because of premature convergence in the design procedure. We consider a strategy that restricts the search range of coefficients of the state transition matrix. Thus the following Gershgorin circle theorem [8] is introduced to maintain enough stable chromosomes during the design procedure.

Theorem 2 (**Gershrogin Circle Theorem**) If $\boldsymbol{A} = (a_{ij})$ is a complex square matrix and S_i is the disk in the complex plane, then all the eigenvalues of $\boldsymbol{A}$ lie in the union of certain disks $S_i\,(i = 1, 2, \cdots, n)$, whose centers are the values along the diagonal and whose radii are the sum of the absolute values of the off-diagonal entries in a given row.

This theorem identifies a region in the complex plane that contains all the eigenvalues of a square matrix $\boldsymbol{A}$. When we check the stability of all chromosomes whose coefficients of $\boldsymbol{A}_1$ and $\boldsymbol{A}_4$ are randomly generated inside $(-0.7, 0.7)$, the number of unstable chromosomes is reduced to one-tenth, and a stable filter with small approximation error can be obtained as demonstrated in Figs. 1 and 2. Consequently, we confine the search range of the coefficients of $\boldsymbol{A}_1$ and $\boldsymbol{A}_4$ to $(-0.7, 0.7)$ to maintain enough stable chromosomes and to obtain a stable filter with small approximation error.

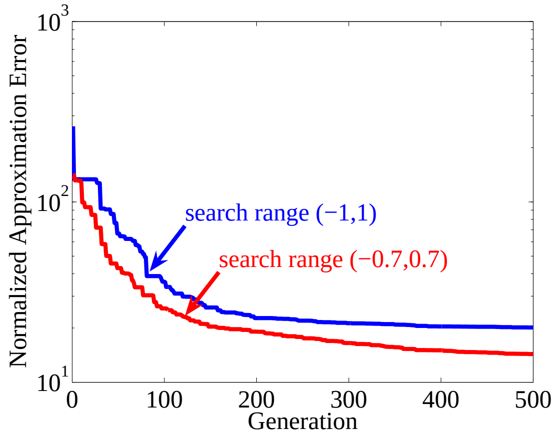

Fig. 1 Relation between search range and the number of unstable chromosomes.

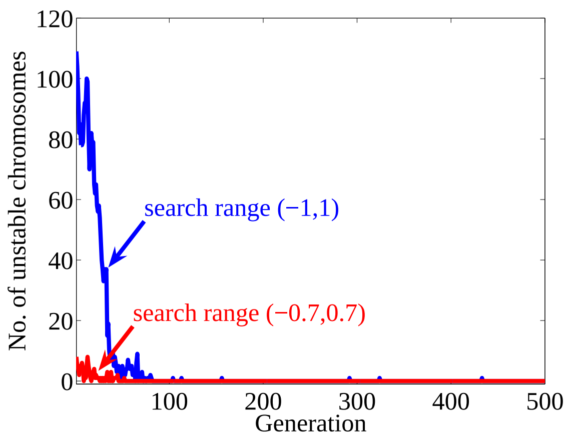

Fig. 2 Relation between search range and approximation error.

6. Numerical Example

Consider the problem of designing a 2-D circularly symmetric lowpass filter. Let the magnitude specification $H_d(\omega_1, \omega_2)$ be given as

$$H_d(\omega_1,\omega_2) = \begin{cases} 1.0, & 0.0 \leqq r \leqq 0.1 \\ 0.8, & 0.1 < r \leqq 0.2 \\ 0.44, & 0.2 < r \leqq 0.3 \\ 0.14, & 0.3 < r \leqq 0.4 \\ 0.03, & 0.4 < r \leqq 0.5 \\ 0.002, & 0.5 < r \leqq 0.6 \\ 0.001, & 0.6 < r \leqq 1.0 \end{cases} \tag{23}$$

where $r = \sqrt{\omega_{1m}^2 + \omega_{2n}^2}/\pi$ and the passband is the region specified by $r \leqq 0.3$. The group delay is

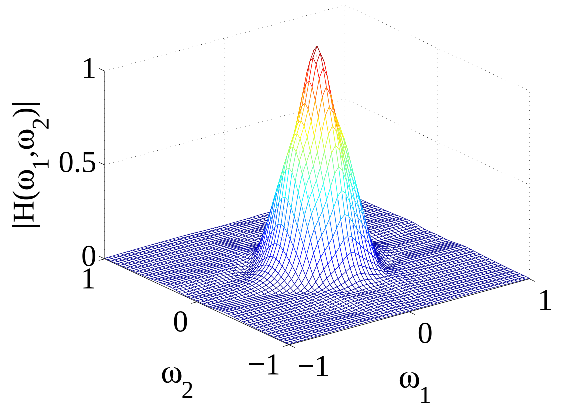

Fig. 3 Magnitude response of the resulting filter.

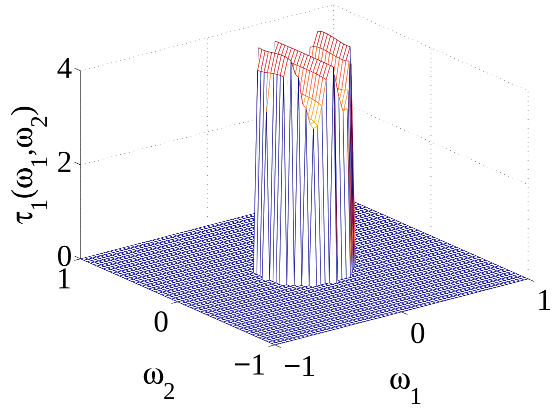

Fig. 4 Passband group delay with respect to ω_1.

specified as

$$\tau_{d_k}(\omega_1,\omega_2) = \begin{cases} 4.0, & 0.0 \leqq r \leqq 0.3 \\ 0.0, & \text{otherwise} \end{cases} \quad (k=1,2) \tag{24}$$

and the support region S is

$$S = \{(\omega_1,\omega_2)|\omega_1 = \pi m_1/10, \ -10 \leqq m_1 \leqq 10; \ \omega_2 = \pi n/10, \ 0 \leqq n \leqq 10\}. \tag{25}$$

For the given specifications, we design a $(4,4)$-order digital filter using MATLAB running on a XEON 3.2 [GHz] machine. The GA parameters and the weighting factor in Eq. (21) are used as follows:

weighting factor	$\alpha = 0.6$
population size	$N = 200$
crossover rate	$P_c = 0.95$
mutation rate	$P_m = 0.01$
terminal condition	$C_{tc} = 500$
bit length	$B = 16$.

Figures 3, 4 and 5 show the magnitude and group delay responses of the resultant filter. The design procedure stops after 1,488 iterations. The CPU time required for the design procedure is 57.71 minutes. Furthermore, for each generation, the calculations of genetic operators (selection, crossover and mutation operation) and the objective function take on average 0.29 sec and 1.42 sec, respectively

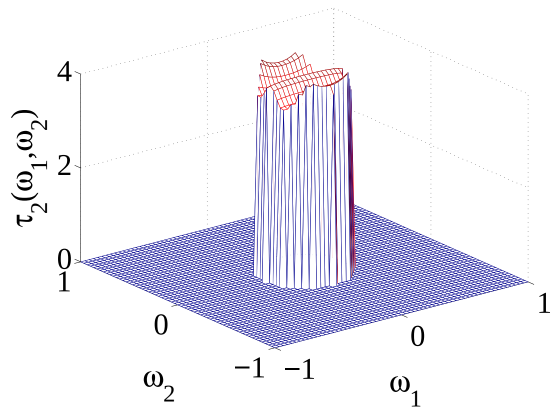

Fig. 5 Passband group delay with respect to ω_w..

Table 1 is given to compare the normalized approximation error analysis of our design with that of the filter designed in [10] and [9] in both group delay and the magnitude characteristics.

7. Conclusion

In this paper, we have proposed a GA-based design method of 2-D separable denominator state space digital filters which meet simultaneously with the magnitude response and a constant group delay specifications. In this proposed approach, we have first formulated the design problem. And then, in order to apply the GA to the design problem, we have presented the coding of all coefficients and an objective function. To check the stability of the

Table 1 Approximation error analysis.

	Order	ε_m	ε_{τ_1}	ε_{τ_2}
our design	$(4,4)$	16.39%	3.48%	1.98%
design in [9]	$(4,4)$	15.58%	0.69%	0.69%
design in [10]	$(4,4)$	24.36%	9.32%	8.18%

resulting 2-D state space digital filter, we embedded the stability test routine in the design procedure. Therefore, the stability of the resulting filter is guaranteed. A numerical example has shown that the normalized approximation error of the resulting filter is smaller than or gives somewhat comparable to those of the other methods.

References

1) Jae S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice-Hall, 1990.

2) W. S. Lu and A. Antoniou, *Two-Dimensional Digital Filters*, Marcel Dekker, 1992.

3) R. P. Roesser, "A discrete state-space model for linear image processing," *IEEE Transactions of Automatic Control*, vol. AC-20, pp. 1–10, Feb. 1975.

4) D. E. Goldberg, *Genetic Algorithm in Search, Optimization, and Machine Learning*, Addison Wesley, Massachusetts, 1989.

5) G. Syswerda, "Uniform crossover in genetic algorithms," *Proceedings of the third International Conference on Genetic Algorithms*, pp. 2–9, 1989.

6) K. F. Man, K. S. Tang, S. Kwong, and W. A. Halang, *Genetic Algorithms for Control and Singal Processing*, Springer, 1997.

7) D. Whitley, "The GENITOR algorithm and selection pressure: Why rank-based allocation of reproductive trials is best," *Proceedings of the third International Conference on Genetic Algorithms*, pp. 116–121, 1989.

8) Gene H. Golub and Charles F. Van Loan, *Matrix Computations*, Johns Hopkins, 3 edition, 1996.

9) Tian-Bo Deng and Masayuki Kawamata, "Frequency-domain desing of 2-D digital filters using the iterative singular value decomposition," *IEEE Transaction on Circuit and System*, vol. CAS-38, no. 10, pp. 1225–1228, Nov. 1991.

10) Samy A. H. Aly and Moustafa M. Fahmy, "Design of two-dimensional recursive digital filters with specified magnitude and group delay characteristics," *IEEE Transaction on Circuit and System*, vol. CAS-25, no. 11, pp. 908–915, Nov. 1978.