Design of Two-Dimensional Digital Filters Using Singular-Value Decomposition and Balanced Approximation Method

Wu-Sheng Lu, *Senior Member, IEEE*, Hui-Ping Wang, and Andreas Antoniou, *Fellow, IEEE*

***Abstract*—It is shown that the singular-value decomposition (SVD) of the sampled amplitude response of a two-dimensional (2-D) digital filter possesses a special structure: every singular vector is either mirror-image symmetric or antisymmetric with respect to its midpoint. Consequently, the SVD can be applied along with 1-D finite-impulse response (FIR) techniques for the design of linear-phase 2-D filters with arbitrary prescribed amplitude responses which are symmetrical with respect to the origin of the (ω_1, ω_2) plane. In the second half of the paper, the well-known balanced approximation method is applied to linear-phase 2-D FIR filters of the type that may be obtained by using the SVD method. The method leads to economical and computationally efficient filters, usually infinite-impulse response filters, which have prescribed amplitude responses and whose phase responses are approximately linear.**

I. Introduction

Linear-phase two-dimensional (2-D) digital filters find many applications in image processing where phase distortion is particularly objectionable [1]. Filters of this type can be designed by using the window method [2]–[4], the McClellan transformation [5]–[8], and as was shown recently, by using the singular-value decomposition (SVD) [9], [10]. The last method offers two important advantages. First, the design can be accomplished by designing a set of 1-D subfilters using well-known 1-D techniques and, second, the resulting 2-D filter can be fast since the 1-D subfilters form a parallel structure which allows extensive pipelining. A limitation of the SVD method as reported in [10] is that the amplitude response of the filter is required to have quadrantal symmetry with respect to the (ω_1, ω_2) plane. This restricts the application of the method to some extent.

In this paper, it is shown that the SVD of the sampled amplitude response of a 2-D digital filter possesses a special structure: every singular vector is either mirror-image symmetric or antisymmetric about its midpoint. As a result, the SVD method can be applied along with 1-D FIR filter techniques for the design of linear-phase 2-D filters with arbitrary amplitude responses which are symmetrical with respect to the origin of the (ω_1, ω_2) plane. A quantitative error analysis is then described which can facilitate the determination of the number of subfilters to be used and the required design accuracy for each of the 1-D subfilters to achieve a desired approximation error.

In the second half of the paper, a model reduction technique known as balanced approximation (BA) method [11]–[14] is applied in the design of 2-D digital filters. The design starts with a linear-phase 2-D FIR filter of the type that may be obtained by using the proposed general SVD method and concludes with a lower order separable 2-D digital filter, usually an infinite-impulse response (IIR) filter. As will be shown, the approximation error in the phase angle is bounded by the sum of the neglected Hankel singular values of the filter. Consequently, the phase response of the resulting filter is approximately linear over the passband region provided that only small Hankel singular values are neglected. Under these circumstances, the resulting 2-D filter is nearly balanced, which implies that the filter has low roundoff noise as well as low parameter sensitivity. Furthermore, the 2-D filter obtained is more economical and computationally more efficient than the original 2-D FIR filter. Therefore, combining the general SVD and BA methods leads to excellent results as is demonstrated by designing a low-pass filter with a rotated elliptical passband.

II. General SVD Method

A. Design

A 2-D FIR single-input single-output (SISO) digital filter with support in the rectangle defined by $-N_i/2 \le n_i \le N_i/2$, $i = 1, 2$, can be characterized by the transfer function

$$H(z_1, z_2) = \sum_{n_1=-N_1/2}^{N_1/2} \sum_{n_2=-N_2/2}^{N_2/2} h(n_1, n_2) z_1^{-n_1} z_2^{-n_2} \quad (1)$$

where $h(n_1, n_2)$ is the impulse response of the filter. If $h(n_1, n_2) = h(-n_1, -n_2)$ and $h(n_1, n_2)$ is real then the frequency response denoted by

$$X(\omega_1, \omega_2) = H(e^{j\omega_1 T_1}, e^{j\omega_2 T_2}) \quad (2)$$

Manuscript received November 26, 1988; revised September 12, 1990. This work was supported by the Natural Sciences and Engineering Research Council of Canada.

The authors are with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, V8W 3P6, Canada.

IEEE Log Number 9101880.

1053-587X/91/1000-2253$01.00 © 1991 IEEE

is a real function which is symmetrical with respect to the origin of the (ω_1, ω_2) plane, i.e.,

$$X(\omega_1, \omega_2) = X(-\omega_1, -\omega_2). \tag{3}$$

A 2-D FIR filter having an arbitrary amplitude response satisfying (3) can readily be designed by using a parallel arrangement of K 2-D FIR sections each comprising two 1-D subfilters in cascade as will now be demonstrated. Such an arrangement can be represented by the transfer function

$$H(z_1, z_2) = \sum_{i=1}^{K} F_i(z_1) G_i(z_2) \tag{4}$$

where $F_i(z_1)$ and $G_i(z_2)$ are the transfer functions of two cascaded 1-D SISO subfilters. If these subfilters are FIR filters with support in the rectangle defined by $-N_i/2 \le n_i \le N_i/2$, $i = 1, 2$, we have

$$F_i(z_1) = \sum_{n_1=-N_1/2}^{N_1/2} f_i(n_1) z_1^{-n_1} \tag{5}$$

and

$$G_i(z_2) = \sum_{n_2=-N_2/2}^{N_2/2} g_i(n_2) z_2^{-n_2} \tag{6}$$

and if $F_i(z_1)$ and $G_i(z_2)$ are assumed to represent zero-phase or $\pi/2$-phase filters, then their frequency responses are given by

$$F_i(e^{j\omega_1 T_1}) = \Phi_i(\omega_1) e^{j\theta_i} \tag{7}$$

and

$$G_i(e^{j\omega_2 T_2}) = \Gamma_i(\omega_2) e^{j\theta_i}. \tag{8}$$

If $f_i(n_1)$ and $g_i(n_2)$ are mirror-image symmetric, then $\theta_i = 0$ in (7) and (8) and $\Phi_i(\omega_1)$ and $\Gamma_i(\omega_2)$ are real functions which are even with respect to ω_1 and ω_2, respectively; if $f_i(n_1)$ and $g_i(n_2)$ are mirror-image antisymmetric, then $\theta_i = \pi/2$ and $\Phi_i(\omega_1)$ and $\Gamma_i(\omega_2)$ are real functions which are odd with respect to ω_1 and ω_2, respectively. Under these circumstances, a zero-phase 2-D filter is obtained whose frequency response is given by (4)–(8) as

$$X(\omega_1, \omega_2) = \sum_{i=1}^{K} \pm\, \Phi_i(\omega_1) \Gamma_i(\omega_2) \tag{9}$$

where "+" corresponds to $\theta_i = 0$, and "−" corresponds to $\theta_i = \pi/2$.

Now assume that matrix $A = \{a_{l,m}\}$ represents a desired arbitrary frequency response such that (3) is satisfied, i.e.,

$$X(\pi\mu_l, \pi\nu_m) = X(-\pi\mu_l, -\pi\nu_m) = a_{l,m} \tag{10}$$

where μ_l, ν_m for $1 \le l \le L$ and $1 \le m \le M$ are normalized frequencies ranging over the baseband. Let

$$A = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T = \sum_{i=1}^{r} \tilde{\boldsymbol{u}}_i \tilde{\boldsymbol{v}}_i^T \tag{11}$$

be the SVD of matrix $\boldsymbol{A}$ where σ_i are the singular values of $\boldsymbol{A}$ such that $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$, and $\tilde{\boldsymbol{u}}_i = \sigma_i^{1/2} \boldsymbol{u}_i$, $\tilde{\boldsymbol{v}}_i = \sigma_i^{1/2} \boldsymbol{v}_i$. An important property of a class of FIR filters can now be stated in terms of the following theorem.

Theorem 1: If the frequency response of an FIR filter satisfies (3), then vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ in (11) are either mirror-image symmetric or antisymmetric simultaneously for $i = 1, 2, \cdots, r$.

Proof: Let $\tilde{\boldsymbol{H}} = \{\tilde{h}_{i,j},\ 1 \le i, j \le 2N\}$ be a $2N \times 2N$ arbitrary matrix with real elements such that

$$\tilde{h}_{i,j} = \tilde{h}_{2N+1-i, 2N+1-j} \tag{12}$$

and, for the sake of simplicity, assume that the matrix is of square even size and has distinct singular values. If matrices $\hat{\boldsymbol{I}}$, $\tilde{\boldsymbol{I}}$, and $\boldsymbol{H}$ are defined by

$$\hat{\boldsymbol{I}} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix}, \qquad \tilde{\boldsymbol{I}} = \begin{bmatrix} \boldsymbol{I}_N & \boldsymbol{0} \\ \boldsymbol{0} & \hat{\boldsymbol{I}}_N \end{bmatrix}, \text{ and}$$

$$\boldsymbol{H} = \tilde{\boldsymbol{I}} \tilde{\boldsymbol{H}} \tilde{\boldsymbol{I}} = \begin{bmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_2 \\ \boldsymbol{H}_2 & \boldsymbol{H}_1 \end{bmatrix}$$

respectively, where the sizes of $\hat{\boldsymbol{I}}_N$, $\boldsymbol{I}_N$, $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are $N \times N$, and the sizes of $\tilde{\boldsymbol{I}}$ and $\boldsymbol{H}$ are $2N \times 2N$, respectively, then matrix $\tilde{\boldsymbol{H}}$ can be decomposed as

$$\begin{aligned} \tilde{\boldsymbol{H}} &= \tilde{\boldsymbol{I}} \boldsymbol{H} \tilde{\boldsymbol{I}} = \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^T \\ &= [\boldsymbol{u}_1 \cdots \boldsymbol{u}_i \cdots \boldsymbol{u}_{2N}] \boldsymbol{\Sigma} [\boldsymbol{v}_1 \cdots \boldsymbol{v}_i \cdots \boldsymbol{v}_{2N}]^T \end{aligned} \tag{13}$$

where $\boldsymbol{u}_i$, $\boldsymbol{v}_i$ are normalized eigenvectors of $\tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^T$ and $\tilde{\boldsymbol{H}}^T\tilde{\boldsymbol{H}}$, respectively, i.e.,

$$\tilde{\boldsymbol{H}} \tilde{\boldsymbol{H}}^T \boldsymbol{u}_i = \sigma_i \boldsymbol{u}_i. \tag{14}$$

Substituting (13) into (14), we have

$$\boldsymbol{H} \boldsymbol{H}^T \tilde{\boldsymbol{I}} \boldsymbol{u}_i = \sigma_i \tilde{\boldsymbol{I}} \boldsymbol{u}_i. \tag{15}$$

If we let

$$\boldsymbol{u}_i = \begin{bmatrix} \boldsymbol{u}_{i1} \\ \boldsymbol{u}_{i2} \end{bmatrix}$$

then

$$\tilde{\boldsymbol{I}} \boldsymbol{u}_i = \begin{bmatrix} \boldsymbol{u}_{i1} \\ \hat{\boldsymbol{I}} \boldsymbol{u}_{i2} \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{x}_{i1} \\ \boldsymbol{x}_{i2} \end{bmatrix}$$

and, therefore (15) becomes

$$\boldsymbol{H} \boldsymbol{H}^T \begin{bmatrix} \boldsymbol{x}_{i1} \\ \boldsymbol{x}_{i2} \end{bmatrix} = \sigma_i \begin{bmatrix} \boldsymbol{x}_{i1} \\ \boldsymbol{x}_{i2} \end{bmatrix}. \tag{16}$$

Note that

$$\boldsymbol{H} \boldsymbol{H}^T \equiv \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{bmatrix}$$

where A is positive semidefinite and B is symmetric. Therefore, (16) can be expressed as

$$\begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix} = \sigma_i \begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix}. \tag{17}$$

By writing (17) in another matrix notation, we have

$$\begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{bmatrix} x_{i2} \\ x_{i1} \end{bmatrix} = \sigma_i \begin{bmatrix} x_{i2} \\ x_{i1} \end{bmatrix}. \tag{18}$$

Now on comparing (18) with (17), we note that both vectors

$$\begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix} \text{ and } \begin{bmatrix} x_{i2} \\ x_{i1} \end{bmatrix}$$

are eigenvectors of matrix HH^T associated with the same eigenvalue σ_i and, therefore, the two vectors must be linearly dependent, i.e., they must satisfy

$$\begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix} = \pm \begin{bmatrix} x_{i2} \\ x_{i1} \end{bmatrix}$$

which implies that

$$x_{i1} = x_{i2}, \quad \text{or } x_{i1} = -x_{i2}.$$

If $x_{i1} = x_{i2} \equiv x_i$, we can write

$$\hat{I}u_i = \begin{bmatrix} x_i \\ x_i \end{bmatrix}, \quad \text{and } u_i = \begin{bmatrix} x_i \\ \hat{I}x_i \end{bmatrix}$$

which means that u_i is mirror-image symmetric. On the other hand, if $x_{i1} = -x_{i2} \equiv x_i$, we have

$$u_i = \begin{bmatrix} x_i \\ -\hat{I}x_i \end{bmatrix}$$

which implies that u_i is mirror-image antisymmetric. Furthermore, by (13)

$$V = \hat{I}H^T\hat{I}U\Sigma^{-1}. \tag{19}$$

If

$$u_i = \begin{bmatrix} x_i \\ \hat{I}x_i \end{bmatrix}$$

then (19) implies

$$v_i = \sigma_i^{-1}\hat{I}H^T\hat{I}u_i = \sigma_i^{-1} \begin{bmatrix} (H_1 + H_2)^T x_i \\ \hat{I}(H_1 + H_2)^T x_i \end{bmatrix} = \begin{bmatrix} y_i \\ \hat{I}y_i \end{bmatrix}$$

where $y_i = \sigma_i^{-1}(H_1 + H_2)^T x_i$. If

$$u_i = \begin{bmatrix} x_i \\ -\hat{I}x_i \end{bmatrix}$$

then (19) implies

$$v_i = \sigma_i^{-1}\hat{I}H^T\hat{I}u_i = \sigma_i^{-1} \begin{bmatrix} (H_2 - H_1)^T x_i \\ -\hat{I}(H_2 - H_1)^T x_i \end{bmatrix} = \begin{bmatrix} \hat{y}_i \\ -\hat{I}\hat{y}_i \end{bmatrix}$$

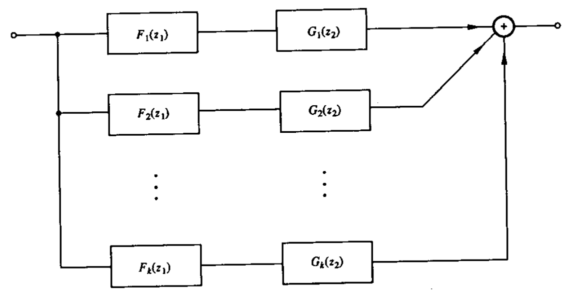


Fig. 1. Parallel realization of 2-D digital filter.

where $\hat{y}_i = \sigma_i^{-1}(H_1 - H_2)^T x_i$. This shows that the two vectors u_i and v_i have the same symmetry properties simultaneously, that is, they are either both mirror-image symmetric or antisymmetric. □

On comparing (9) with (11), $\tilde{u}_i$ and $\tilde{v}_i$ may be taken to be sampled versions of the frequency responses $\Phi_i(\omega_1)$ and $\Gamma_i(\omega_2)$, respectively. By designing the 1-D FIR filters characterized by $F_i(z_1)$ and $G_i(z_2)$ ($1 \le i \le K$, $1 \le K \le r$) as zero-phase or $\pi/2$-phase filters and then interconnecting the filters obtained as in Fig. 1, a zero-phase 2-D filter is obtained whose amplitude response is a good approximation to the desired amplitude response in the minimal mean-square-error sense [9]. The impulse response of the resulting filter is given by

$$h(n_1, n_2) = \sum_{i=1}^{K} f_i(n_1) g_i(n_2).$$

A 2-D causal linear-phase filter can readily be obtained by shifting the impulse response by $N_1/2$ and $N_2/2$ samples with respect to the n_1 axis and n_2 axis, respectively. This can be accomplished by multiplying $F_i(z_1)$ and $G_i(z_2)$ by $z_1^{-N_1/2}$ and $z_2^{-N_2/2}$, respectively. The design of the 2-D filter can be completed by using any one of the standard methods for the design of 1-D FIR filters. Using the Fourier series method in conjunction with window techniques [15], designs can be obtained very quickly with a small amount of computational effort. These designs are not optimal although the approximation error can be made arbitrarily small by increasing the order of the 1-D filters used. On the other hand, by using methods based on the Remez algorithm [16], it may be possible to obtain optimal designs although a large amount of computation would be required.

In the above design approach, the sampling density must be high enough in order for matrix A to represent the ideal amplitude response well. However, since the dimension of A, namely $L \times M$, determines the dimensions of vectors $\hat{u}_i$ and $\hat{v}_i$, a high sampling density leads to high-order FIR filters. Furthermore, the SVD of matrix A is time consuming and may entail numerical ill conditioning, particularly if L or $M \ge 100$. Based on a variety of designs we carried so far, a sampling density $L \times M$ with $60 < L, M < 90$ is appropriate if the order of the FIR filter to be designed is in the range 25×25 to 41×41.

B. Error Analysis

Assume that the SVD method has been used to design a 2-D FIR filter for a desired frequency response A and that the transfer function of the filter $H(z_1, z_2)$ is given by (4). Let $\tilde{f}_i$ and $\tilde{g}_i$ be the column vectors obtained by evaluating $\Phi_i(\omega_1)$ and $\Gamma_i(\omega_2)$ at frequencies $\omega_1 = \pi\mu_l/T_1$ and $\omega_2 = \pi\nu_m/T_2$, where $1 \le l \le L$ and $1 \le m \le M$, and T_1 and T_2 are the sampling periods in the horizontal and vertical directions, respectively. The amplitude response of the 2-D filter at frequency point $(\omega_1, \omega_2) = (\pi\mu_l/T_1, \pi\nu_m/T_2)$ is the (l, m) entry of matrix $\Sigma_1^K \tilde{f}_i\tilde{g}_i^T$ and, therefore, the approximation error at this frequency point is the (l, m) entry in the error matrix E defined by

$$\begin{aligned} E &= \{e_{l,m}\} \\ &= \sum_{i=1}^{K} \tilde{f}_i\tilde{g}_i^T - A \\ &= \sum_{i=1}^{K} (\tilde{f}_i\tilde{g}_i^T - \tilde{u}_i\tilde{v}_i^T) - \sum_{i=K+1}^{r} \sigma_i u_i v_i^T. \end{aligned} \quad (20)$$

If we define

$$\Delta\tilde{u}_i = \tilde{f}_i - \tilde{u}_i, \; \Delta\tilde{v}_i = \tilde{g}_i - \tilde{v}_i$$

then $\Delta\tilde{u}_i$ and $\Delta\tilde{v}_i$ represent the approximation errors in the 1-D frequency responses $\Phi_i(\omega_1)$ and $\Gamma_i(\omega_2)$, respectively. From (20) if follows that

$$e_{l,m} = \sum_{i=1}^{K} [\sigma_i^{1/2}(\Delta\tilde{u}_{il} v_{im} + u_{il}\Delta\tilde{v}_{im}) + \Delta\tilde{u}_{il}\Delta\tilde{v}_{im}] - \sum_{i=K+1}^{r} \sigma_i u_{il} v_{im}$$

where u_{il} denotes the lth component of vector u_i, etc. Since vectors u_i and v_i are unit vectors, an upper bound on the magnitude of $e_{l,m}$ can be obtained as

$$|e_{l,m}| \le \sum_{i=1}^{K} [\sigma_i^{1/2}(\epsilon_{1i} + \epsilon_{2i}) + \epsilon_{1i}\epsilon_{2i}] + \sum_{i=K+1}^{r} \sigma_i \quad (21)$$

where ϵ_{1i} and ϵ_{2i} represent the maximum approximation errors in the frequency responses $\Phi_i(\omega_1)$ and $\Gamma_i(\omega_2)$, respectively.

If the approximation error in each 1-D filter design is reasonably small, then the high-order terms $\epsilon_{1i}\epsilon_{2i}$ in inequality (21) can be neglected. We note also that the right-hand side in (21) is independent of l and m, and hence (21) holds also for the maximum of $|e_{l,m}|$. We have, therefore, obtained an upper bound for the maximum approximation error over the set of sampled frequency points as

$$e_\infty = \max_{1 \le l \le L, 1 \le m \le M} |e_{l,m}| \le \epsilon_{pK} + \epsilon_{rK} \quad (22)$$

where the principal error ϵ_{pK} and the residual error ϵ_{rK} are defined by

$$\epsilon_{pK} = \sum_{i=1}^{K} \sigma_i^{1/2}(\epsilon_{1i} + \epsilon_{2i})$$

and

$$\epsilon_{rK} = \sum_{i=K+1}^{r} \sigma_i \quad (23)$$

respectively.

The error bound given by (22) shows clearly how the choice of K and the approximation error introduced by a specific 1-D design technique affect the overall approximation error in the 2-D filter. As $K \to r$, the residual error ϵ_{rK} will tend to zero and the approximation error in the design of the 2-D filter is reduced to that introduced in the design of the 1-D FIR filters but the number of multiplications required becomes very large. Hence, in practice, K should be chosen sufficiently small to keep the number of parallel sections small and the residual error ϵ_{rK} acceptable. Having determined the value of K, the principal error ϵ_{pK} can also be made acceptable by controlling the 1-D design errors ϵ_{1i} and ϵ_{2i} for $1 \le i \le K$, by increasing the orders of the 1-D filters or by using a better design method for the 1-D filters. From (23) it follows that small approximation error should be obtained in the design of those 1-D filters that correspond to the large singular values.

C. Example 1

In this section, a 2-D low-pass FIR filter with a rotated elliptical passband is designed to illustrate the effectiveness of the proposed method.

The desired amplitude response of the 2-D low-pass FIR filter, shown in Fig. 2(a), is specified by

$$|H(e^{j\omega_1 T_1}, e^{j\omega_2 T_2})| = \begin{cases} 1 & (\tilde{\omega}_1/\omega_{p_1})^2 + (\tilde{\omega}_2/\omega_{p2})^2 \le 1 \\ 0 & (\tilde{\omega}_1/\omega_{a_1})^2 + (\tilde{\omega}_2/\omega_{a2})^2 > 1 \end{cases}$$

where

$$\tilde{\omega}_1 = \omega_1 \cos\alpha + \omega_2 \sin\alpha$$
$$\tilde{\omega}_2 = -\omega_1 \sin\alpha + \omega_2 \cos\alpha$$

and

$$\alpha = \pi/6, \; \omega_{p_1} = 0.32\pi, \quad \omega_{p_2} = 0.52\pi$$
$$\omega_{a_1} = 0.48\pi, \quad \omega_{a_2} = 0.68\pi.$$

With $L = M = 61$, the corresponding sampled amplitude response $A = |H(e^{j\pi\mu_l}, e^{j\pi\nu_m})|$ can be expressed as

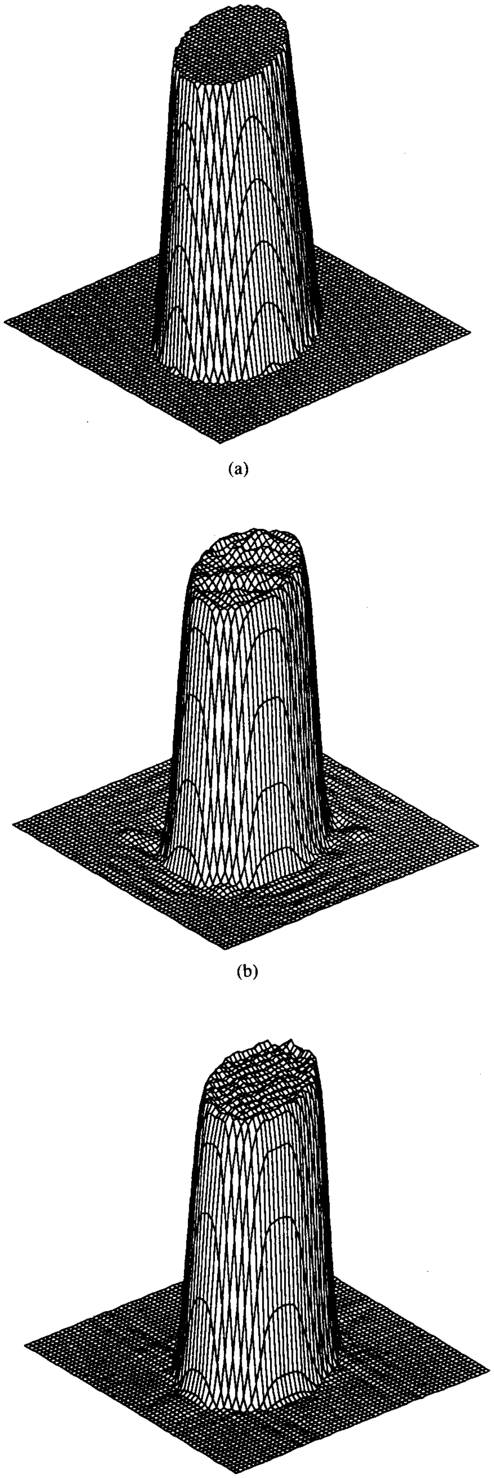
$$|H(e^{j\pi\mu_l}, e^{j\pi\nu_m})| = \begin{cases} 1 & (\tilde{\mu}_l/\omega_{c_1})^2 + (\tilde{\nu}_m/\omega_{c_2})^2 \le 1/\pi^2 \\ 0 & \text{otherwise} \end{cases}$$

where

$$\tilde{\mu}_l = \mu_l \cos\alpha + \nu_m \sin\alpha, \quad 1 \le l \le 61$$
$$\tilde{\nu}_m = -\mu_l \sin\alpha + \nu_m \cos\alpha, \quad 1 \le m \le 61$$

and

$$\omega_{c_1} = \tfrac{1}{2}(\omega_{a_1} + \omega_{p_1}), \quad \omega_{c_2} = \tfrac{1}{2}(\omega_{a_2} + \omega_{p_2}).$$



(a)

(b)

(c)

Fig. 2. Amplitude response of 2-D filter with rotated elliptical passband: (a) ideal amplitude response; (b) amplitude response of 2-D FIR filter obtained by using SVD method ($N_1 = N_2 = 29$, $K = 12$); (c) amplitude response of 2-D IIR filter obtained by using SVD and BA methods ($N_1 = 13$, $N_2 = 15$).

TABLE I
MAXIMUM PASSBAND AND STOPBAND ERRORS FOR FIR FILTER DESIGNED BY USING THE GENERAL SVD METHOD

K	Passband	Stopband
4	0.1360	0.1025
5	0.0854	0.0771
12	0.0393	0.0255
14	0.0365	0.0129
15	0.0295	0.0121
16	0.0150	0.0115
18	0.0124	0.0113
25	0.0117	0.0102

The software package MATLAB has been used to obtain the SVD of matrix $\boldsymbol{A}$. The 1-D FIR filters were designed by using the Fourier series method along with the Kaiser window function. As may be expected, the higher the order of the 1-D filters, the lower the approximation error. By trial and error, it has been found that a value of 29 for both N_1 and N_2 gives satisfactory results.

There were 25 nonzero singular values resulting from the SVD of matrix $\boldsymbol{A}$. The resulting amplitude response of the low-pass 2-D FIR filter for $K = 12$ is shown in Fig. 2(b). The maximum passband and stopband errors for $K = 4, 5, 12, 14, 15, 16, 18$, and 25 are given in Table I.

III. Application of Balanced Approximation Method

In what follows, separable IIR 2-D digital filters with linear phase response are designed by applying the balanced approximation (BA) method to high-order FIR 2-D filters. Some properties associated with the proposed approach are also presented and proved. Further, a design example is included to illustrate the method and to address some computation issues.

A. Design

The BA method has been applied extensively in the past in both 1-D as well as 2-D dynamical systems [17]–[23]. The method leads to more economical systems with reduced computational complexity. In this section, it is shown that the BA method can be put to good use in the design of 2-D digital filters as well.

Let

$$H(z_1, z_2) = \sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} h(n_1, n_2) z_1^{-n_1} z_2^{-n_2} \tag{24}$$

be the transfer function of a 2-D FIR filter of order (N_1, N_2). In Roesser's local state-space characterization [24], (24) can be represented by

$$\begin{aligned}\begin{bmatrix} \boldsymbol{x}^h(i+1, j) \\ \boldsymbol{x}^v(i, j+1) \end{bmatrix} &= \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}^h(i, j) \\ \boldsymbol{x}^v(i, j) \end{bmatrix} + \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} u(i, j) \\ &\equiv \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}u \\ y(i, j) &= [\boldsymbol{c}_1 \quad \boldsymbol{c}_2] \begin{bmatrix} \boldsymbol{x}^h(i, j) \\ \boldsymbol{x}^v(i, j) \end{bmatrix} + du(i, j) \\ &\equiv \boldsymbol{c}\boldsymbol{x} + du \end{aligned} \tag{25}$$

where [25]

$$A_1 = \begin{bmatrix} \mathbf{0} & I_{N_1-1} \\ 0 & \mathbf{0} \end{bmatrix}, \quad A_2 = \begin{bmatrix} h_{1N_2} & \cdots & h_{11} \\ \vdots & \cdots & \vdots \\ h_{N_1N_2} & \cdots & h_{N_11} \end{bmatrix}$$

$$A_3 = \mathbf{0}, \quad A_4 = \begin{bmatrix} \mathbf{0} & I_{N_2-1} \\ 0 & \mathbf{0} \end{bmatrix}$$

$$b = [h_{10} \cdots h_{N_10} | 0 \cdots 0 \quad 1]^T$$

$$c = [1 \quad 0 \cdots 0 | h_{0N_2} \cdots h_{01}] \quad \text{and } d = h_{00}.$$

If

$$f(z_1, z_2) = [I(z_1, z_2) - A]^{-1} b$$

$$g(z_1, z_2) = c[I(z_1, z_2) - A]^{-1}$$

where $I(z_1, z_2) = z_1 I_{N_1} \oplus z_2 I_{N_2}$, then the generalized reachability and observability gramians of (25) are defined as

$$K = \frac{1}{(2\pi j)^2} \oint_{|z_1|=1} \oint_{|z_2|=1} f(z_1, z_2) f^*(z_1, z_2) \frac{dz_1}{z_1} \frac{dz_2}{z_2}$$

and

$$W = \frac{1}{(2\pi j)^2} \oint_{|z_1|=1} \oint_{|z_2|=1} g^*(z_1, z_2) g(z_1, z_2) \frac{dz_1}{z_1} \frac{dz_2}{z_2}$$

respectively. Further denote the $N_1 \times N_1$ upper left blocks and the $N_2 \times N_2$ lower left blocks of K and W by K_{11}, W_{11} and K_{22}, W_{22}, respectively. System (25) is said to be (locally) balanced if

$$K_{11} = W_{11} = \text{diag}\,(\sigma_{11}, \sigma_{12}, \cdots, \sigma_{1N_1})$$

$$K_{22} = W_{22} = \text{diag}\,(\sigma_{21}, \sigma_{22}, \cdots, \sigma_{2N_2})$$

where σ_{1i} $(1 \le i \le N_1)$ and σ_{2i} $(1 \le i \le N_2)$ are called the Hankel singular values of the system. Note that $A_3 = 0$ means that the transfer function of the filter has a separable denominator. In addition, both A_1 and A_4 are nilpotent since $A_1^{N_1} = \mathbf{0}$ and $A_4^{N_2} = \mathbf{0}$. By using the Lyapunov approach [14], [18] in conjunction with the aforementioned properties, we obtain

$$K_{11} = \sum_{i=0}^{N_1-1} A_1^i P (A_1^T)^i \tag{26}$$

where

$$P = H_b H_b^T$$

with

$$H_b = \begin{bmatrix} h_{10} & h_{11} & \cdots & h_{1N_2} \\ h_{20} & h_{21} & \cdots & h_{2N_2} \\ \vdots & \vdots & \cdots & \vdots \\ h_{N_10} & h_{N_11} & \cdots & h_{N_1N_2} \end{bmatrix}.$$

In addition

$$K_{22} = I_{N_2}$$

$$W_{11} = I_{N_1}$$

and

$$W_{22} = \sum_{i=0}^{N_2-1} (A_4^T)^i Q A_4^i \tag{27}$$

where

$$Q = H_c^T H_c$$

with

$$H_c = \begin{bmatrix} h_{01} & h_{02} & \cdots & h_{0N_2} \\ h_{11} & h_{12} & \cdots & h_{1N_2} \\ \vdots & \vdots & \cdots & \vdots \\ h_{N_11} & h_{N_12} & \cdots & h_{N_1N_2} \end{bmatrix}.$$

Once K_{ii} and W_{ii} $(i = 1, 2)$ are found, a balancing transformation $T = T_1 \oplus T_2$ with T_1 and T_2 nonsingular can be computed such that the system realization $(\hat{A}, \hat{b}, \hat{c}, \hat{d}) = (T^{-1}AT, T^{-1}b, cT, d)$ is locally balanced. A reduced state-space model (A_r, b_r, c_r, d) of order (r_1, r_2) can be obtained by partitioning $\hat{A}$, $\hat{b}$, and $\hat{c}$ as

$$\hat{A} = \begin{bmatrix} \hat{A}_1 & \hat{A}_2 \\ \hat{A}_3 & \hat{A}_4 \end{bmatrix} = \begin{bmatrix} A_{1r} & A_{12} & A_{2r} & A_{22} \\ A_{13} & A_{14} & A_{23} & A_{24} \\ \mathbf{0} & \mathbf{0} & A_{4r} & A_{42} \\ \mathbf{0} & \mathbf{0} & A_{43} & A_{44} \end{bmatrix}$$

$$\hat{b} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \end{bmatrix} = \begin{bmatrix} b_{1r} \\ * \\ b_{2r} \\ * \end{bmatrix}$$

and

$$\hat{c} = [\hat{c}_1 \mid \hat{c}_2] = [c_{1r} \quad * \mid c_{2r} \quad *] \tag{28}$$

and then letting

$$A_r = \left[\begin{array}{c|c} A_{1r} & A_{2r} \\ \hline \mathbf{0} & A_{4r} \end{array}\right], \quad b_r = \left[\begin{array}{c} b_{1r} \\ \hline b_{2r} \end{array}\right] \quad \text{and}$$

$$c_r = [c_{1r} \mid c_{2r}].$$

B. Properties

The balanced approximation has been used extensively for the design of automatic controllers of 1-D linear systems owing to its good approximation accuracy. The maximum error introduced in 2-D separable discrete systems can be estimated as [23]

$$e_r \equiv \max_{|z_1|=1, |z_2|=1} |H_r(z_1, z_2) - H(z_1, z_2)| \le \alpha_1 \sum_{i=r_1+1}^{N_1} \sigma_{1i} + \alpha_2 \sum_{i=r_2+1}^{N_2} \sigma_{2i} \tag{29}$$

where

$$H_r(z_1, z_2) = c_r[I_r(z_1, z_2) - A_r]^{-1} b_r + d$$

σ_{1i} ($r_1 + 1 \le i \le N_1$) and σ_{2i} ($r_2 + 1 \le i \le N_2$) are the Hankel singular values of the original FIR filter, and α_1, α_2 are constants determined by $\{h_{ij}\}$. An important property of the BA method is that the stability of the reduced-order 2-D system is guaranteed provided that the original system is stable and has a separable denominator [23]. Since our design approach starts with an FIR filter, the above-mentioned properties apply. In what follows, we present two more properties of the BA method, which are especially desirable for digital filters.

Theorem 2: If $\varphi_r(\omega_1, \omega_2)$ and $\varphi(\omega_1, \omega_2)$ are the phase responses of the reduced and the original filter, respectively, and Ω_p denotes the passband region of the original filter, then

$$\max_{\Omega_p} |\varphi_r(\omega_1, \omega_2) - \varphi(\omega_1, \omega_2)| \approx \delta_m \le \frac{2e_r}{1 - \epsilon_p} \tag{30}$$

provided that $e_r << 1$, where e_r is defined in (29), ϵ_p is the maximum error over the passband region for the original FIR filter

$$\delta_m = \max_{\Omega_p} \left| 1 - \frac{X(\omega_1, \omega_2)|X_r(\omega_1, \omega_2)|}{X_r(\omega_1, \omega_2)|X(\omega_1, \omega_2)|} \right|.$$

$X(\omega_1, \omega_2)$ is the frequency response of the original filter given by (3), and

$$X_r(\omega_1, \omega_2) = H_r(e^{j\omega_1 T_1}, e^{j\omega_2 T_2}).$$

Proof: Writing

$$X(\omega_1, \omega_2) = M(\omega_1, \omega_2) e^{j\varphi(\omega_1 T_1, \omega_2 T_2)}$$
$$X_r(\omega_1, \omega_2) = M_r(\omega_1, \omega_2) e^{j\varphi_r(\omega_1 T_1, \omega_2 T_2)}$$

the approximation error in the phase angle can be expressed as

$$|\varphi_r(\omega_1, \omega_2) - \varphi(\omega_1, \omega_2)| = \left| \ln \left(\frac{X(\omega_1, \omega_2) M_r(\omega_1, \omega_2)}{X_r(\omega_1, \omega_2) M(\omega_1, \omega_2)} \right) \right| = |\ln (1 + \delta)|$$

where δ can be estimated as

$$|\delta| = \left| 1 - \frac{XM_r}{X_r M} \right| = \frac{|X_r M - X_r M_r + X_r M_r - XM_r|}{M_r M} \le \frac{|X_r(M - M_r)| + |M_r(X_r - X)|}{M_r M} = \frac{\| X| - |X_r \| + |X_r - X|}{M} \le \frac{2|X_r - X|}{M}.$$

Hence

$$\delta_m = \max_{\Omega_p} |\delta| \le \frac{N_{\max}}{D_{\min}} \le \frac{2e_r}{1 - \epsilon_p} \tag{31}$$

where

$$N_{\max} = 2 \max_{\Omega_p} |X_r - X|$$

and

$$D_{\min} = \min_{\Omega_p} M.$$

Since for a good FIR filter $\epsilon_p \approx 0$, (31) indicates that $|\delta| << 1$ if $e_r << 1$. Under these circumstances, we have

$$\max_{\Omega_p} |\varphi_r(\omega_1, \omega_2) - \varphi(\omega_1, \omega_2)| = \max_{\Omega_p} |\ln (1 + \delta)| \approx \max_{\Omega_p} |\delta| \le \frac{2e_r}{1 - \epsilon_p}. \qquad \square$$

If the BA method is used to approximate a high-order FIR filter, then by (29) the approximation error can be made very small if both

$$\sum_{i=r_1+1}^{N_1} \sigma_{1i} \quad \text{and} \quad \sum_{i=r_2+1}^{N_2} \sigma_{2i}$$

are very small. In such a case, the phase-response linearity of the original filter will be well preserved in the reduced filter by virtue of Theorem 2.

It is known that a locally balanced 2-D state-space digital filter has the lowest output roundoff noise and further, its sensitivity to coefficient quantization is low [18], [26], [27]. The second property, stated in terms of Theorem 3 below, ensures that, like the original filter, the reduced filter has low roundoff noise and low sensitivity to coefficient quantization since it is nearly balanced.

Theorem 3: If $\Sigma_{1r_1} = \text{diag}(\sigma_{11}, \cdots, \sigma_{1r_1})$, $\Sigma_{1c} = \text{diag}(\sigma_{1r_1+1}, \cdots, \sigma_{1N_1})$, $\Sigma_{2r_2} = \text{diag}(\sigma_{21}, \cdots, \sigma_{2r_2})$, $\Sigma_{2c} = \text{diag}(\sigma_{2r_2+1}, \cdots, \sigma_{2N_2})$, then

$$\begin{aligned} A_{1r}\Sigma_{1r_1}A_{1r}^T - \Sigma_{1r_1} &\approx -(b_{1r}b_{1r}^T + A_{2r}\Sigma_{2r_2}A_{2r}^T) \\ A_{1r}^T\Sigma_{1r_1}A_{1r} - \Sigma_{1r_1} &\approx -c_{1r}^T c_{1r} \\ A_{4r}\Sigma_{2r_2}A_{4r}^T - \Sigma_{2r_2} &\approx -b_{2r}b_{2r}^T \\ A_{4r}^T\Sigma_{2r_2}A_{4r} - \Sigma_{2r_2} &\approx -(c_{2r}^T c_{2r} + A_{2r}^T\Sigma_{1r_1}A_{2r}) \end{aligned} \tag{32}$$

provided that

$$\sigma_{1r_1+1} << \sigma_{1r_1} \quad \text{and} \quad \sigma_{2r_2+1} << \sigma_{2r_2}. \tag{33}$$

Proof: Since $(\hat{A}, \hat{b}, \hat{c}, \hat{d})$ given by (28) represents a balanced realization of the original system, it satisfies [18] the relations

$$\hat{A}_1\Sigma_1\hat{A}_1^T - \Sigma_1 = -(\hat{b}_1\hat{b}_1^T + \hat{A}_2\Sigma_2\hat{A}_2^T) \tag{34}$$

$$\hat{A}_1^T\Sigma_1\hat{A}_1 - \Sigma_1 = -\hat{c}_1^T\hat{c}_1 \tag{35}$$

$$\hat{A}_4\Sigma_2\hat{A}_4^T - \Sigma_2 = -\hat{b}_2\hat{b}_2^T \tag{36}$$

$$\hat{A}_4^T\Sigma_2\hat{A}_4 - \Sigma_2 = -(\hat{c}_2^T\hat{c}_2 + \hat{A}_2^T\Sigma_1\hat{A}_2) \tag{37}$$

where

$$\Sigma_1 = \begin{bmatrix}\Sigma_{1r_1} & 0\\ 0 & \Sigma_{1c}\end{bmatrix} \text{ and } \Sigma_2 = \begin{bmatrix}\Sigma_{2r_2} & 0\\ 0 & \Sigma_{2c}\end{bmatrix}.$$

By taking the upper left $r_1 \times r_1$ submatrices from (34) and (35), and the upper left $r_2 \times r_2$ submatrices from (36) and (37), we obtain

$$\begin{aligned}
A_{1r}\Sigma_{1r_1}A_{1r}^T - \Sigma_{1r_1} &= -(b_{1r}b_{1r}^T + A_{2r}\Sigma_{2r_2}A_{2r}^T) - E_{1b}\\
A_{1r}^T\Sigma_{1r_1}A_{1r} - \Sigma_{1r_1} &= -c_{1r}^Tc_{1r} - E_{1c}\\
A_{4r}\Sigma_{2r_2}A_{4r}^T - \Sigma_{2r_2} &= -b_{2r}b_{2r}^T - E_{2b}\\
A_{4r}^T\Sigma_{2r_2}A_{4r} - \Sigma_{2r_2} &= -(c_{2r}^Tc_{2r} + A_{2r}^T\Sigma_{1r_1}A_{2r}) - E_{2c}.
\end{aligned} \tag{38}$$

Since the balanced realization $(\hat{A}, \hat{b}, \hat{c}, \hat{d})$ is obtained from a canonic realization of the FIR filter which is locally reachable and observable [28], it is also locally reachable and observable. Moreover it can be shown the $\|\hat{A}_1\| \leq 1$, $\|\hat{A}_4\| \leq 1$ and $\|\hat{A}_2\| \leq 1$ (see [23, theorem 5.2.3 and appendix 5C]) which imply that $\|A_{12}\| \leq 1$, $\|A_{42}\| \leq 1$ and $\|A_{22}\| \leq 1$, respectively. Therefore, matrices E_{ib} and E_{ic} $(i = 1, 2)$ can be estimated as

$$\begin{aligned}
\|E_{1b}\| &= \|A_{12}\Sigma_{1c}A_{12}^T + A_{22}\Sigma_{2c}A_{22}^T\| \leq \|A_{12}\|^2\|\Sigma_{1c}\|\\
&\quad + \|A_{22}\|^2\|\Sigma_{2c}\| \leq \sigma_{1r_1+1} + \sigma_{2r_2+1}\\
\|E_{1c}\| &= \|A_{12}^T\Sigma_{1c}A_{12}\| \leq \|A_{12}\|^2\|\Sigma_{1c}\| \leq \sigma_{1r_1+1}\\
\|E_{2b}\| &= \|A_{42}\Sigma_{2c}A_{42}^T\| \leq \|A_{42}\|^2\|\Sigma_{2c}\| \leq \sigma_{2r_2+1}\\
\|E_{2c}\| &= \|A_{42}^T\Sigma_{2c}A_{42} + A_{22}^T\Sigma_{1c}A_{22}\|\\
&\leq \sigma_{1r_1+1} + \sigma_{2r_2+1}.
\end{aligned}$$

Consequently, matrices E_{ib} and E_{ic} $(i = 1, 2)$ can be neglected from (38) if the conditions stated in (33) are satisfied. This modifies (38) to (32). □

C. Example 2

In order to demonstrate the effectiveness of the BA method, it was applied to the linear-phase 2-D FIR digital filter designed for $K = 25$ in Section II-C. Following the steps of the algorithm described above, the balanced realization was first obtained. Then by examining the two sets of singular values and neglecting the insignificant ones, reduced IIR realizations of orders $(r_1, r_2) = (10, 12)$ and $(13, 15)$ were obtained. In the first realization, all σ_{1i} and σ_{2i} that are less than 0.4% of σ_{11} and σ_{21}, respectively, were ignored and in the second realization all σ_{1i} and σ_{2i} that are less than 0.02% of σ_{11} and σ_{21}, respectively, were ignored. For the sake of comparison, the BA method was also applied to a (31, 31) FIR filter to obtain an IIR filter of order (15, 17). The maximum passband and stopband errors for the three designs are given in Table II. The amplitude response of the reduced filter for the case $(r_1, r_2) = (13, 15)$ is depicted in Fig. 2(c).

TABLE II
MAXIMUM PASSBAND AND STOPBAND ERRORS FOR REDUCED REALIZATION

(r_1, r_2)	Passband	Stopband
(10, 12)	0.0915	0.0716
(13, 15)	0.0230	0.0202
(15, 17)	0.0102	0.0087

The group delays of the 2-D IIR filter, namely

$$\tau_1 = -\frac{\partial\{\arg[H(e^{j\omega_1T_1}, e^{j\omega_2T_2})]\}}{\partial\omega_1}$$

and

$$\tau_2 = -\frac{\partial\{\arg[H(e^{j\omega_1T_1}, e^{j\omega_2T_2})]\}}{\partial\omega_2}$$

were used to construct the contour plots of Fig. 3(a) and 3(b), respectively, for the case $(r_1, r_2) = (13, 15)$, assuming fifty layers. It is clearly seen that the linear-phase characteristic of the original FIR filter is well preserved over the entire passband of the 2-D filter. The maximum relative errors in the group delays are given in Table III.

We conclude this section with a brief discussion on the efficiency of the filters obtained by the BA method. We observe in Tables II and III that if the errors in the amplitude response and the group delays are required to be less than 3.0% and 1.5%, respectively, an IIR filter of order (13, 15) is required. The same amplitude-response accuracy can be achieved by using an FIR filter of order (29, 29) comprising 15 parallel sections, as can be seen in Table I. The number of multiplications needed for a direct realization of the FIR filter is $(15 + 15) \times 15 = 450$ as compared to $(13 \times 15) + (13 + 15) = 223$ multiplications required by a direct realization of the reduced IIR filter. If the FIR filter is implemented using 1-D fast convolution with block size 512×512, then the total multiplications required to process an array of size 512×512 is 166×512^2 [3]. On the other hand, by writing the transfer function of the (13×15) IIR filter as

$$H_r(z_1, z_2) = \frac{N_r(z_1, z_2)}{D_{r1}(z_1)D_{r2}(z_2)}$$

the filter can be implemented using the scheme of Fig. 4. Through the SVD of the 13×15 coefficient matrix of $N_r(z_1, z_2)$, implementation can be achieved by using at most 13 parallel 1-D sections. Consequently, the application of 1-D fast convolution to the FIR filters in both feedforward and feedback paths leads to an implementation of the IIR filter which requires a total of 156×512^2 multiplications in the worst case. Experience with a variety of practically useful 2-D filters has shown that some of the singular values of the coefficient matrix of $N_r(z_1, z_2)$ are often negligible (less than 10^{-4}) and, therefore,

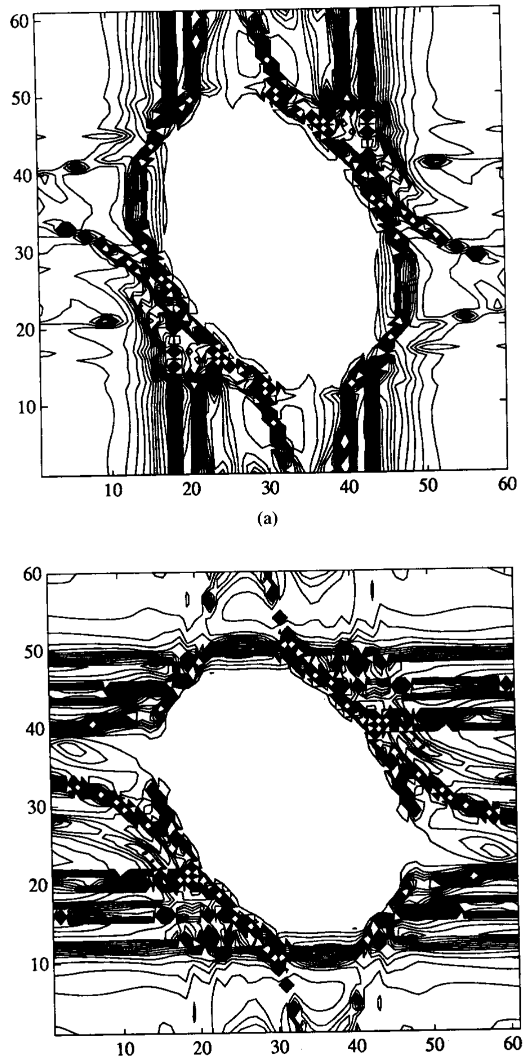


Fig. 3. Contour plots of group delays of IIR filter: (a) group delay with respect to ω_1; (b) group delay with respect to ω_2.

TABLE III
MAXIMUM RELATIVE ERRORS IN GROUP DELAYS FOR REDUCED REALIZATION

(r_1, r_2)	τ_1	τ_2
(10, 12)	6.30%	5.52%
(13, 15)	1.48%	1.49%
(15, 17)	0.79%	0.93%

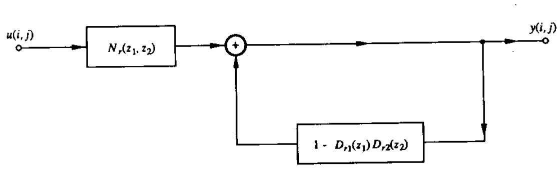


Fig. 4. A realization of $H_r(z_1, z_2)$.

the number of multiplications required is usually less than that of the worst case.

Now if the errors in the amplitude response and the group delays are required to be less than 1.2% and 1.0%, respectively, an IIR filter of order (15, 17) is required according to Tables II and III. The same amplitude-response accuracy can be achieved by using an FIR filter of order (29, 29) comprising 25 parallel sections as seen in Table I. In this case, the implementation of the FIR filter by 1-D fast convolution requires 266×512^2 multiplications while its IIR counterpart requires only 176×512^2 multiplications.

At the other extreme, if the error in the amplitude response is allowed to be as high as 9%, an IIR filter of order (10, 12) is required. The same accuracy can be achieved by using an FIR filter of order (29, 29) comprising only five parallel sections. In this case, the FIR implementation is more efficient than its IIR counterpart.

In effect, if the amplitude-response approximation error is required to be low, say less than 5%, and a small variation in the group delays of the order of 1 to 2% can be tolerated, the BA method is very likely to yield a more efficient IIR design. Furthermore, a nearly balanced state-space realization is achieved which has low roundoff output noise as well as low sensitivities to coefficient quantization, according to Theorem 3.

IV. CONCLUSIONS

In the first half of the paper, the SVD has been applied in conjunction with 1-D FIR filter techniques for the design of 2-D, causal, linear-phase FIR filters with arbitrary amplitude responses. The method is relatively simple to apply and leads to a parallel arrangement of pairs of cascade sections. This configuration is, therefore, suitable for parallel processing and is amenable to VLSI implementation.

In the second half of the paper, the BA method has been applied for the design of 2-D digital filters. In this approach, the design starts with a 2-D causal, linear-phase, FIR filter of the type that can be obtained using the SVD and concludes with a design of reduced order, which is almost always an IIR design. The method leads to computationally efficient designs and tends to preserve the linear phase response of the original FIR filter irrespective of whether an FIR or an IIR design is obtained. Furthermore, the designs obtained are causal and locally quasi balanced, and in cases where IIR designs are obtained, stability is guaranteed. In other words, the method is highly suitable for the design of 2-D, causal, linear-phase IIR filters, which are very difficult to design by other methods.

The design approach has two important advantages. First, it leads to more efficient implementations in applications where the amplitude-response approximation error is required to be low. Second, nearly balanced state-space realizations are achieved which have low roundoff output noise as well as low sensitivity to coefficient quantization.

References

[1] T. S. Huang, J. W. Burnett, and A. G. Deczky, "The importance of phase in image processing filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, pp. 529–542, Dec. 1975.

[2] T. S. Huang, "Two-dimensional windows," *IEEE Trans. Audio Electroacoust.*, vol. AU-20, pp. 88–89, Mar. 1972.

[3] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[4] T. C. Speake and R. M. Mersereau, "A note on the use of windows for two-dimensional FIR filter design," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 125–127, Feb. 1981.

[5] J. H. McClellan and T. W. Parks, "Equiripple approximation of fan filters," *Geophysics*, vol. 7, pp. 573–583, 1972.

[6] J. H. McClellan, "The design of two-dimensional filters by transformations," in *Proc. 7th Annu. Princeton Conf. Inform. Sci. Syst.*, 1973, pp. 247–251.

[7] R. M. Mersereau, W. F. G. Mecklenbrauker, and T. F. Quatieri, Jr., "McClellan transformations for two-dimensional digital filtering: I—Design," *IEEE Trans. Circuits Syst.*, vol. CAS-23, pp. 405–413, July 1976.

[8] W. F. G. Mecklenbrauker and R. M. Mersereau, "McClellan transformations for two-dimensional digital filtering: II—Implementation," *IEEE Trans. Circuits Syst.*, vol. CAS-23, pp. 414–422, July 1976.

[9] A. Antoniou and W.-S. Lu, "Design of two-dimensional digital filters by using the singular value decomposition," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 1191–1198, Oct. 1987.

[10] W.-S. Lu, H.-P. Wang, and A. Antoniou, "Design of two-dimensional FIR digital filters by using the singular value decomposition," in *Proc. 1988 ComCon Conf.* (Baton Rouge, LA), Oct. 1988, pp. 546–556.

[11] B. C. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 17–32, Feb. 1981.

[12] H. Kimurd and Y. Honoki, "Balanced approximation of digital FIR filter with linear phase characteristic," in *IEEE Proc. Int. Symp. Circuits Syst.*, 1985, pp. 283–286.

[13] W.-S. Lu, E. B. Lee, and Q.-T. Zhang, "Model reduction for 2-D systems," in *Proc. 1986 IEEE Int. Symp. Circuits Syst.*, May 1986, pp. 79–82.

[14] W.-S. Lu, E. B. Lee, and Q.-T. Zhang, "Balanced approximation of two-dimensional and delay-differential systems," in *Proc. 25th Conf. Decision Contr.* (Athens, Greece), Dec. 1986, pp. 917–922.

[15] A. Antoniou, *Digital Filters: Analysis and Design*. New York: McGraw-Hill, 1979.

[16] T. W. Parks and J. H. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," *IEEE Trans. Circuits Theory*, vol. CT-19, pp. 189–194, March 1972.

[17] B. Lashgari, L. M. Silverman, and J. F. Abramatic, "Approximation of 2-D separable in denominator filters," *IEEE Trans. Circuits Syst.*, vol. CAS-30, pp. 107–121, Feb. 1983.

[18] M. Kawamata and T. Higuchi, "Synthesis of 2-D separable denominator digital filters with minimum roundoff noise and no overflow oscillations," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 365–372, Apr. 1986.

[19] E. I. Jury and K. Premaratne, "Model reduction of two-dimensional discrete systems," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 558–562, May 1986.

[20] K. Glover, R. F. Curtain, and J. R. Partington, "Realization and approximation of linear infinite dimensional systems with error bounds," Univ. of Cambridge, Dep. Eng., Rep. CUED/F-CAMS/TR258, 1986.

[21] A. Kumar, F. W. Fairman, and J. R. Sveinsson, "Separately balanced realization and model reduction of 2-D separable denominator transfer functions from input data," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 233–239, Mar. 1987.

[22] K. Premaratne and E. I. Jury, "Model reduction of two-dimensional discrete systems via balanced realization," in *Proc. 21st Asilomar Conf. Signal, Syst., Comput.*, Nov. 1987.

[23] K. Premaratne, "Model reduction of two-dimensional discrete time and delay systems," Ph.D. dissertation, Univ. of Miami, Aug. 1988.

[24] S. Y. Kung, B. C. Levy, M. Morf, and T. Kailath, "New results in 2-D system theory, Part II: 2-D state-space model—realization and notions of controllability, observability, and minimality," *Proc. IEEE*, vol. 65, pp. 945–961, 1977.

[25] T. Hinamoto and F. W. Fairman, "Separable denominator state-space realization of two-dimensional filters using a canonic form," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 846–853, Aug. 1981.

[26] W.-S. Lu and A. Antoniou, "Synthesis of 2-D state-space fixed point digital-filter structures with minimum roundoff noise," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 965–973, Oct. 1986.

[27] A. Zilouchian and R. L. Carroll, "A coefficient sensitivity bound in 2-D state-space digital filtering," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 665–667, July 1986.

[28] R. P. Roesser, "A discrete state-space model for linear image processing," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 1–10, Feb. 1975.

Wu-Sheng Lu (S'81–M'85–SM'90) received the B.S. degree in mathematics from Fudan University, China, in 1964 and the Ph.D. degree in control science from the University of Minnesota, Minneapolis, in 1984.

He was a postdoctoral fellow at the University of Victoria, Victoria, BC, Canada, in 1985 and held a visiting assistant professorship at the University of Minnesota in 1986. Since 1987, he has been with the University of Victoria where he is currently Professor. His teaching and research interests are in the areas of multidimensional digital signal processing and robotics. He was Associate Editor of the *Canadian Journal of Electrical and Computer Engineering* in 1989 and has been Editor of the same journal since 1990.

Hui-Ping Wang received the B.Eng. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 1984, and the M.A.Sc. degree in 1987 from the University of Victoria, Victoria, BC, Canada. She is currently working toward the Ph.D. degree in electrical engineering at the University of Victoria.

She has been working as a Research and Teaching Assistant, and has received University of Victoria fellowships during the period 1986–1989, the Charles S. Humphrey Award for 1987 and 1988, and a British Columbia Advanced Systems Foundation Graduate Scholarship Award for 1989. Currently, she is with Hydrographic Division, Terra Surveys, Sidney, BC, working on signal processing techniques for bathymetric data.

Andreas Antoniou (M'69–SM'79–F'82) received the B.Sc. (Eng.) and Ph.D. degrees in electrical engineering from London University in 1963 and 1966, respectively.

From 1966 to 1969 he was Senior Scientific Officer at the Post Office Research Department, London, and from 1969 to 1970, he was a Member of the Scientific Staff at the R & D Laboratories of Northern Electric Company Ltd., Ottawa, Ont., Canada. From 1970 to 1983 he served in the Department of Electrical Engineering, Concordia University, Montreal, Que., Canada, as Professor from June 1973 and as Chairman since December 1977. He served as founding Chairman of the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada, during the period July 1, 1983 to June 30, 1990 and is now Professor in the same department. His teaching and research interests are in the areas of electronics, network synthesis, digital system design, active and digital filters, and digital signal processing. He is the author of *Digital Filters: Analysis and Design* (McGraw-Hill). One of his papers on gyrator circuits was awarded the Ambrose Fleming Premium by the Institution of Electrical Engineers, U.K.

Dr. Antoniou is a Member of the Association of Professonal Engineers of British Columbia, and a Fellow of the Institution of Electrical Engineers. He was elected a Fellow of the IEEE for contributions to active and digital filters, and to electrical engineering education. He served as Associate Editor of the IEEE Transactions on Circuits and Systems during the period June 1983 to May 1985 and as Editor during the period June 1985 to May 1987.