PARALLEL AND SEPARABLE 2-D FIR DIGITAL FILTER DESIGN

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering

in

The Department of Electrical and Computer Engineering

by
Vidya Venkatachalam
B.E., Anna University, India, 1992
May, 1995

Acknowledgments

I express my most sincere thanks to my advisor, Dr. Jorge L. Aravena, for his time, considerable patience, encouragement and continual support, while guiding me throughout the period of this work. I am especially grateful to him for being a constant source of inspiration, both professionally and personally. I thank Dr. Guoxiang Gu, and Dr. Kemin Zhou, for consenting to be on my defense committee, and also for the knowledge they have imparted to me through various courses, without which this work would not have been possible. I am appreciative of the valuable suggestions given by Dr. Wu-Sheng Lu, which form the basis of the Section 4.1.1. I also thank my mother and my sister for their support at all times, and my friends Karthik, Aruna, and Shobhana, among others, for making my stay here very enjoyable.

Table of Contents

List of Tables

List of Figures

List of symbols

D	Desired 2-D frequency response matrix
w_{ik_i}	Frequency sample points, defined as $\frac{\pi(2k_i+1)}{2M_i}$, where M_i is the total number of such points
h	Designed nonseparable filter coefficient matrix
a, b	Designed separable filter coefficient vectors
x	2-D input image matrix
$D \bullet X$	Hadamard product
y	2-D output image matrix after filtering
$A \in E^{M_1 \times M_2}$	2-D complex matrix A of size M_1 by M_2 with the Euclidean metric
$A \otimes B$	Kronecker product of A and B
A^*	The conjugate transpose of the matrix A
A^T	The transpose of the matrix A
A^c	The conjugate of the matrix A
$AB_{i,j}$	$AB(i,j)$
X_v	1-D Vector obtained by stacking the columns of the matrix X from left to right
$D(\bullet)$	Forming a diagonal matrix from the vector $\bullet$
$D'(\bullet)$	Forming a vector from the diagonal matrix $\bullet$
X_{k_i}	The $i\ th$ element in the k component (column/row) of X

$tr\{A\}$	Sum of the elements of $diag(A)$
$diag(X)$	The main diagonal of the 2-D matrix X
$norm[X]$	Frobenius norm, defined as $\sqrt{tr\{(X^*X)\}}$
$\langle x, y \rangle$	The inner product of x and y, defined as $\sum_i x_i y_i$
$J(\bullet, \bullet)$	Cost function which is dependent on the variables $\bullet$ and $\bullet$
Ω_i	Discrete frequency matrix, with each element (k_1, k_2) defined as $e^{-jw_{ik_1}k_2}$
$\mathcal{W}$	Weight map, representing the weighting function
$\mathcal{F}(\bullet)$	Frequency map, used to obtain the discrete frequency response from $\bullet$
$\mathcal{S}$	Stacking map, representing the matrix to vector conversion
$\mathcal{T}$	Permutation map, representing a rearrangement of elements
$\delta J_\bullet$	Partial derivative of the function J with respect to the variable $\bullet$
$\Re(\bullet)$	The matrix of the real part of each element of $\bullet$
$max(\bullet)$	Maximum of $\bullet$
$\parallel a \parallel$	The Euclidean norm of the vector a, defined by $\sqrt{a^T a}$
$X(m_1 : m_2, n_1 : n_2)$	The matrix comprising the $m_1, \ldots, m_2$ rows, and the $n_1, \ldots, n_2$ columns of the matrix X
$X(:, 1 : n)$	The matrix comprising all the rows, and the first n columns of the matrix X
$X(1 : m, :)$	The matrix comprising the first m rows, and all the columns of the matrix X
$\mathcal{O}(i)$	Of the order of i, defined as $\lim_{i \to \infty} \frac{\mathcal{O}(i)}{i}$ = constant

Abstract

This thesis presents the results of a research which develops a technique to design $2-D$ filters by approximating an ideal frequency response with sums of separable FIR components. The technique is independent of the nature of the ideal response, and can accommodate the inclusion of a weighting function. This approach gives the designer flexibility in selecting the $1-D$ filter orders and the number of separable filters to be used for best results. The problem is solved for the weighted least mean squares case, and a rigorous mathematical analysis is used to formulate the separable design algorithm. This work includes a brief analysis of the computational complexity of the formulated technique, and simulation results demonstrating the effectiveness of the design algorithm. It also offers suggestions for further research in this area.

Chapter 1

Introduction

Computational speed is of prime importance in image processing. Processing here refers to the filtering of the image in order to reduce noise and enhance the image. The design of fast-acting $2-D$ FIR digital filters is thus a much researched area in Digital Signal Processing. Several papers have been written on the subject ([2]-[4], [23]-[30]), which propose new algorithms to achieve good quality designs with reduced computational complexity. A *good quality design* is obtained by finding the *optimal* filter coefficients that satisfy a given constraint. *Reduced computational complexity* is obtained by eliminating redundant operations, making acceptable approximations and by putting to use the inherent symmetry properties in the desired filter response. Unfortunately, these two requirements are conflicting in nature, and there is always a trade-off between them in any standard design technique.

1.1 Historical Background

Consider a linear shift invariant ideal filter discrete frequency response $D(w_{1k_1}, w_{2k_2})$ of size $M_1 \times M_2$. The filtering operation on the image is then defined by the model $Y = D \bullet X$ where X and Y are the Discrete Fourier Transforms (DFT) of the input and output images respectively, and the Hadamard product $D \bullet X$ represents the element by element multiplication of the matrices D and X. By convention, one uses discretized frequencies

$$w_{ik_i} = \frac{\pi(2k_i + 1)}{2M_i}; \quad k_i = -M_i, -M_i + 1, \ldots, M_i - 1; i = 1, 2 \tag{1.1}$$

The design of $2 - D$ digital filters involves the following steps:

i) Approximation

ii) Realization

iii) Implementation

The filtering application imposes certain constraints on the frequency response of the filter which need to be satisfied. Based on the specifications prescribed, one must first make a choice of the most appropriate type (usually whether recursive (IIR) or nonrecursive (FIR)) of filter to be used. The *approximation* problem must then be solved. This is the process of finding a stable transfer function with real coefficients such that the required specifications regarding the frequency response of the filter, are satisfied. This is the most important aspect of filter design, and in most cases, the solution to the approximation problem is considered the design of the filter.

The *realization* problem is one of converting the transfer function of the filter obtained in the first step, into a digital filter network. The *implementation* problem is application dependent. Real time processing applications require a hardware implementation, while for applications where speed is not of primary concern, a software implementation is more appropriate.

The design of FIR and IIR filters are two distinct problems due to the vastly different properties of these two filter types. Since this work addresses the design of

FIR filters, the definition of an FIR filter and some of its important properties are listed below.

A $2-D$ FIR single-input single-output digital filter with support in the rectangle defined by $-N_i \leq n_i \leq N_i,\ i = 1,2$ is characterized by the transfer function

$$H(z_1, z_2) = \sum_{n_1=-N_1}^{N_1} \sum_{n_2=-N_2}^{N_2} h(n_1, n_2) z_1^{-n_1} z_2^{-n_2} \tag{1.2}$$

where $h(n_1, n_2)$ represents the impulse response of the filter.

This filter has the following very important properties:

i) It is always stable.

ii) A linear phase response with respect to w_1 and w_2 can be easily achieved. This property is particularly important for image processing applications ([13]).

The subsequent sections present a review of the different approaches adopted in $2-D$ FIR digital filter design.

1.1.1 Conventional approaches

The simplest approach to design the $2-D$ FIR digital filter is based on the application of the Fourier series. Since the frequency response of a nonrecursive filter is a periodic function of frequency, it can be expressed in terms of the Fourier series. In this method, the Fourier series representation is used in conjunction with a special class of functions called windowing functions to give reasonably good results ([14]-[17]). The operation consists of taking the $2-D$ Discrete Fourier Transform (DFT) of the ideal response D, and convolving it with the windowing function chosen. The designed filter is then given by the inverse DFT on the convolved result. This entire exercise, though undoubtedly simple, is slow and places a great strain on the hardware due to excessive memory requirements owing to the large filter size. Also, it gives sub-optimal results. It has been observed that certain filter types, notably quadrantally symmetric and antisymmetric filters, always have real coefficients. The Discrete Cosine Transform (DCT) ([18],[19]) and the Discrete Sine Transform (DST) ([20],[21]) have been developed, and they operate exclusively on the symmetric and antisymmetric responses respectively, with a significant reduction in computational complexity. Now, any $2-D$ FIR filter D can be separated into two components, D_c and D_s, where D_c is the symmetric part and D_s is the antisymmetric part. The DCT and the DST operate on D_c and D_s respectively, replacing the original DFT operation on D. This results in a reduction in overall complexity. Particularly, for filters having some special symmetry properties, one can achieve very efficient designs by the use of these transforms.

Linear phase quadrantally symmetric filters have special properties which allow for approximations which are easier to work with and which give good quality designs. This is the motivation behind filter design by use of suitable transformations. The most widely used design technique based on transformations is the one based on the the McClellan Transformation ([22],[23]) which gives very efficient designs. Here, the zero-phase $2-D$ FIR filter is first designed, and then the McClellan Transformation is applied to the transfer function obtained. The McClellan Transformation attempts to approximate the cosine terms of the designed zero phase response with a series of terms, leading to linear phase, quadrantally symmetric, FIR filter designs. There are other design techniques which aim to produce optimal designs (as defined by the nature of the application). One example of this is the Chebyshev design algorithm ([24]-[27]) where the aim is to achieve an equiripple approximation to a desired frequency response. The minimax design method ([28]-[30]) is yet another example where optimization techniques are used. Here, the design of the nonrecursive filter is transformed into an unconstrained minimax optimization problem by defining the objective function as the error function which is to be minimized in the minimax sense. All the above are some of the most widely used design procedures in conventional FIR filter design.

1.1.2 Alternate approaches

With the advances in VLSI technology and the advent of high speed processors which allow a high degree of parallelism, there is new interest ([6]-[11]) in digital filter design algorithms which readily lend themselves to a parallel architecture. Such algorithms provide for a fast implementation *without deterioration in filter quality* by allowing for several operations to be performed concurrently, thus reducing the trade-off inherent in the standard design techniques.

The approach taken here is to *approximate a desired filter with sums of simpler and faster filters*. The $2-D$ filtering action is now accomplished by several pairs of $1-D$ filtering actions, all acting *concurrently* on the image, with each $1-D$ filter in a pair acting either in the w_1 or the w_2 directions. The transfer function of the $k\ th$ such separable FIR filter is then given by

$$H(z_1, z_2) = \sum_{n_1=-N_1}^{N_1} \sum_{n_2=-N_2}^{N_2} a_k(n_1) b_k(n_2) z_1^{-n_1} z_2^{-n_2} \tag{1.3}$$

[6],[7] give the details of one approach to design such filter pairs. This approach entails the use of the *Singular Value Decomposition* (SVD) of the desired response to find the optimal separable responses. These responses are then approximated by $1-D$ FIR filters, using standard design algorithms (like for example, the Remez algorithm ([31])).

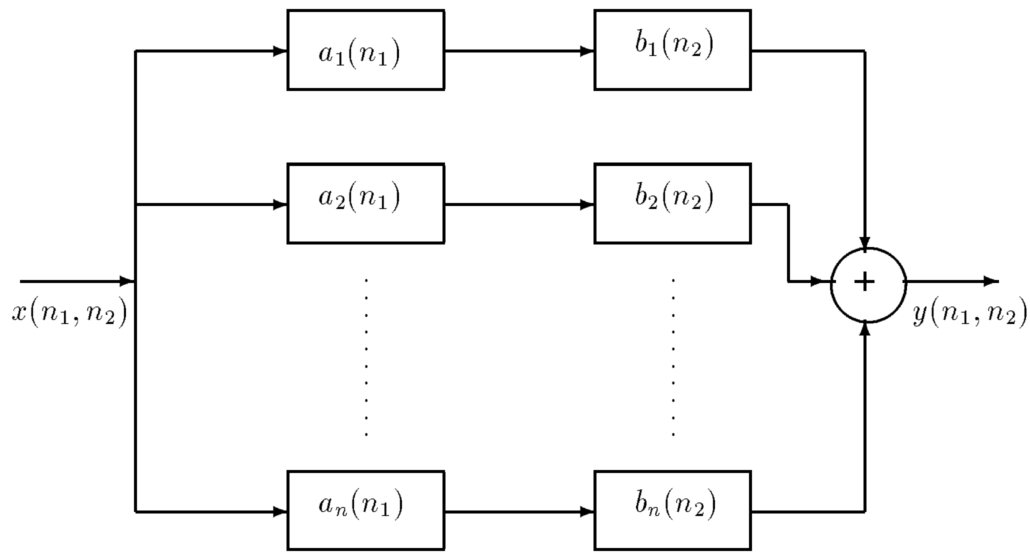

Figure 1.1: Implementation of filtering action using parallel, separable FIR filters

Once the $1-D$ filter pairs have been designed, the filtering action is implemented in real time as shown in Figure 1.1 ([5]). In the figure, $x(n_1, n_2)$ represents the pixels of the input image, and $y(n_1, n_2)$ the output image pixels after the filtering action. The $\{a_k\}$ represent the $1-D$ FIR filter coefficients which act on all the columns of the image concurrently. The $\{b_k\}$ represent the $1-D$ FIR filters which act concurrently on all the rows of the partially filtered image to produce the final output image. Thus, there are n pairs of separable filters, all acting in parallel on the image. This architecture is most suitable for shared memory machines and provides enormous speed-up in real-time image processing.

The main drawback of the SVD design procedure is that it involves a further *approximation* by FIR filters, once the optimal $1-D$ frequency responses are determined. This leads to a final design which is suboptimal. There is therefore a need to formulate an algorithm to design truly optimal $1-D$ FIR filters. The following section provides a synopsis of the research done in this direction, and its implications.

1.2 Synopsis of work

This work develops a parallel separable structure for $2-D$ digital filtering and studies the effect of the type of the desired response on the extent of simplification achievable. The $1-D$ filters are constrained to be FIR and a complete mathematical analysis is done to provide a formal algorithm to obtain the optimal $1-D$ filter coefficients, without using any approximation involved in the SVD design method. These optimal coefficients are obtained as the minimizing solutions of a least mean square merit index. The problem formulation is general in that it allows for the inclusion of a weighting function to permit the presence of a transition band in the filter response.

The most significant aspect of this algorithm is that it seeks to approximate the ideal filter, with a sum of separable FIR filters, *where the length of each* $1-D$ *filter can be selected independently.* This form permits fast implementation in a shared memory architecture with a less complicated implementation owing to its separable nature, and offers the possibility of very high throughput, while remaining very flexible in its strucuture.

The succeeding chapters describe the development of this separable design algorithm.

Chapter 2 discusses the results leading to the feasibility of obtaining separable approximations, and studies the effect of the nature of the desired $2-D$ frequency response on the separable filters.

Chapter 3 describes the mathematical development of the algorithm to determine the optimal 1-term separable filter.

Chapter 4 extends this algorithm to multiple terms, and outlines the modifications to be made when designing filters with certain special properties.

Chapter 5 illustrates the effectiveness of the algorithm with suitable examples, and gives an analysis of the computational complexity of this design process.

Chapter 6 provides a discussion of the results and offers suggestions for future research.

Chapter 2

Preliminary Studies

It was observed in Chapter 1 that computational speed-up can be achieved if some form of parallelism can be incorporated in the processing of the image. The attempt, in this study, is to outline the basis for incorporating parallelism in the filtering action, and to carefully study the properties of the commonly used filters and determine what influences the extent to which they would lend themselves to a parallel structure implementation. Essentially, the $2-D$ filter is separated out into several $1-D$ component filters in the frequency domain, using the approach given in [6],[7]. These component filters are then analyzed in detail to determine their properties, which are then experimentally verified using some commonly used filter configurations.

2.1 Obtaining the $1-D$ component filters

This section presents a review of the mathematical basis behind the feasibility of the separable design approach. The development uses the matrix Kronecker product which is defined below.

If $A \in E^{M_1 \times M_2}$ and $B \in E^{M_3 \times M_4}$, the Kronecker product of A and B is defined by,

$$A \otimes B = [AB_{i,j}] \in E^{M_1 M_3 \times M_2 M_4} \tag{2.1}$$

Consider the $M_1 \times M_2$ digital image filter matrix D, representing the ideal frequency response. One can form the diagonal matrix T from D by stacking the columns of

D from left to right along the diagonal of T. T is thus an $M_1M_2 \times M_1M_2$ diagonal matrix.

It is known that T admits the parallel Kronecker decomposition ([1]) :

$$T = \sum_{k=1}^{q} L_k \otimes R_k; \quad q \leq M_2 \tag{2.2}$$

One convenient definition for the L_k and R_k matrices is H_k and E_k, where H_k is the $k\ th$ $M_1 \times M_1$ diagonal matrix of T and E_k is an $M_2 \times M_2$ matrix of the form

$$E_k(i,j) = \begin{cases} 1 & \text{if i=j=k} \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

Now, q is determined by the number of $H_k^{\prime}s$ that form an independent set. The maximum possible number of such independent decompositions of T is $n = min(M_1, M_2)$. Then, from eqs. 2.2 and 2.3, we obtain

$$T = \sum_{k=1}^{q} H_k \otimes E_k; \quad q \leq n \tag{2.4}$$

Let X be the $M_1 \times M_2$ matrix of the Discrete Fourier Transform of the digital image. The $M_1M_2 \times 1$ vector X_v is formed from X by stacking the columns of X from left to right. The output of the filtering action on X_v by T is given by :

$$Y_v = TX_v$$

$$= (\sum_{k=1}^{q} H_k \otimes E_k)X_v; \quad q \leq n \tag{2.5}$$

where Y_v is an $M_1M_2 \times 1$ vector. The above is equivalent to the following representation, derived from [1] :

$$Y = D \bullet X = \sum_{k=1}^{q} H_k X E_k; \quad q \leq n \tag{2.6}$$

where Y is the $M_1 \times M_2$ matrix representation of vector Y_v. The component filters here are the (H_k, E_k) which act in parallel on the image.

In order to reduce computational complexity, one needs to reduce the number (denoted here by q) of such parallel components. The filter matrix D , as it is usually structured, is non - zero in only certain specified locations. Hence the diagonal matrix T has many zeros along its diagonal. This leads one to explore the possibility of reducing the number of parallel components by studying the nature of D and extracting the set of q independent significant component filters, where $q \leq n$. Thus, one needs to approximate the diagonal filter matrix T as :

$$T = \sum_{k=1}^{q} \Delta_k \otimes \Phi_k; \quad q \leq n \tag{2.7}$$

where Δ_k and Φ_k are diagonal matrices of sizes M_1 and M_2 respectively. Then, the output image is :

$$Y = \sum_{k=1}^{q} \Delta_k X \Phi_k; \quad q \leq n \tag{2.8}$$

The above represents a row - column operation on the elements of X with the Δ_k matrices acting on the columns of X and the Φ_k matrices acting on the rows of X. This suggests the decomposition of D into two matrices Δ and Φ such that

$$D = \Delta\Phi \tag{2.9}$$

where Δ is the $M_1 \times q$ matrix obtained by stacking the diagonal elements of each Δ_k along the $k\ th$ column, and Φ is the $q \times M_2$ matrix obtained by stacking the diagonal elements of each Φ_k along the $k\ th$ row.

The Singular Value Decomposition (SVD) can be used to find one explicit definition of the matrices Δ and Φ. From the SVD of D, one obtains :

$$D = USV^* \tag{2.10}$$

Denoting $D(x)$ as the diagonal $n \times n$ matrix whose main diagonal is the vector $x \in E^n$, it is easily inferred that

$$\Delta_k = \sqrt{S(k,k)}D(U_k) \tag{2.11}$$

$$\Phi_k = \sqrt{S(k,k)}D(V_k^*), \tag{2.12}$$

where U_k is the $k\ th$ column of U and V_k^* is the $k\ th$ row of V^*.

The $2-D$ filtering action $D \bullet X$ has thus been reduced to a sum of $1-D$ filtering actions. Using the notation $D'(X)$ to denote the vector obtained from the main diagonal of the $n \times n$ matrix X, the $1-D$ filters can be represented as the $D'(\Delta_k)\ 's$ and $D'(\Phi_k)\ 's$. This gives the most optimal (in the least mean squares sense) separable decomposition of D in the frequency domain. (Refer to Appendix A for the proof).

2.2 Reduction of the number of component filters

Reduction of the number of the $1-D$ component filters Δ_k and Φ_k is the next objective. The simplest way in which this can be achieved is by the truncation of S, neglecting all singular values σ_k *below a specified tolerance* . The tolerance value chosen for truncation depends on the nature of the application for which the filter is to be used. S is thus truncated to a $q \times q$ diagonal matrix of *significant* (as defined by the nature of the operation) singular values. U is truncated to an $M_1 \times q$ matrix and V^* is truncated to a $q \times M_2$ matrix $(q < n)$. Hence,

$$\Delta_q = \Delta_{approx} = [D'(\Delta_1) \;\; D'(\Delta_2) \;\; \ldots \;\; D'(\Delta_q)]; \tag{2.13}$$

$$\Phi_q = \Phi_{approx} = \begin{bmatrix} D'(\Phi_1)^T \\ D'(\Phi_2)^T \\ \vdots \\ D'(\Phi_q)^T \end{bmatrix} \tag{2.14}$$

$$D_q = D_{approx} = \Delta_{approx}\Phi_{approx} \tag{2.15}$$

The value of q, from now on to be referred to as the *reduced order* , may be further reduced by using an error criterion in the filter approximation. The least square error can be used for this purpose. The approximated filter D_q is compared with the actual filter D, and truncation is done if the normalized least square error of D_q is below a specified limit. It is common, in filtering operations, to specify a transition region in the frequency response. This could be incorporated in the filter to obtain further reduction in the number of the $1-D$ components.

2.2.1 Factors affecting the order of reduction : Study of symmetry

It is both interesting and important to study what determines the extent of reduction achievable in a given filter. This knowledge can then be used in the designing process to reduce the number of filter components to the smallest number possible. Certain filter types appear to lend themselves more easily than others to reduction, without noticeable distortion in the output. It can be rather easily inferred that reduction is directly related to the rank of the filter matrix D since this determines the number of non-zero singular values. Occurrence of a large number of dependent rows (columns) suggests *symmetry in the matrix structure*. The symmetry properties of the D, and the U, S and V matrices have thus to be studied in detail, to better understand the factors affecting the order of reduction.

Symmetry in the filter response implies filters whose frequency responses are quadrantally symmetric, antisymmetric, or half - plane symmetric. Again, if the frequency response is simply defined, such as square or rectangular, a great deal of symmetry would be present in the filter matrix D. The singular values, in all these cases, would be such that $S(q,q) < tol$, for $q << n$. In this discussion, the order of D is assumed to be even. This implies that D is structured to be perfectly quadrantally symmetric, antisymmetric or half - plane symmetric as the case may be. This is achieved by assigning no values in the frequency response when either w_1 or w_2 is 0.

It is known that the U, S, V matrices resulting from the SVD of a symmetrical response matrix have some special properties ([8]). The following is a discussion on the effect of this on the separable filter nature, and the simplifications achievable on this account.

Consider the $k\ th$ filter component fil_k obtained from the SVD of D. From eq. 2.10, this can be obtained as the outer product of U_k and V_k^*, weighted by the $k\ th$ singular value $\sigma_k = S(k,k)$.

$$fil_k = \begin{bmatrix} U_{k_1}\sigma_k V_{k_1}^* & \dots & U_{k_1}\sigma_k V_{k_{M_2}}^* \\ \vdots & \vdots & \vdots \\ U_{k_{M_1}}\sigma_k V_{k_1}^* & \dots & U_{k_{M_1}}\sigma_k V_{k_{M_2}}^* \end{bmatrix} \tag{2.16}$$

Now,

$$D = \begin{bmatrix} d_{1,1} & d_{1,2} & \dots & d_{1,M_2} \\ \vdots & \vdots & \vdots & \vdots \\ d_{M_1,1} & d_{M_1,2} & \dots & d_{M_1,M_2} \end{bmatrix} \tag{2.17}$$

If D is quadrantally symmetric, then it satisfies the following condition.

$$D(w_1,w_2) = D(-w_1,w_2) = D(-w_1,-w_2) = D(w_1,-w_2) \tag{2.18}$$

If D is antisymmetric, then it satisfies the following condition.

$$D(w_1,w_2) = -D(-w_1,w_2) = D(-w_1,-w_2) = -D(w_1,-w_2) \tag{2.19}$$

The following theorem is then valid for any D satisfying eqs. 2.18 or 2.19 and any of its components fil_k.

Theorem 2.1 : *The filter fil_k obtained from the SVD of D and defined as given in eq. 2.16 preserves the symmetry properties of the filter D, as long as $\sigma_k \neq 0$.*

Proof :

(i) Consider the case of a quadrantally symmetric filter matrix D. Such a filter has the properties outlined by eq. 2.18 and can be implemented by using just one quadrant response. Since the filter D is quadrantally symmetric (D_{qs}),

$$d_{i,j} = d_{i,M_2+1-j},$$

$$d_{i,j} = d_{M_1+1-i,j},$$

$$d_{i,j} = d_{M_1+1-i,M_2+1-j}, \quad where \;\; d_{i,j} = D(i,j) \tag{2.20}$$

Also, as long as $\sigma_k \neq 0$, the vectors U_k and V_k are mirror image symmetric, i.e., for all $1 \leq i \leq M_1; 1 \leq j \leq M_2; 1 \leq k \leq q; \sigma_k \neq 0$,

$$U_{k_i} = U_{k_{M_1+1-i}} \; and \; V_{k_j} = V_{k_{M_2+1-j}} \tag{2.21}$$

Hence, on the condition that $\sigma_k \neq 0$,

$$U_{k_i}\sigma_k V^*_{k_j} = U_{k_i}\sigma_k V^*_{k_{M_2+1-j}}$$

$$U_{k_i}\sigma_k V^*_{k_j} = U_{k_{M_1+1-i}}\sigma_k V^*_{k_j}$$

$$U_{k_i}\sigma_k V^*_{k_j} = U_{k_{M_1+1-i}}\sigma_k V^*_{k_{M_2+1-j}}$$

which implies that the separable filter fil_k is also quadrantally symmetric.

Hence,

$$D_{qs} \rightarrow fil_{k_{qs}} \tag{2.22}$$

(ii) Next, consider the case of D being antisymmetric (D_{as}). Such a filter has the property outlined by eq. 2.19 and can also be implemented by using just one quadrant response. For such a filter, for all $1 \leq i \leq M_1; 1 \leq j \leq M_2$,

$$d_{i,j} = -d_{i,M_2+1-j},$$

$$d_{i,j} = -d_{M_1+1-i,j},$$

$$d_{i,j} = d_{M_1+1-i,M_2+1-j}, \tag{2.23}$$

In this case, as long as $\sigma_k \neq 0$, the vectors U_k and V_k are mirror image antisymmetric, i.e., for all i, j such that $1 \leq i \leq M_1; 1 \leq j \leq M_2; 1 \leq k \leq q; \sigma_k \neq 0$,

$$U_{k_i} = -U_{k_{M_1+1-i}} \;\; and \;\; V_{k_j} = -V_{k_{M_2+1-j}} \tag{2.24}$$

Then, for $\sigma_k \neq 0$,

$$U_{k_i}\sigma_k V_{k_j}^* = -U_{k_i}\sigma_k V_{k_{M_2+1-j}}^*,$$

$$U_{k_i}\sigma_k V_{k_j}^* = -U_{k_{M_1+1-i}}\sigma_k V_{k_j}^*,$$

$$U_{k_i}\sigma_k V_{k_j}^* = U_{k_{M_1+1-i}}\sigma_k V_{k_{M_2+1-j}}^*,$$

which implies that the separable filter fil_k is also antisymmetric. Hence,

$$D_{as} \rightarrow fil_{as} \tag{2.25}$$

From eqs. 2.22 and 2.25, one can conclude that the component filters fil_k of D obtained in this manner, preserve the symmetry of D.

There is a class of filters which is neither quadrantally symmetric nor antisymmetric, but whose response is symmetric with respect to the origin of the (w_1, w_2) plane.

These filters have the following property :

$$D(w_1, w_2) = D(-w_1, -w_2) \tag{2.26}$$

These are called half - plane symmetric filters. Thus, for a half - plane symmetric filter D,

$$d_{i,j} = d_{M_1+1-i, M_2+1-j}, \quad 1 \leq i \leq M_1; 1 \leq j \leq M_2 \tag{2.27}$$

If fil_k is assumed to have the symmetry of D, it would imply that U_k and V_k need to both be either mirror image symmetric or antisymmetric, which is true of only certain filters in this category, like for example, the rotated elliptical filter. Thus, only in these cases do the component filters have the symmetry properties of D. These filters can be designed using only the first two quadrants of the desired response D.

As an extension, consider the component filters in pairs. The $i\ th$ filter pair is :

$$fil_k = \begin{bmatrix} U_{k_1}\sigma_k V^*_{k_1} + U_{k+1_1}\sigma_{k+1} V^*_{k+1_1} & \cdots \\ \vdots & \vdots \\ \cdots & U_{k_{M_1}}\sigma_k V^*_{k_{M_2}} + U_{k+1_{M_1}}\sigma_{k+1} V^*_{k+1_{M_2}} \end{bmatrix} \tag{2.28}$$

Now, the U, V matrices in some special classes of half - plane symmetric filters, like the fan and triangular filters, have the property that their columns, when considered in pairs, k and $k+1$, (k : *odd, k+1 : even*) are pairwise mirror image symmetric or antisymmetric, with the singular values also varying in pairs. Thus, for these filters,

$$U_{k_i} = \pm U_{k+1_{M_1+1-i}}, \quad 1 \leq i \leq M_1$$

$$V_{k_j} = \pm V_{k+1_{M_2+1-j}}, \;\; 1 \leq j \leq M_2$$

and

$$\sigma_k = \sigma_{k+1}, \;\; 1 \leq k \leq q$$

This implies that fil_k as defined in eq. 2.28 is half - plane symmetric.

Thus for these half - plane symmetric cases, the component filters when considered in pairs, k and $k+1$ (*k : odd, k+1 : even*), preserve the symmetry of D.

2.2.2 Implications

The implications of the results obtained in the last section are significant on the computational complexity of the design algorithm. If the desired filter response D is symmetric in any one of the three senses described earlier, *it is usually sufficient to design each component filter in either one or two quadrants only.* This leads to an enormous reduction in the memory requirements. There is also a huge computational speed-up since one is required to work with a smaller number of frequency sample points. For filters which do not possess any kind of symmetry, it is possible to decompose them into their symmetrical and antisymmetrical components, and design for each component separately. Thus, the particular cases of symmetry analyzed above may be extended to any class of filters by using a suitable decomposition.

2.3 Experimental results

In this section, experimental results are presented on the effects of symmetry, filter shape, the transition region, and the truncation criterion used, on the order of reduction. The nature of the U, S and V matrices were studied for the three cases of symmetry in D. In each case, the special properties attributed to these matrices were verified, as long as the singular values were non - zero. This establishes that the component filters do exhibit the symmetry of D, which was further verified from a simulation of the individual filter components.

2.3.1 Effect of filter configuration

Figure 2.1 presents the filters used for the study of the effect of filter configurations. In each case, a 128×128 ideal response filter matrix was considered.

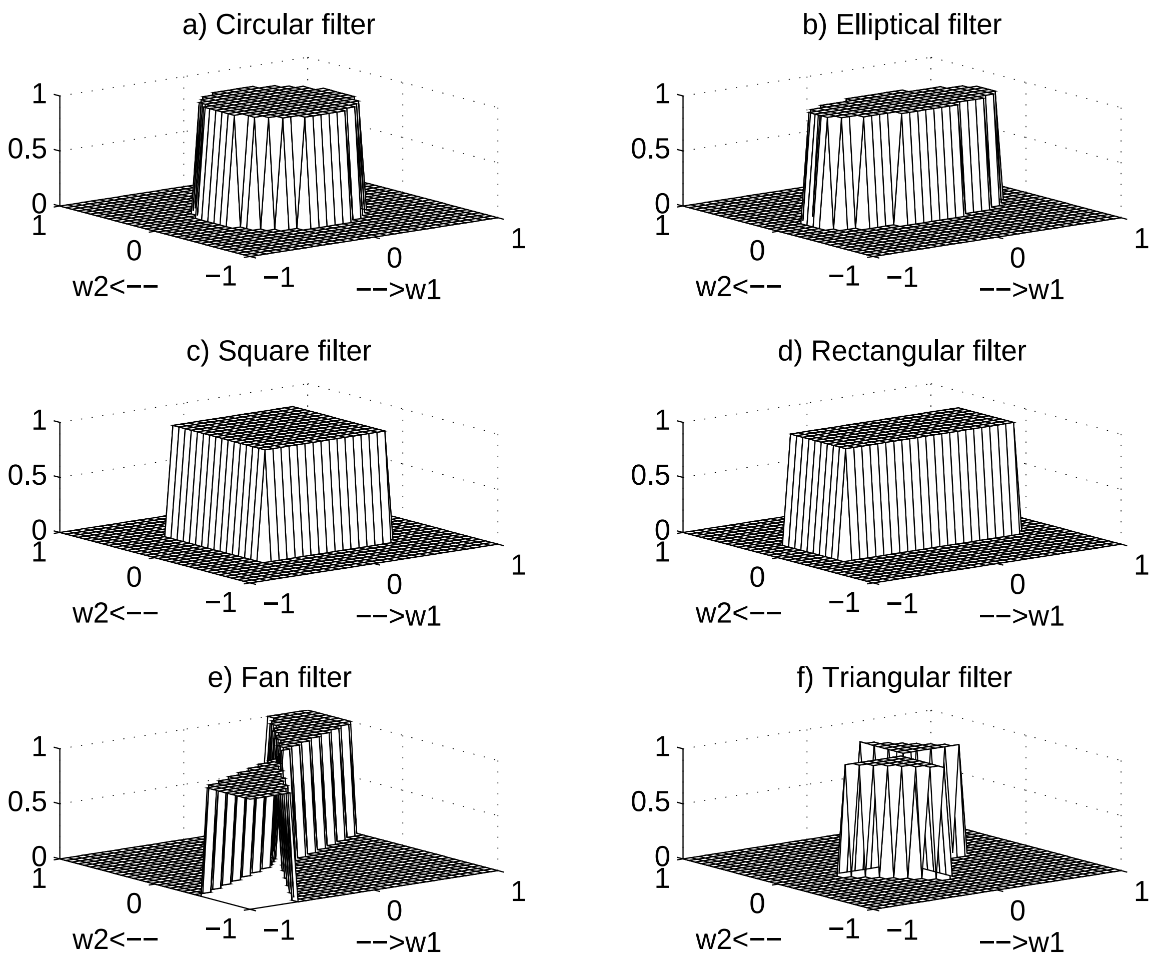

Figure 2.1: Filter configurations

The reduced order was first determined based on truncation if $\sigma_k < tol$, where tol $= 0.001$. The values obtained for the different cases are listed in Table 2.1.

Table 2.1: Reduced order based on truncation

Case	q
a	18
b	22
c	3
d	3
e	52
f	88

The reduced order q was determined next by truncation based on the least mean square error using the Frobenius norm. The criterion for truncation was $\%err < 3$, where

$$\%err = \frac{norm[D - \sum_{k=1}^{q} D'(\Delta_k)D'(\Phi_k)]}{norm[D]} \tag{2.29}$$

The reduced order values are given in Table 2.2.

Table 2.2: Reduced order based on norm of error

Case	q
a	16
b	21
c	1
d	1
e	51
f	76

Finally, the effect of a transition bandwidth in the filter D was evaluated. The value used for the transition region was 0.1π, which is 10% of the defined region of the frequency response. Table 2.3 gives the reduced orders for the different cases. It can be observed from the data presented that filters which are quadrantally symmetric like a, b, c, d have lower q values than e, f. Among these, c, d have a reduced order of 1, based on the last two truncation criteriea used. It is to be noted that these are the most simply structured filter matrices used, having many rows (columns) which

Table 2.3: Reduced order based on incorporation of a transition band

Case	q
a	16
b	16
c	1
d	1
e	46
f	76

are entirely 0, and only 1 independent row (column). Thus, symmetry in the filter configuration and simplicity of shape contributes to a large reduction in the number of separable filters required.

Chapter 3

Mathematical Analysis

3.1 Introduction

Once it has been established that parallel filter implementation is both feasible and desirable, it is necessary to design an algorithm to obtain the optimal parallel FIR filters. [3] describes an algorithm to compute a $2-D$ non-separable FIR digital filter which is optimal in the least mean squares (LMS) sense. [4] details the extension of this algorithm for the case when we have a general weighting function. In the absence of any constraints, the SVD of the optimal non-separable filter designed using [4] gives the $1-D$ filters. This approach however, leads at most to a *suboptimal* solution since each $1-D$ response must then be approximated by FIR filters. In order to obtain an optimal solution, one needs to incorporate the FIR constraint in the optimization problem. The problem definition and its solution is developed in the next section. The formulation is quite general and permits the inclusion of weighting functions which can help to overcome some of the known limitations of the LMS approach and allow the inclusion of transition bands.

3.1.1 Review

Properties of the *trace*, the *outer product*, and the Kronecker product functions are used extensively in this analysis. For completeness, this section offers a brief review of the relevant concepts.

$$\langle x, Py \rangle = \langle P^* x, y \rangle \tag{3.1}$$

$$\langle A, B\rangle = \langle B, A\rangle \tag{3.2}$$

$$tr\{AB^*\} = tr\{BA^*\} \tag{3.3}$$

$$tr\{AB^*\} = \langle A, B\rangle \tag{3.4}$$

As stated previously in eq. 2.1, if $A \in E^{M_1 \times M_2}$ and $B \in E^{M_3 \times M_4}$, the Kronecker product of A and B is defined by,

$$A \otimes B = [AB_{i,j}] \in E^{M_1 M_3 \times M_2 M_4} \tag{3.5}$$

3.1.2 Definitions used in analysis

The following linear transformations are used in the mathematical development. The *weight* map $\mathcal{W}$ is given by

$$\mathcal{W}(E) = W \bullet E. \tag{3.6}$$

The next development shows that the map is positive semidefinite.

$$\langle B, \mathcal{W}(A)\rangle = tr\{(\mathcal{W}(A))^* B\} = tr\{B(\mathcal{W}(A))^*\}$$

$$= tr\{B(W^* \bullet A^*)\}$$

$$= \sum b_{i,j}(w^c_{i,j} a^c_{i,j})$$

$$= \sum (b_{i,j} w^c_{i,j}) a^c_{i,j}$$

$$= tr\{(B \bullet W^c)A^*\}$$

$$= tr\{(W^c \bullet B)A^*\}$$

$$= \langle (W^c \bullet B), A \rangle$$

$$= \langle \mathcal{W}^c(B), A \rangle$$

$$\Rightarrow \mathcal{W}^c(B) = W^c \bullet B = \mathcal{W}^*(B) \tag{3.7}$$

and if $\mathcal{W}(B)$ is real, then $\mathcal{W}^*(B) = \mathcal{W}(B)$. Also, by setting $A = B$, positivity follows immediately whenever $w_{i,j} \geq 0$.

The *frequency* map $\mathcal{F}$ is defined as

$$\mathcal{F}(A) = \Omega_1 A \Omega_2^T, \tag{3.8}$$

This map gives the discrete frequency response of a filter with the coefficient matrix A, and Ω_1 and Ω_2 represent the discrete frequency matrices defined as :

$$\Omega_i(k_1, k_2) = e^{-\jmath w_{ik_1} k_2}; \tag{3.9}$$

$$k_1 = -M_i, \ldots, M_i - 1, k_2 = -N_i, \ldots, N_i, i = 1, 2.$$

Using a procedure similar to the one outlined before, the adjoint can be computed and is given by

$$\mathcal{F}^*(A) = \Omega_1^* A \Omega_2^c \tag{3.10}$$

If $X \in E^{N_1 \times N_2}$ is an $N_1 \times N_2$ matrix, one can define an $N_1 N_2 \times 1$ vector X_v formed by stacking the columns of X from left to right. This *stacking* operation is denoted as

$$\mathcal{S}(X) = X_v$$

In particular,

$$\mathcal{S}(ab^T) = a \otimes b \tag{3.11}$$

The vectors $a \otimes b$ and $b \otimes a$ have the same components but arranged in a different order. Hence, they can be related by a suitably defined permutation $\mathcal{T}$

$$\mathcal{T}(a \otimes b) = b \otimes a \tag{3.12}$$

Using the same technique as before, the adjoints of the *stacking* and the *rearranging* operations can be shown to be

$$\mathcal{S}^*(X_v) = X \tag{3.13}$$

$$\mathcal{T}^*(b \otimes a) = a \otimes b \tag{3.14}$$

We observe from the above results, that

$$\mathcal{S}\mathcal{S}^* = \mathcal{S}^*\mathcal{S} = I$$

$$\mathcal{T}\mathcal{T}^* = \mathcal{T}^*\mathcal{T} = I$$

Finally, from the definition of $a \otimes b$, one can establish the identity,

$$a \otimes b = (a \otimes I_{N_2})b, \tag{3.15}$$

where I_{N_2} is the identity matrix of order N_2, and $(a \otimes I_{N_2})$ is an $N_1N_2 \times N_2$ matrix.

3.2 Best separable approximation to a $2-D$ FIR filter

The *best separable FIR filter* approximation to the ideal filter response D minimizes the cost function

$$J(a,b) = tr\{(W \bullet E)E^*\}, \tag{3.16}$$

where $a \in E^{N_1}$, $b \in E^{N_2}$ are the vectors of filter coefficients, and $E = D - \Omega_1 ab^T \Omega_2^T$ is the error function matrix. The discrete frequency matrices Ω_1 and Ω_2 have been defined in eq. 3.9.

3.2.1 Development of the necessary conditions

Using the notation defined earlier, the cost function defined in eq. 3.16 can be written as:

$$J = \left\langle \mathcal{W}(D - \mathcal{F}(ab^T)), (D - \mathcal{F}(ab^T)) \right\rangle$$

$$= \left\langle \mathcal{W}(D) - \mathcal{W}(\mathcal{F}(ab^T)), (D - \mathcal{F}(ab^T)) \right\rangle$$

Expanding the inner product, one obtains,

$$J(a,b) = \langle \mathcal{W}(D), D \rangle - \left\langle \mathcal{W}(\mathcal{F}(ab^T)), D \right\rangle - \left\langle \mathcal{W}(D), \mathcal{F}(ab^T) \right\rangle$$

$$+ \left\langle \mathcal{W}(\mathcal{F}(ab^T)), \mathcal{F}(ab^T) \right\rangle \tag{3.17}$$

Using adjoints, the above expression simplifies to

$$J(a,b) = \langle \mathcal{W}(D), D \rangle - \left\langle \mathcal{F}(ab^T), \mathcal{W}^*(D) \right\rangle - \left\langle \mathcal{F}^*\mathcal{W}(D), ab^T \right\rangle$$

$$+\left\langle \mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)), ab^T \right\rangle$$

$$= \langle \mathcal{W}(D), D\rangle - \left\langle ab^T, \mathcal{F}^*\mathcal{W}^*(D)\right\rangle - \left\langle \mathcal{F}^*\mathcal{W}(D), ab^T\right\rangle$$

$$+\left\langle \mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)), ab^T \right\rangle \tag{3.18}$$

Taking variations with respect to the vector a (b fixed), one gets,

$$\delta J_a = -\left\langle \delta ab^T, \mathcal{F}^*\mathcal{W}^*(D)\right\rangle - \left\langle \mathcal{F}^*\mathcal{W}(D), \delta ab^T\right\rangle + \left\langle \mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)), \delta ab^T\right\rangle$$

$$+\left\langle \mathcal{F}^*\mathcal{W}(\mathcal{F}(\delta ab^T)), ab^T\right\rangle$$

$$= -\left\langle \delta ab^T, \mathcal{F}^*\mathcal{W}^*(D)\right\rangle - \left\langle \mathcal{F}^*\mathcal{W}(D), \delta ab^T\right\rangle + \left\langle \mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)), \delta ab^T\right\rangle$$

$$+\left\langle \delta ab^T, \mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T))\right\rangle \tag{3.19}$$

Using the trace function, and incorporating the fact that the $\mathcal{W}(D)$ operation is real, eq. 3.19 can be written as,

$$\delta J_a = -2tr\{\mathcal{F}^*\mathcal{W}(D)b^c\delta a^*\} + tr\{\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T))b^c\delta a^*\} + tr\{\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T))b^c\delta a^*\} \tag{3.20}$$

$$= -2tr\{\delta a^*\mathcal{F}^*\mathcal{W}(D)b^c\} + 2tr\{\delta a^*\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T))b^c\}$$

Hence, $\delta J_a = 0$ if and only if

$$tr\{\delta a^*\mathcal{F}^*\mathcal{W}(D)b^c\} = tr\{\delta a^*\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T))b^c\}$$

Now, the quantities within the trace are scalars. Hence,

$$\delta a^*\mathcal{F}^*\mathcal{W}(D)b^c = \delta a^*\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T))b^c$$

Since δa^* is arbitrary, this equation is satisfied if and only if

$$(\mathcal{F}^*\mathcal{W}(D))b^c = (\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)))b^c \tag{3.21}$$

Now taking variations with respect to the vector b (a fixed), and proceeding in a similar fashion, one gets the second necessary condition for optimality.

$$a^*\mathcal{F}^*\mathcal{W}(D) = a^*\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T))$$

Performing the transpose operation on both sides, one gets

$$(\mathcal{F}^*\mathcal{W}(D))^T a^c = (\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)))^T a^c \tag{3.22}$$

3.2.2 Solution procedure

Equations 3.21 and 3.22 represent a set of non-linear equations for which no direct solution exists. They are solved using an iterative technique which is explained in this section. However, first, one needs an efficient representation for the map $\mathcal{F}^*\mathcal{W}\mathcal{F}$. Such an expression is derived below using the SVD of the weighting function matrix W. This computation is required to be done only once for a given weight matrix, and does not increase the design time significantly.

From eq. 2.10, one has,

$$W = U_w S_w V_w^*. \tag{3.23}$$

The operation $\mathcal{W}(E)$ can then be obtained from eqs. 2.8 and 3.6 as :

$$\mathcal{W}(E) = \sum_{k=1}^{q} L_{w_k} E R_{w_k}; \quad q \leq min(M_1, M_2); \tag{3.24}$$

where the matrices L_{w_k} and R_{w_k} are obtained from eqs. 2.11 and 2.12 as :

$$L_{w_k} = \mathcal{D}(U_{w_k})\sqrt{S_w(k,k)} \tag{3.25}$$

$$R_{w_k} = \sqrt{S_w(k,k)}\mathcal{D}(V^*_{w_k}) \tag{3.26}$$

Then, from eqs. 3.6, 3.8, 3.10, and 3.24,

$$(\mathcal{F}^*\mathcal{W}(D))b^c = \Omega_1^* \sum_{k=1}^{q} L_{w_k} D R_{w_k} \Omega_2^c b^c; \tag{3.27}$$

$$(\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)))b^c = \Omega_1^* \sum_{k=1}^{q} L_{w_k} \Omega_1 a b^T \Omega_2^T R_{w_k} \Omega_2^c b^c \tag{3.28}$$

Note that $b^T\Omega_2^T R_{w_k}\Omega_2^c b^c$ is a scalar. Eq. 3.28 can therefore be written as :

$$(\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)))b^c = \sum_{k=1}^{q} b^T \Omega_2^T R_{w_k} \Omega_2^c b^c \Omega_1^* L_{w_k} \Omega_1 a \tag{3.29}$$

$$= \sum_{k=1}^{q} b^T R_k b^c L_k a \tag{3.30}$$

where

$$L_k = \Omega_1^* L_{w_k} \Omega_1,$$

$$R_k = \Omega_2^T R_{w_k} \Omega_2^c \tag{3.31}$$

$$(\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)))^T a^c = \sum_{k=1}^{q} (\Omega_1^* L_{w_k} \Omega_1 a b^T \Omega_2^T R_{w_k} \Omega_2^c)^T a^c \tag{3.32}$$

$$= \sum_{k=1}^{q} \Omega_2^* R_{w_k} \Omega_2 b a^T \Omega_1^T L_{w_k} \Omega_1^c a^c$$

Noting that $a^T\Omega_1^T L_{w_k}\Omega_1^c a^c$ is a scalar, one gets,

$$(\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)))^T a^c = \sum_{k=1}^{q} a^T\Omega_1^T L_{w_k}\Omega_1^c a^c \Omega_2^* R_{w_k}\Omega_2 b \tag{3.33}$$

Now,

$$L_k^T = \Omega_1^T L_{w_k}\Omega_1^c = L_k^c$$

$$R_k^T = \Omega_2^* R_{w_k}\Omega_2 = R_k^c \tag{3.34}$$

Eq. 3.33 then reduces to :

$$(\mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)))^T a^c = \sum_{k=1}^{q} a^T L_k^c a^c R_k^c b, \tag{3.35}$$

The following are defined for convenience of notation.

$$G = \Omega_1^* \sum_{k=1}^{q} L_{w_k} D R_{w_k}\Omega_2^c \tag{3.36}$$

$$X = \sum_{k=1}^{q} b^T R_k b^c L_k \tag{3.37}$$

$$Y = \sum_{k=1}^{q} a^T L_k^c a^c R_k^c \tag{3.38}$$

The set of necessary conditions (eqs. 3.21 and 3.22) is then equivalent to :

$$Gb^c = Xa \tag{3.39}$$

$$G^T a^c = Yb, \tag{3.40}$$

This set of equations can be solved using successive approximations. The initial choice of b is arbitrary and the system of equations is solved for a. This value of a is then used to solve for b. This process is repeated until the minimal optimal index value is reached and convergence results. The following theorem is presented as proof of the existence of a solution to this system.

Theorem 3.1 : *The system of equations defined by eqs. 3.39 and 3.40 has a limiting solution if and only if there exists a unique optimal non-separable FIR filter.*

The proof of this is given in Appendix B. The proof establishes the existence of a sequence of $a'_k s$, all contained on the closed unit ball, and associated with a monotonically decreasing sequence of costs. Extracting a convergent subsequence, and using the limit of this subsequence of $a'_k s$, one can derive an optimal filter coefficient pair (a^0, b^0), which defines the best separable FIR approximation to the desired frequency response in the weighted least mean squares sense.

Chapter 4

Extension to multiple terms

In Chapter 3, the procedure to obtain the *best separable approximation* to any arbitrary $2-D$ ideal frequency response was outlined. In most cases, this 1-term approximation may not be sufficient to meet the given performance specifications. One would need to design more such separable filters, the sum of which would meet all performance requirements. This chapter describes the extension of the algorithm to design any number of separable filters. It also gives the details of the actual implementation of the filtering action using these filters.

4.1 Obtaining multiple terms

The cost function to be minimized is :

$$J = \left\langle \mathcal{W}(D - \mathcal{F}(ab^T)), (D - \mathcal{F}(ab^T)) \right\rangle \tag{4.1}$$

Set $D = D_1$ where D_1 is the ideal frequency response, and solve the above problem to find the optimal $a = a_1$ and $b = b_1$ using the iterative procedure explained in Chapter 3. Then $\mathcal{F}(a_1 b_1^T)$ is the best separable 1-term approximation to D_1. To find the second best separable approximation, set $D = D_1 - \mathcal{F}(a_1 b_1^T) = D_2$, and solve the problem to find $a = a_2$ and $b = b_2$. Then, $\mathcal{F}(a_2 b_2^T)$ is the most optimal approximation to D_2, and $\mathcal{F}(a_1 b_1^T) + \mathcal{F}(a_2 b_2^T)$ is the best 2-term approximation to D_1. To find the *best n term* approximation to D_1, one needs to set $D = D_1 - \sum_{k=1}^{n-1} \mathcal{F}(a_k b_k^T)$, and solve the minimization problem to find a_n and b_n. Then, $\sum_{k=1}^{n} \mathcal{F}(a_k b_k^T)$ represents the best n term approximation to D_1. Thus, the given ideal frequency response is approximated by separable filters using *the method of successive approximation.* The process is terminated when all performance specifications are met.

4.1.1 Reduction of the number of channels using SVD

Once the q separable filters have been designed which meet all performance requirements, one has a set (a_k, b_k) of q coefficient vectors. These vectors represent the q channels of the filter which operate on the input image. The sum of the outputs from each channel is then the output filtered image. It has been observed in the case of FIR filters ([12]) that *the number of channels can be reduced by a significant factor* for any arbitrary filter configuration. For this, one needs to determine the summation of the q rank - 1 matrices $a'_k b'^T_k$, where a'_k and b'_k represent the coefficient vectors adjusted to the same size by padding zeros where required. Thus, one obtains the coefficient matrix,

$$C = \sum_{k=1}^{q} a'_k b'^T_k$$

From the SVD of C, the set of q' significant (which still meets all performance specifications) singular values is obtained, where $q' \leq q$. One can thus obtain the set of q' coefficient vectors (a_k, b_k) from the SVD matrices, which represent the q' channels of the filter, all of which act on the input to produce an output of a similar quality as before, though with less number of channels than used earlier. For filters with highly symmetric responses, the factor of reduction achievable using this technique can be extremely high.

4.2 Modifications for filters with symmetry

4.2.1 Quadrantally symmetric/antisymmetric responses

The filters which are quadrantally symmetric or antisymmetric (eqs. 2.18 and 2.19) always have real coefficients. Hence, the necessary conditions for optimality (eqs. 3.39 and 3.40) may be modified as :

$$\Re(G)b = \Re(X)a \tag{4.2}$$

$$\Re(G^T)a = \Re(Y)b, \tag{4.3}$$

In addition, the optimal coefficient vectors are symmetric (/antisymmetric) about their middle. Hence, it is sufficient to find the coefficients $a_{k_0}, a_{k_1}, \ldots, a_{k_{N_1}}$, and $b_{k_0}, b_{k_1}, \ldots, b_{k_{N_2}}$. For this, the Ω_i matrices are modified to take into account the symmetry of D.

$$\Omega_i = \Omega_i S \tag{4.4}$$

where

$$S = \begin{bmatrix} I_{N_i+1} & \\ I'_{N_i} & 0 \end{bmatrix} \tag{4.5}$$

If D is quadrantally symmetric,

$$I'_N(i,j) = \begin{cases} 1 & \text{if i=N-j+1} \\ 0 & \text{otherwise} \end{cases}$$

and if D is antisymmetric,

$$I'_N(i,j) = \begin{cases} -1 & \text{if i=N-j+1} \\ 0 & \text{otherwise} \end{cases}$$

Only one quadrant of the frequency response is needed to obtain the optimal coefficients, since the response is symmetric. Hence, one considers only the positive frequencies.

$$w_{ik_i} = \frac{\pi(2k_i+1)}{2M_i}; \quad k_i = 0,1,\ldots,M_i-1; i = 1,2$$

4.2.2 Half-plane symmetric responses

Now, consider the case when D_1 is half-plane symmetric. The optimal coefficients here have a special property which is presented in the following theorem. This property enables one to design certain special classes of filters which always have real coefficients.

Theorem 4.1 : *Suppose that $\mathcal{W}(D)$ is real and half-plane symmetric, if (a,b) is an optimal solution to the cost function defined in eq. 4.1, then (a^c, b^c) is also an optimal solution.*

The proof of this is given in Appendix C. W is chosen to be real and half-plane symmetric which ensures that the operation $\mathcal{W}(D)$ is real and half-plane symmetric. This property is used to show that the matrix G defined in 3.36 is real, and the matrices L_k and R_k are either both real or both imaginary, which fact is then used to establish that the optimizing solutions in this case always exist in conjugate pairs. A corollary to this theorem is stated below, and it shows that real coefficients can be obtained under special conditions, with significantly less designing effort, as explained in the appendix.

Corollary : *Under the assumptions of the above theorem, if* $W \bullet D = USV^*$ *is the SVD of* $W \bullet D$*, the structure of the* U, S, V *matrices can be used to obtain real coefficients in the design process.*

Thus, if D is half-plane symmetric, and W is chosen such that $\mathcal{W}(D)$ is real and half-plane symmetric, and satisfies either one of the special conditions stated in the Corollary in Appendix C, then one can always obtain real coefficients. For the case where the optimal solutions exist as sums of conjugate pairs, one can obtain m separable filters by solving the necessary conditions only $m/2$ times. The remaining terms can be obtained just by taking the conjugate of the optimal solutions. In this case, the best separable solution $\mathcal{F}(a_1 b_1^T)$ is first obtained. Then, D is set to $D = D_1 - (\mathcal{F}(a_1 b_1^T) + \mathcal{F}(a_1^c b_1^{cT}))$, and the problem is solved to obtain (a_2, b_2). This process is repeated until all performance specifications are met.

If the ideal response does not possess any symmetry, it is separated into symmetric and antisymmetric components, and the best n term separable approximation to each is obtained. This process ensures that the optimal coefficients are always real, and hence their realization is feasible.

4.3 Implementation of the filtering action in the time domain

While the optimal coefficients have been obtained, as explained above, using frequency response analysis, the actual filtering operation is carried out in the time domain. Figure 1.1 shows the parallel separable implementation of the filtering action. The image x is loaded onto a shared memory machine. The q separable filters are then made to operate on the image in parallel, each a_k acting on all the columns of the image concurrently, after which each b_k acts on all the rows of the image concurrently. The actual filtering action of each of the separable filters is a series of two $1-D$ convolution operations. These operations are defined in eqs. 4.6 and 4.7.

$$y_1(k_1,k_2)=\sum_{k=1}^{q}\sum_{l_1=1}^{k_1}a_k(l_1)x(k_1-l_1,k_2);\quad k_1=-M_1,\ldots,M_1-1, \tag{4.6}$$

$$y(k_1,k_2)=\sum_{k=1}^{q}\sum_{l_2=1}^{k_2}b_k(l_2)y_1(k_1,k_2-l_2);\quad k_2=-M_2,\ldots,M_2-1 \tag{4.7}$$

Now, the entire filtering action is done in real time, which implies that the filter coefficients need to be physically realizable, *i.e.* they need to be real. If the ideal response is quadrantally symmetric or antisymmetric, the coefficients are always real, and hence, their realization is not a problem. If the filter is half-plane symmetric, and satisfies any one of the special conditions outlined in the Corollary to the Theorem in Appendix C, then also the coefficients are real, and can be easily realized. The filters which possess no such symmetry however, pose a problem. One can overcome this problem by constraining the solution to be real, but this has the disadvantage of taking too much computation time. Another approach is to decompose the desired response D into symmetric (D_c) and antisymmetric (D_s) responses in the design stage. The optimal separable solutions are then found for each of these two components separately, and then combined suitably to give the complete response. If (a_{k_c}, b_{k_c}) and (a_{k_s}, b_{k_s}) represent the coefficient vectors for the symmetric components D_c and D_s respectively, then the optimal coefficient vectors for D are given by the SVD of the coefficient matrix h_k obtained from $h_{k_c}=a_{k_c}b_{k_c}^T$ and $h_{k_s}=a_{k_s}b_{k_s}^T$ as:

$$h_k=\frac{h_{k_c}-h_{k_s}}{4e},$$

where e is defined as

$$e(i,j) = \begin{cases} \frac{1}{2} & \text{if i=j=0} \\ \frac{1}{\sqrt{2}} & \text{if i=0 or j=0} \\ 1 & \text{otherwise} \end{cases}$$

Chapter 5

Simulation results

In this chapter, the separable design algorithm enumerated in Chapters 3 and 4 is used to design certain well - known filters. The results of these designs are presented here and are compared with the designs obtained using the weighted least mean square non - separable design algorithm ([3],[4]), with respect to quality of the design and the computational complexity of the design algorithm.

5.1 Experimental results

5.1.1 Design examples

Two aspects of the design algorithm are illustrated here using suitable examples.

(i) For filters which are *almost separable* in nature, excellent designs can be obtained using very few terms.

(ii) Approximating the optimal non-separable design ([4]) *a posteriori* with lesser number of terms, to reduce its complexity, results in inferior filters than those created with this new technique.

Evidently, since the new technique uses only a limited number of separable terms, the design will, in general, yield a result no better than the optimal non-separable, taken as a whole. Consequently, only an equivalent number of terms of the optimal non-separable are considered for comparison. This is done by first finding the optimal non-separable $L_1 \times L_2$ filter response coefficient matrix h, by combining the coefficient matrices of the symmetric and antisymmetric components of the designed filter given by h_c and h_s respectively as follows.

$$h = \frac{h_c - h_s}{4e},$$

where e is defined as given in Chapter 4 as

$$e(i,j) = \begin{cases} \frac{1}{2} & \text{if i=j=0} \\ \frac{1}{\sqrt{2}} & \text{if i=0 or j=0} \\ 1 & \text{otherwise} \end{cases}$$

The SVD of h is then determined and from the U, S and V matrices of this decomposition, the *n term* approximation is obtained as :

$$h_{approx} = U(1:L_1, 1:n)S(1:n, 1:n)V^*(1:n, 1:L_2)$$

Example 1

The first example considers a one quadrant fan filter. This filter is almost completely separable, and is used here to validate this technique. The ideal filter frequency response is shown in Figure 5.1. This filter has the I and the III quadrants as passband and the II and the IV quadrants as the stopband with an internal transition band of width $.1\pi$ rad. Figure 5.3 shows the optimal *one term* separable filter designed using the separable algorithm. For this design, the filter orders N_1 and N_2 were taken equal to one another, and set at a value of 22. One term here refers to the optimal solution to the ideal response and its conjugate, taken together, since this filter has the properties that ensure that the optimal solutions exist as sums of conjugate pairs. The quality of the approximation is evident from the magnitude frequency response of the designed filter. The response is smooth (less ripples) and the passband edges are well defined. The error frequency response is shown in Figure 5.4, and it is seen that the maximum error is well within acceptable error bounds. The number of iterations required to obtain this one term is 105. The value of the cost function is 0.1816. Observe that this filter has a quality far superior to the one term optimal non-separable (Figure 5.2), while using only a small fraction ($\frac{90}{1013} \approx 9\%$) of the coefficients in the design process. Thus, this example validates the technique, giving a very good quality design with highly reduced computational complexity.

Example 2

The second example is that of the elliptical filter, and it was chosen to demonstrate the efficacy of this design algorithm as against that of the optimal non-separable with an equivalent number of terms. Here, the ideal filter (Figure 5.5) is a quadrantally symmetric elliptical filter with axes of $.7\pi$ rad and $.3\pi$ rad and an external transition band of width $.1\pi$ rad. The filter orders in the w_1 and w_2 directions were taken to be $N_1 = N_2 = 22$, as before. The optimal five term non-separable response is shown in Figure 5.6. It uses 1013 independent coefficients for the design. Figure 5.7 gives the best five term separable magnitude frequency response. This was obtained using only one quadrant of the ideal frequency response, and less than half the number of filter coefficients (460 for the 10 terms originally designed). The filter was originally designed using 10 terms, and then the SVD method explained in Chapter 4 was used to reduce the number of channels from 10 to 5, with almost no deterioration in the performance. The contour of the magnitude frequency response is shown in Figure 5.8. From the contour, it can clearly be seen that this filter was obtained by the addition of several rectangles, underlying the basis of this design algorithm. From the error response (Figure 5.9), it can be seen that the maximum

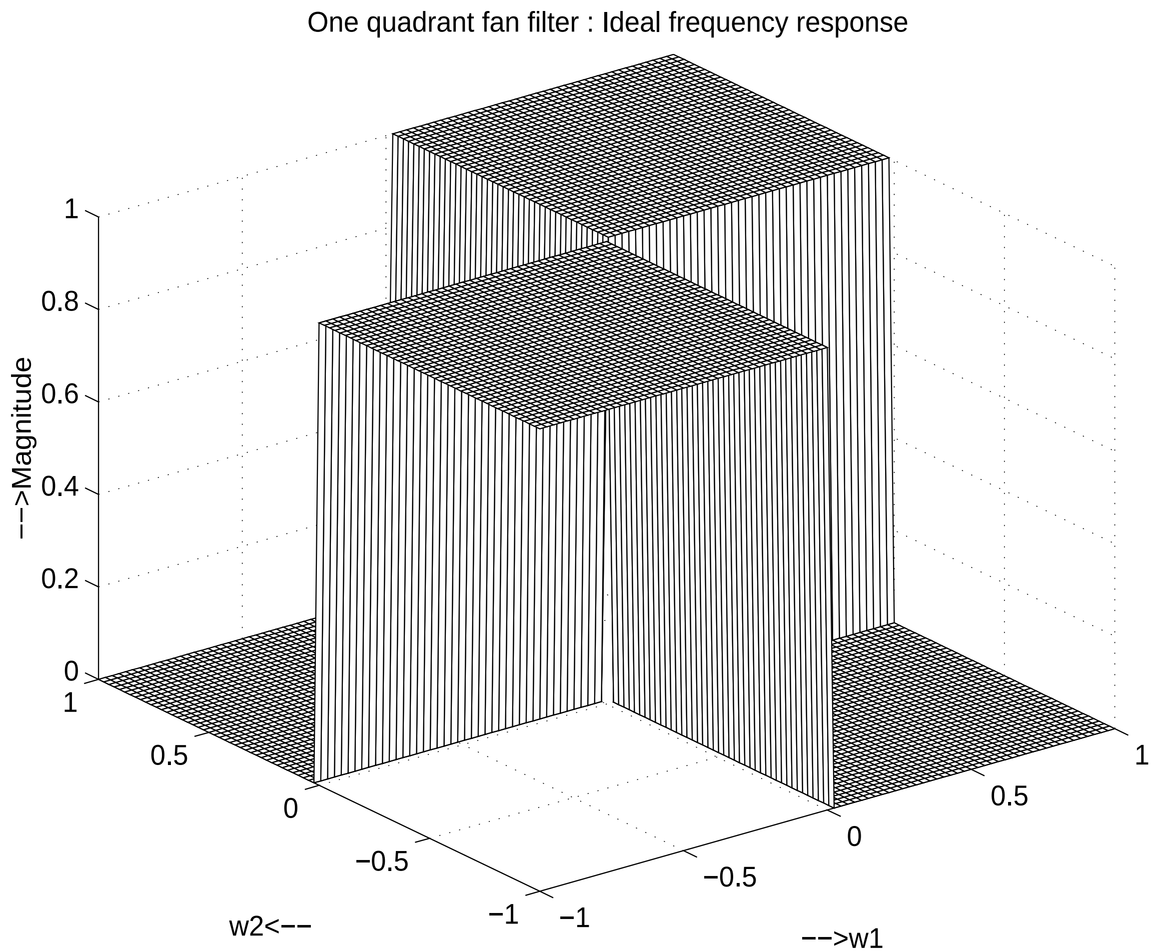

Figure 5.1: Ideal magnitude frequency response of 2-D one quadrant fan filter

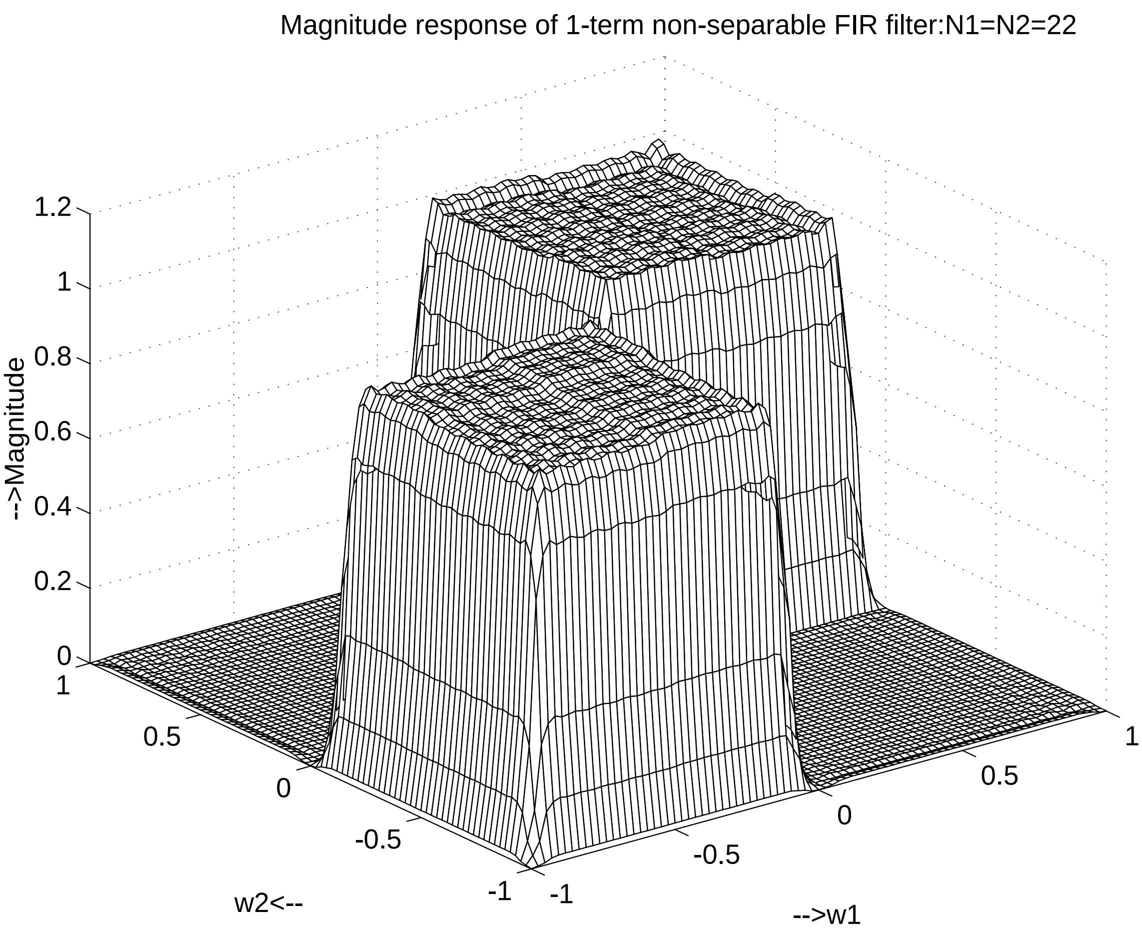

Figure 5.2: Magnitude frequency response of the optimal 1-term non-separable FIR one quadrant fan filter

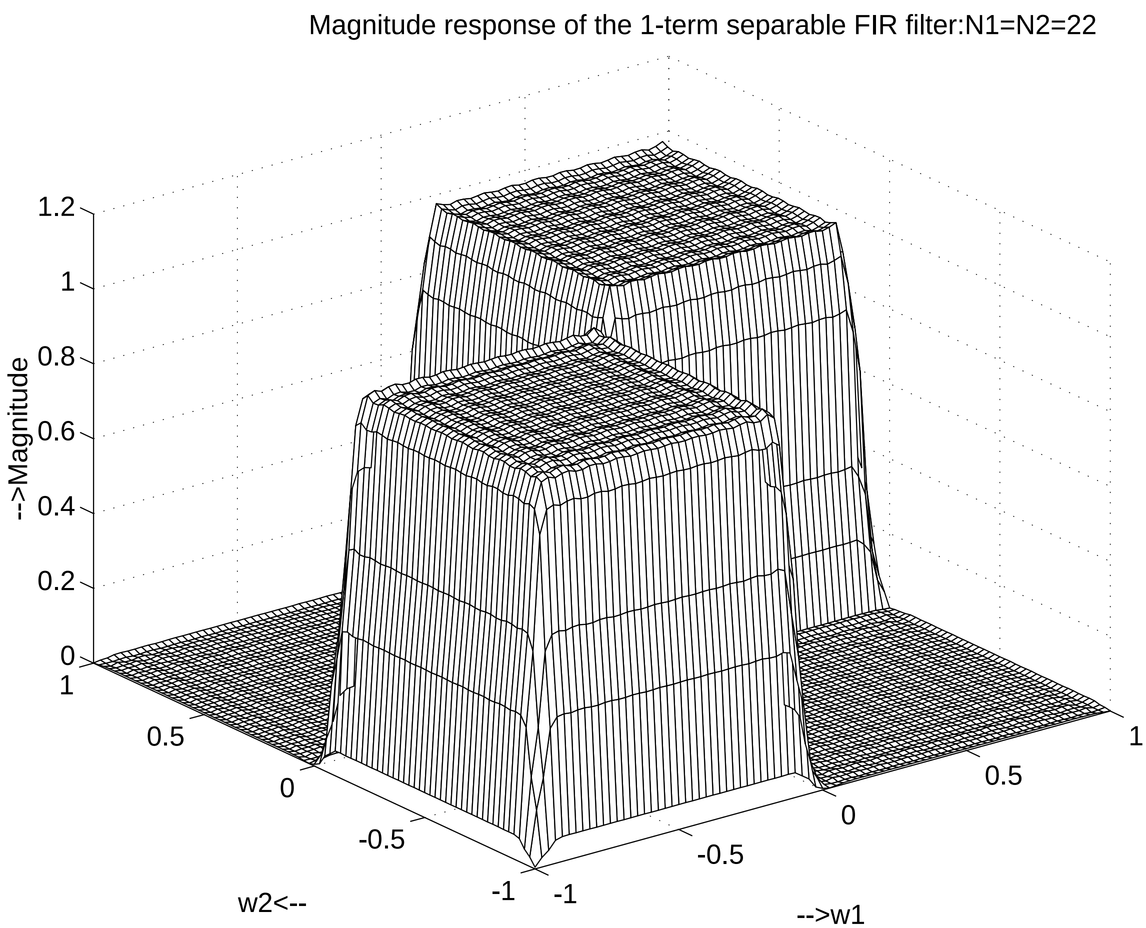

Figure 5.3: Magnitude frequency response of the optimal 1-term separable FIR one quadrant fan filter

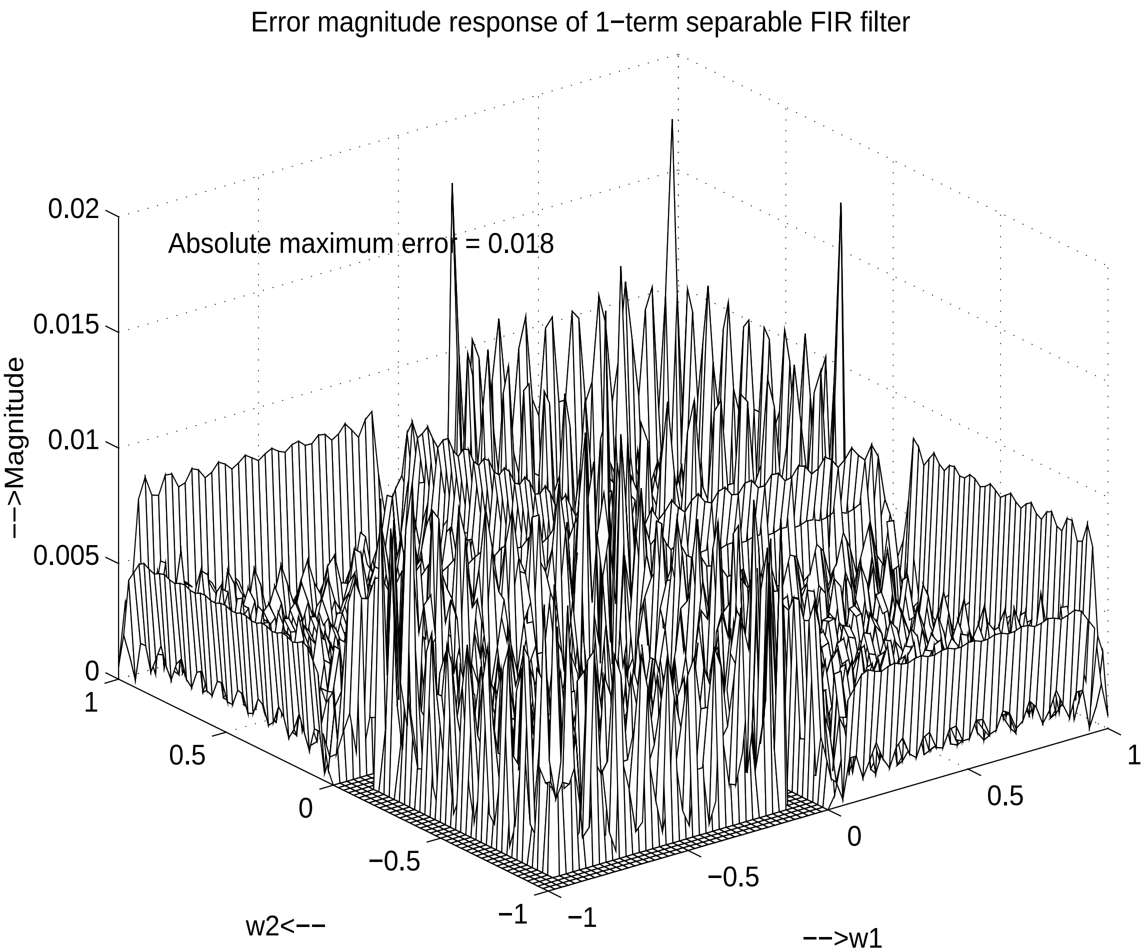

Figure 5.4: Magnitude error frequency response of the optimal 1-term FIR separable one quadrant fan filter

error is less than 10%, even though only very few coefficients have been used. The variation of the cost function and the number of iterations required for each term are shown in Figure 5.10. This example shows that superior designs, with respect to the optimal non-separable, can be obtained using only a few separable terms, with only an average of 20 iterations, and $N_1 + N_2 + 2 = 46$ independent parameters per term.

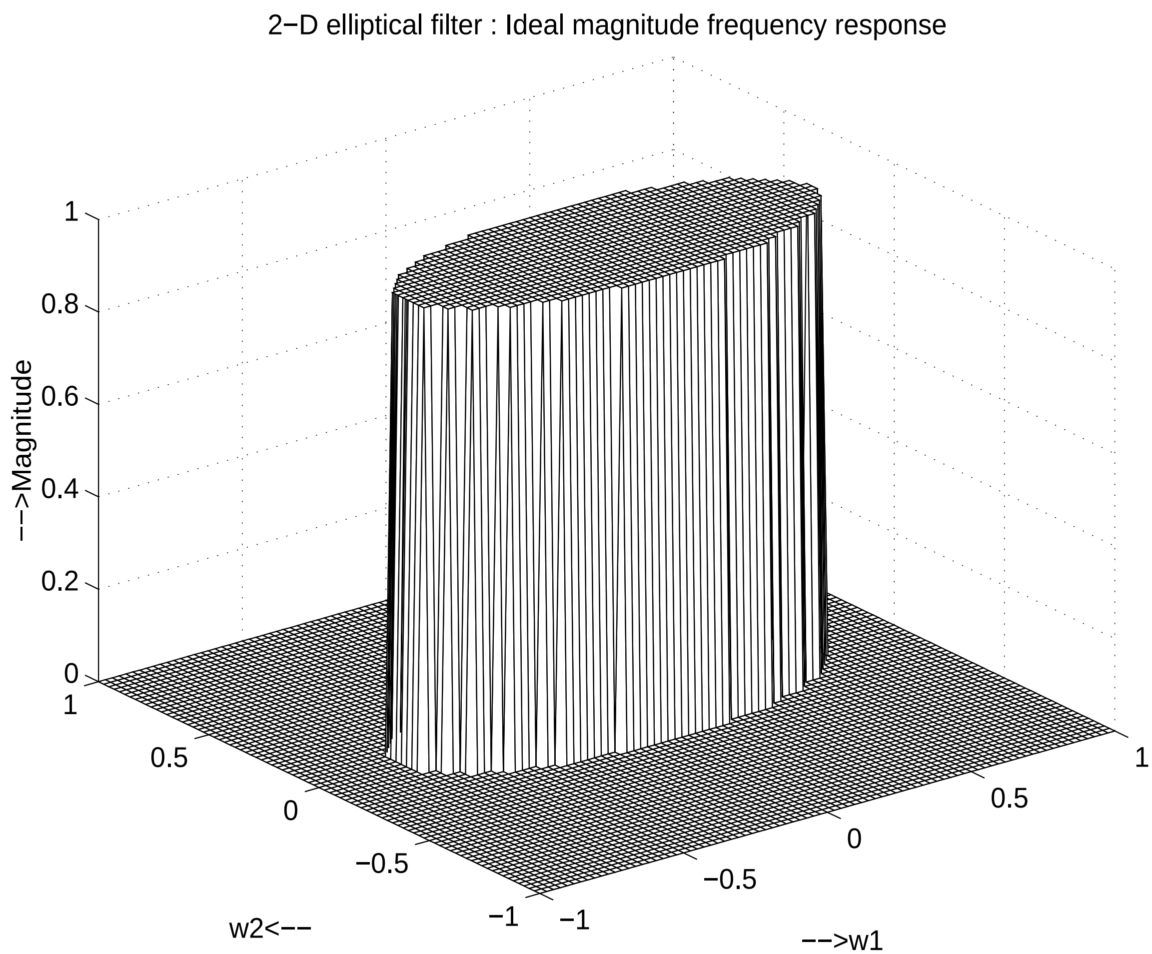

Figure 5.5: Ideal magnitude frequency response of 2-D elliptical filter

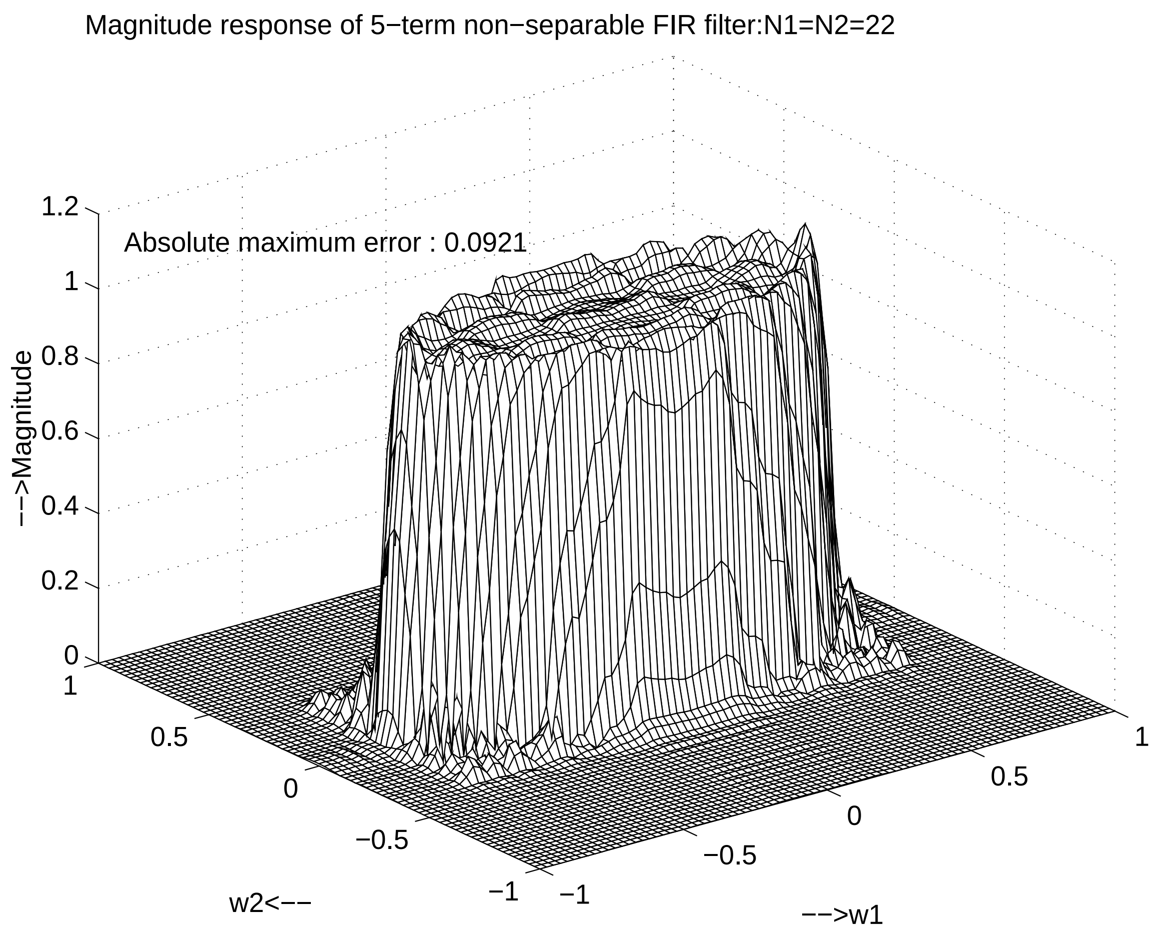

Figure 5.6: Magnitude frequency response of the optimal 5-term non-separable FIR elliptical filter

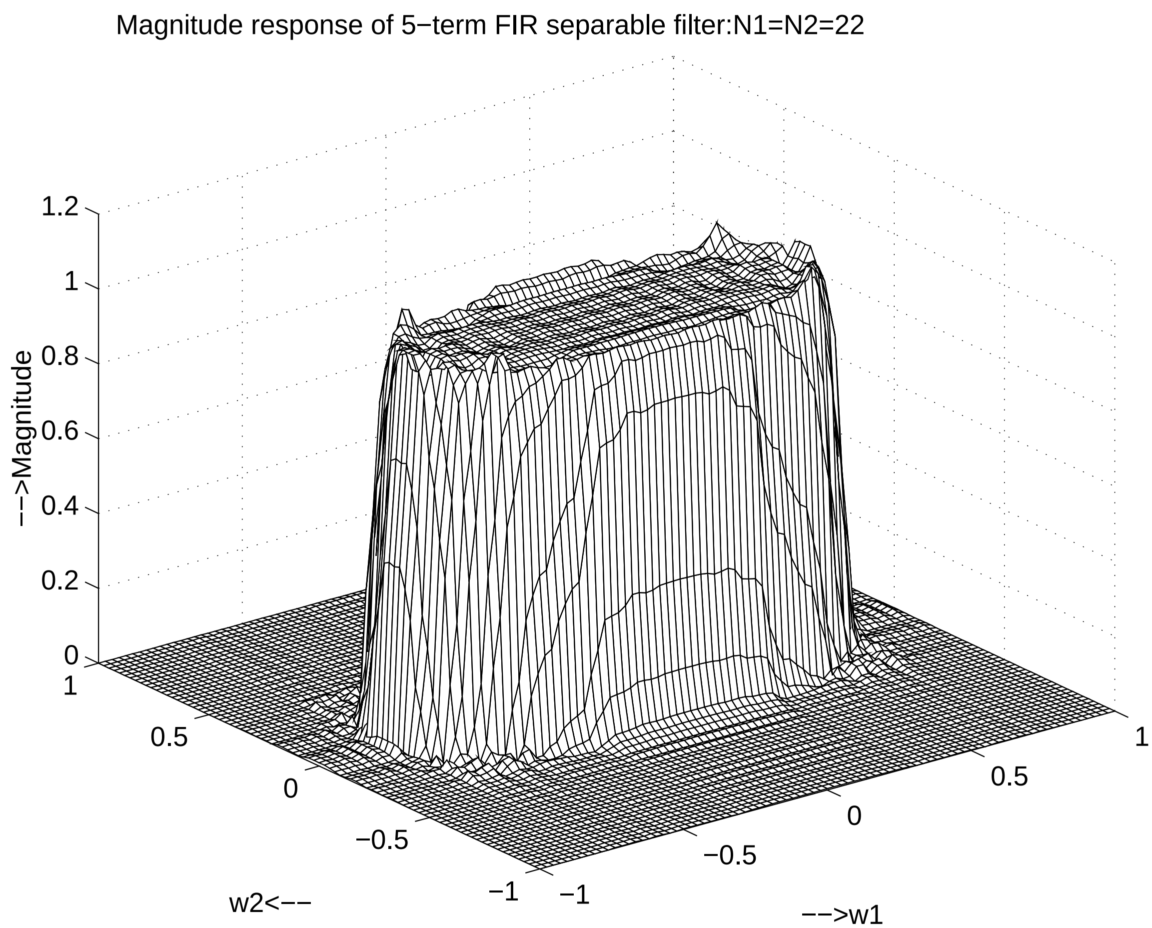

Figure 5.7: Magnitude frequency response of the optimal 5-term separable FIR elliptical filter

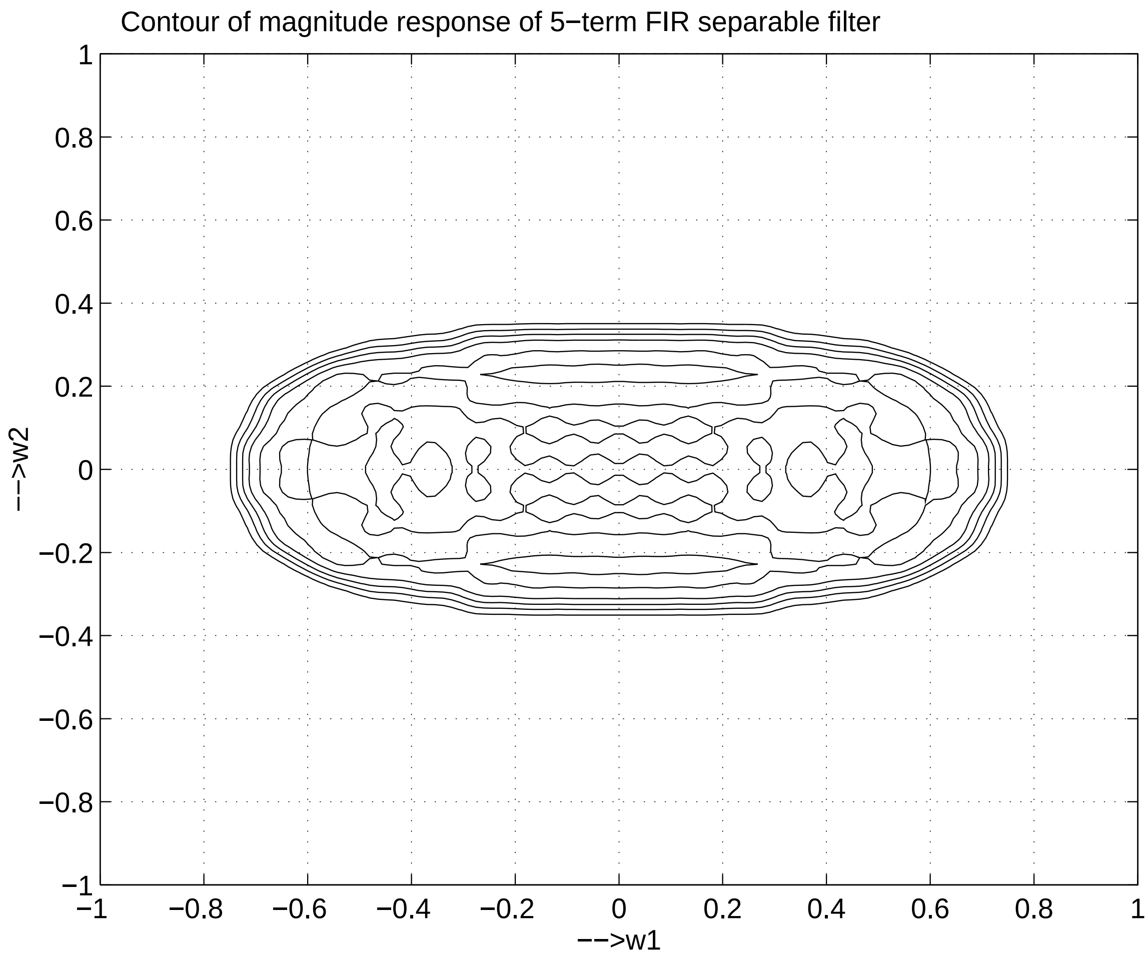

Figure 5.8: Contour of the magnitude frequency response of the optimal 5-term separable FIR elliptical filter

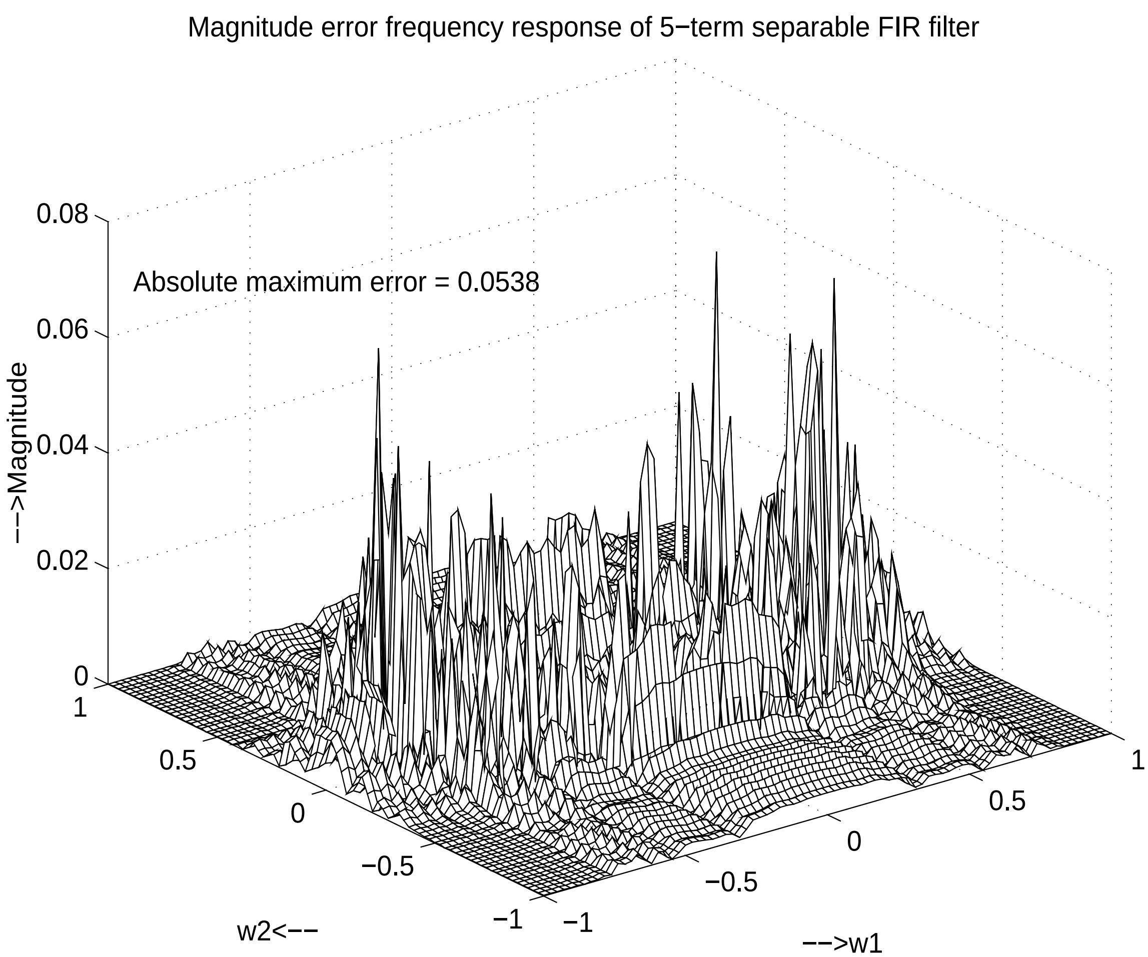

Figure 5.9: Magnitude error frequency response of the optimal 5-term separable FIR elliptical filter

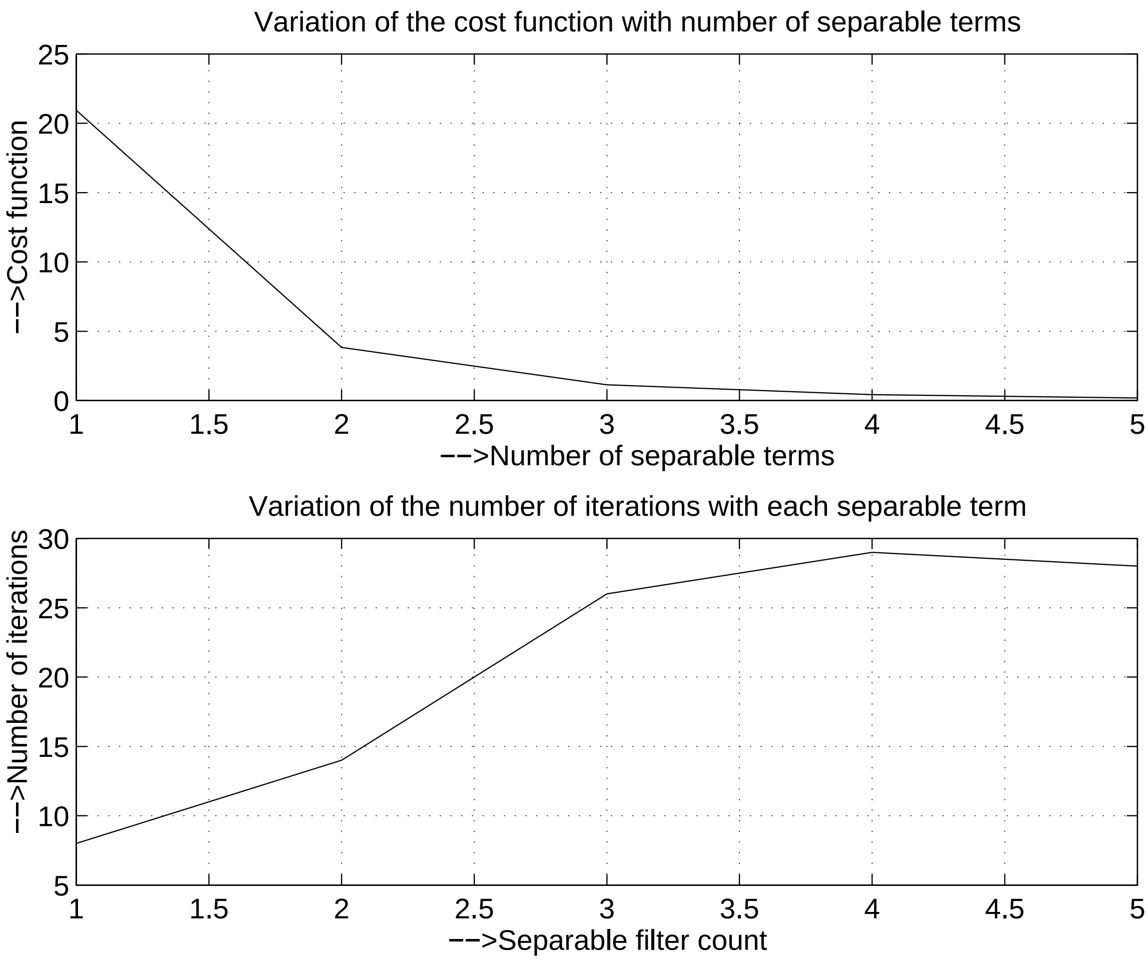

Figure 5.10: Some design results of the optimal 5-term separable elliptical filter

Example 3

Example 3 is a more *difficult* filter to design using this algorithm in that it is almost completely non-separable. The ideal frequency response is shown in Figure 5.11, and it is a half-plane rotated elliptical filter with axes of $.7\pi$ rad and $.3\pi$ rad and an external transition band of width $.1\pi$ rad and rotated by 30^o counterclockwise about the w_1 axis. The optimal 11-term non-separable response is shown in Figure 5.12. This filter has properties such that the Corollary to Theorem 4.1 stated in Appendix C can be applied. It is therefore designed in a fashion similar to the design of the elliptical filter with all coefficients constrained to be real. The optimal 11-term separable response is shown in Figure 5.13, and the corresponding error response is shown in Figure 5.14. This filter was originally designed using 15 terms, and the second level SVD was then used to obtain a reduction of 4 terms. Thus, one can observe that the *SVD does not produce any significant reduction in the number of channels in this case.* From the magnitude responses, it can be seen that this filter compares favourably with the non-separable response using an equivalent number of terms. Thus, even for almost completely non-separable filter configurations, better quality designs are obtained using this algorithm than an equivalent number of optimal non-separable terms.

5.1.2 Performance variation with parameters

The parameters affecting the filter performance, which one can represent here as the cost function, are affected by *the filter orders in the two directions and also the number of separable terms used.* For a fixed number of coefficients, there are several different combinations of filter order and number of terms possible. Of these possible combinations, there might exist one which results in the *minimum cost.* The aim in this section is to study the map of the performance variation with respect to the

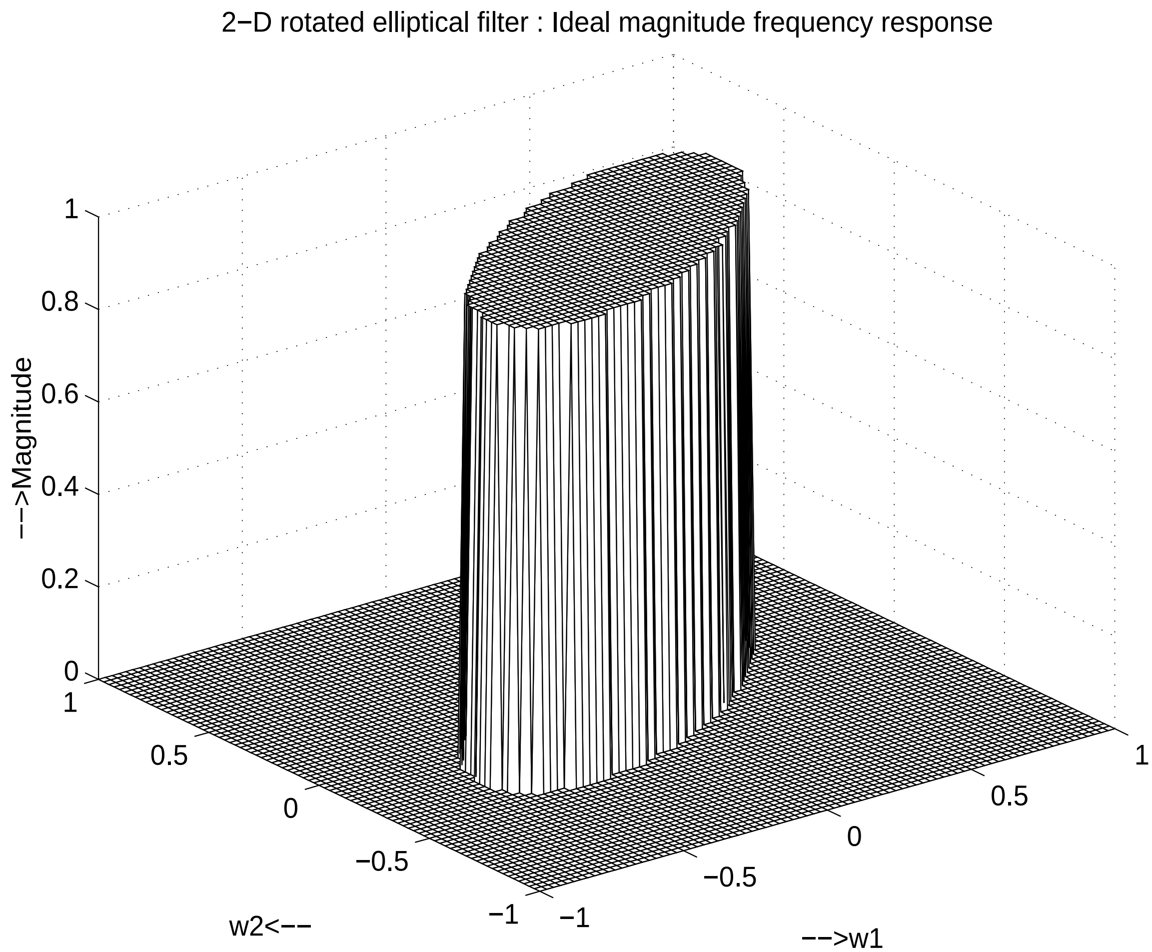

Figure 5.11: Ideal magnitude frequency response of 2-D rotated elliptical filter

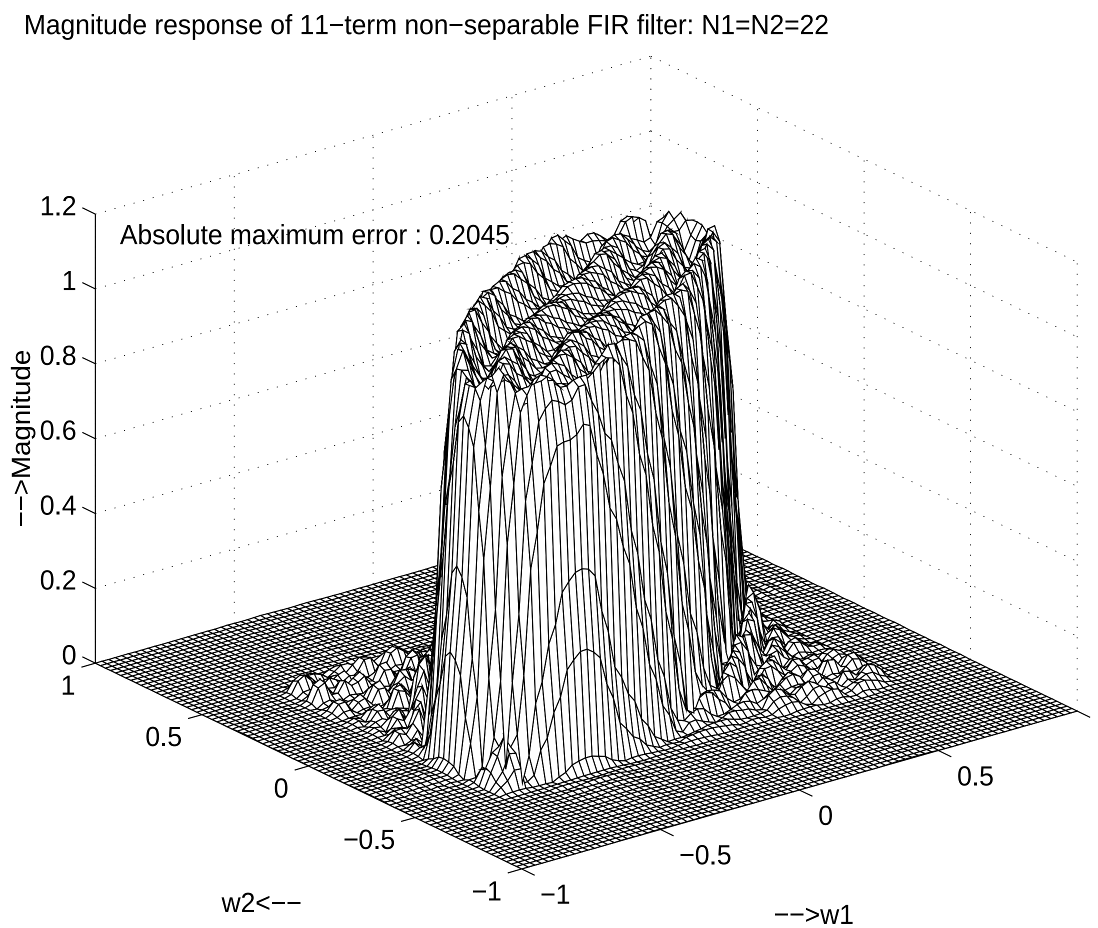

Figure 5.12: Magnitude frequency response of the optimal 11-term non-separable FIR rotated elliptical filter

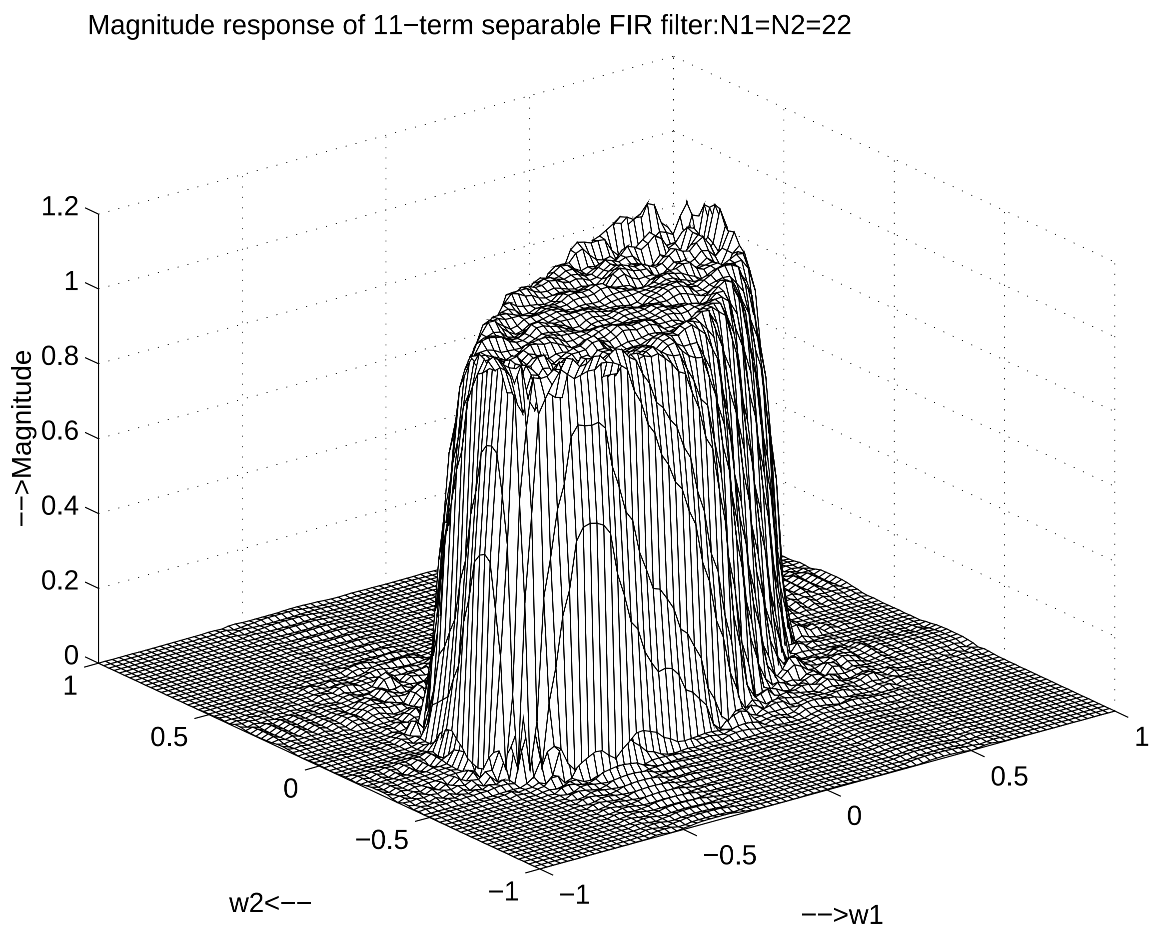

Figure 5.13: Magnitude frequency response of the optimal 11-term separable FIR rotated elliptical filter

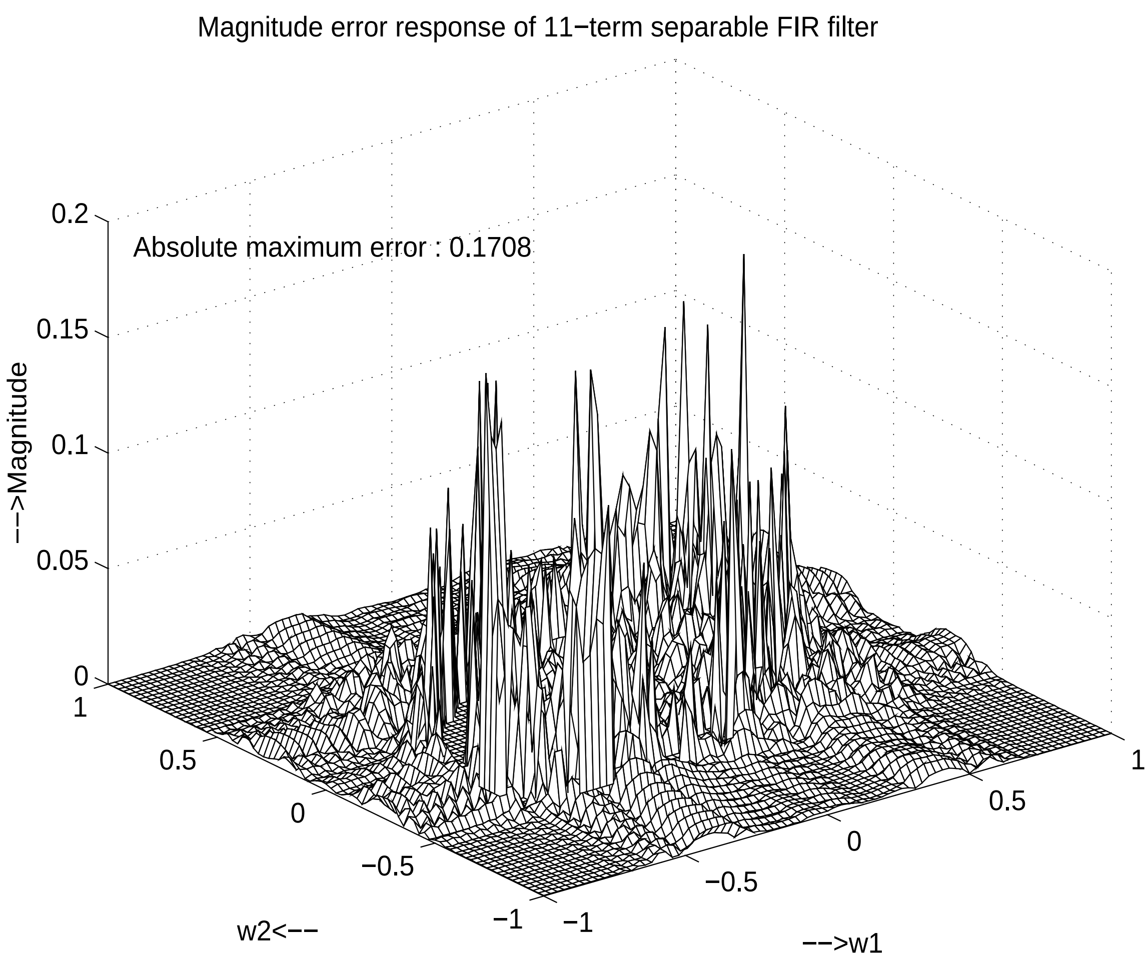

Figure 5.14: Magnitude error frequency response of the optimal 11-term separable FIR rotated elliptical filter

two parameters, for several different coefficient values, in order to determine the optimal combination. The filter orders in the w_1 and w_2 directions are assumed equal, as also the filter order in every term. The map is drawn for the elliptical filter of Example 2, and the coefficients are fixed at 100, 200, ..., 800.

For each given fixed value of coefficients, the number of terms and filter order are varied in order to determine the optimal combination of filter order and number of terms that produces the smallest cost function. The optimal combination values of filter order and number of separable terms, together with the corresponding cost, are tabulated in Table 5.1. It can be observed from the table that the optimal

Table 5.1: Optimal combination values of filter order N and number of separable terms q for minimum cost

No. of coefficients	N	q	Cost
100	12	4	1.5089
200	16	6	0.5284
300	18	8	0.2267
400	19	10	0.1258
500	20	12	0.0849
600	22	13	0.0528
700	20	17	0.0397
800	23	17	0.0320

filter order is around 20 in each direction, with the number of terms varying rather widely, but with a value around 12. Quite naturally, the cost function decreases as the number of coefficients increases, as can be concluded from Figures 5.15 and 5.16. Figure 5.17 shows the variation of the minimum value of the cost with the number of independent coefficients used for the design. It can be observed from the graph that no significant improvement is achieved by increasing the number of coefficients beyond about 500. This result can be used to determine the *optimal combination of number of coefficients, number of separable terms, and order of the* $1-D$ *filters in* *each term*, for each design. This map was obtained without the further reduction using the SVD, which was explained in Chapter 4.

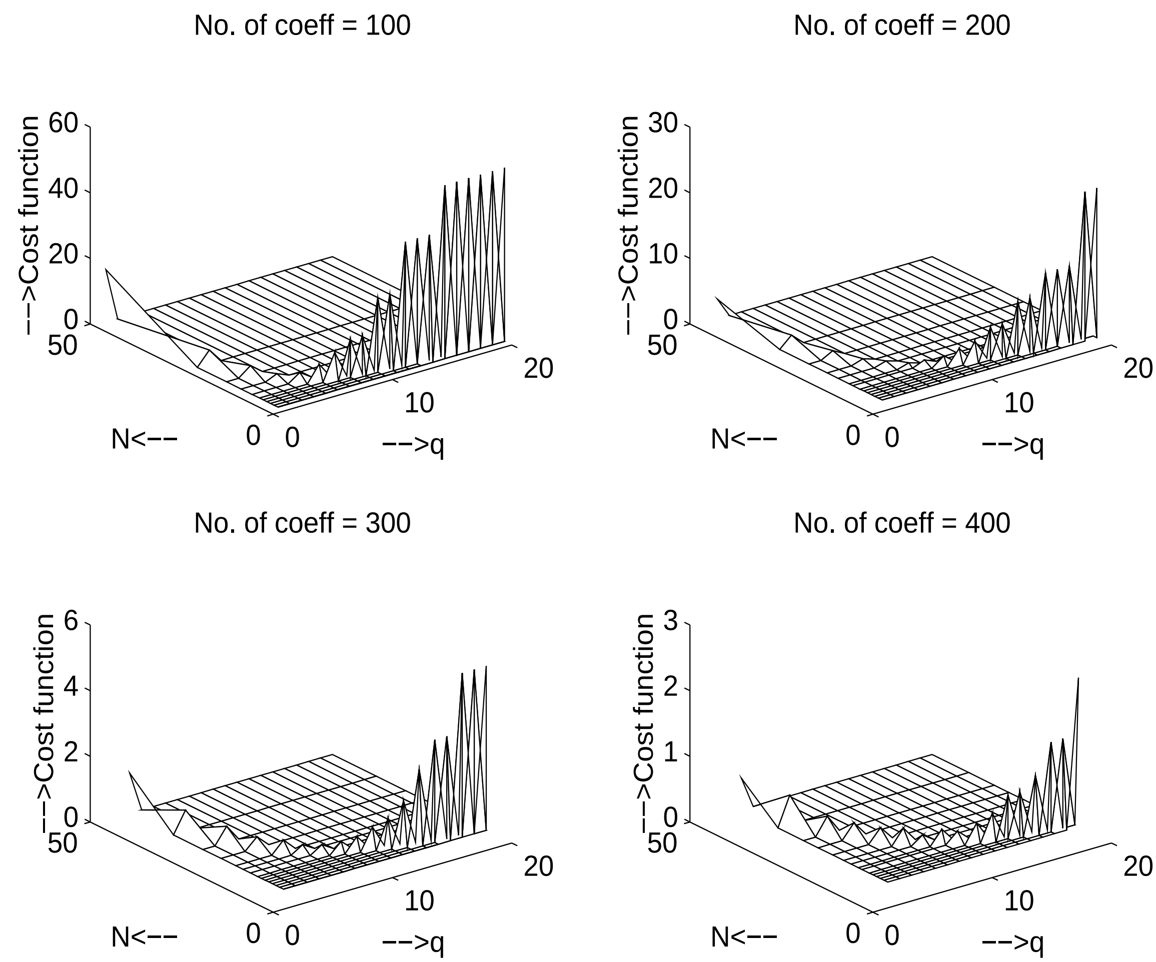

Figure 5.15: Cost function values with variation of filter order and number of channels

5.1.3 Computational complexity considerations of the algorithm

Now, the most serious problem in applying optimization techniques for the design of FIR filters is the fact that these filters have very low *selectivity*. This implies that even a moderately demanding application requires a very high order filter i.e. a large

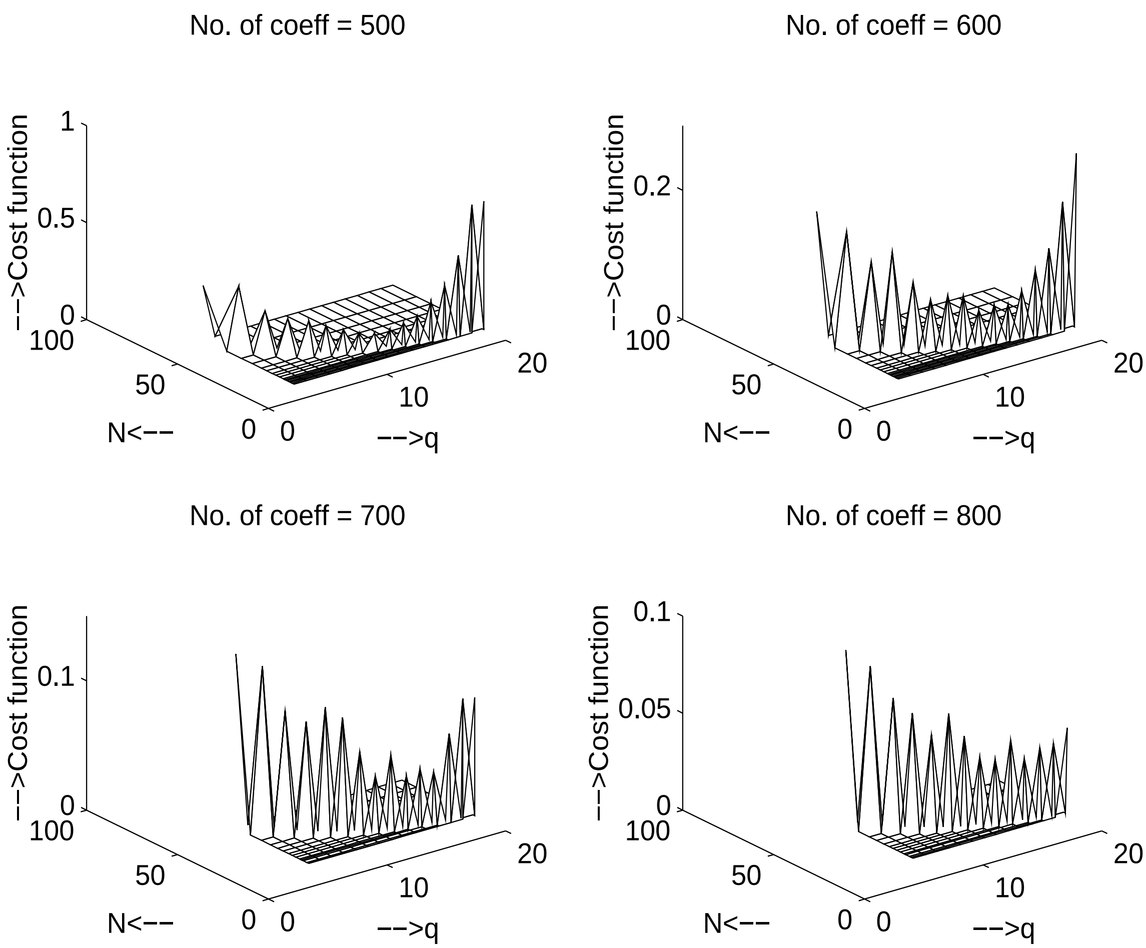

Figure 5.16: Cost function values with variation of filter order and number of channels

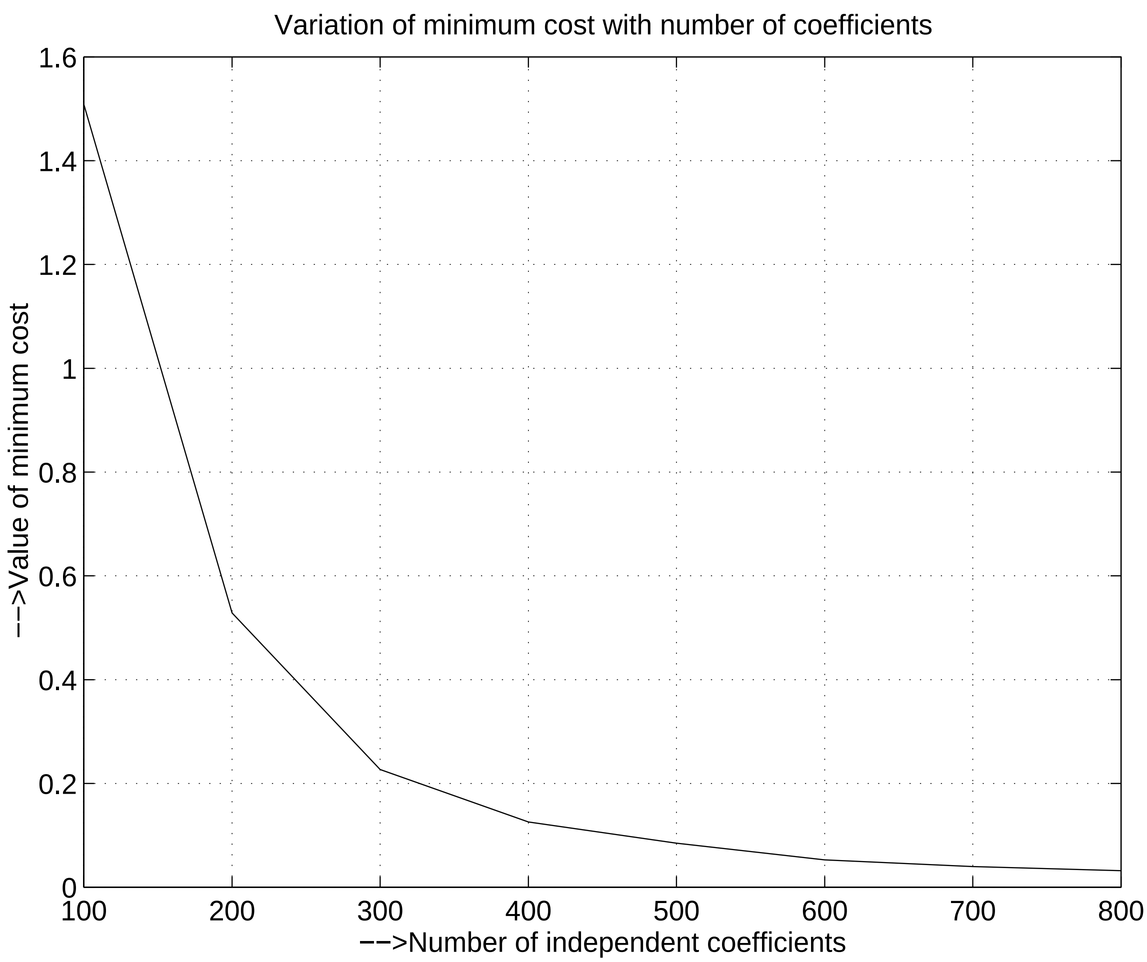

Figure 5.17: Variation of the minimum cost with the number of independent coefficients

number of coefficients. This problem can be reduced by minimizing the number of *independent parameters*, and this, in turn, reduces the computational complexity of the algorithm. This section presents a simplistic analysis of the complexity of this algorithm as a function of the number of independent coefficients, when compared to the weighted least mean square non-separable algorithm.

Consider first the $2-D$ non-separable FIR filter design algorithm ([4]). For a filter of order N_1 and N_2 in the w_1 and w_2 directions respectively, the number of free parameters is $p_f = 2N_1N_2 + N_1 + N_2 + 1$ and one is required to solve a linear system of equations. Hence one can estimate the complexity as $\mathcal{O}(p_f^2)$. In the separable filter design, the number of such parameters is only $p_f = 2N_1 + 2N_2 + 2$ *for each separable filter*. Assuming initially a similar complexity, and $N = N_1 = N_2$, one can see that the two techniques will have the same overall complexity if the single separable filter is solved approximately $N^2/4$ times. This is of course a very simplistic estimate but points out the well known fact that oftentimes the divide-and-conquer-strategy leads to more efficient algorithms. A more complete complexity analysis needs to take into consideration the fact that the non-separable case is a linear problem while the new method requires the solution of non-linear equations. Even with the more complete complexity analysis, one would expect, and the experiments confirm, computational advantages in the separable design. It is to be noted that for an order of say $N = 20$, $N^2/4 = 100$, which is almost 10 times the number of separable filters needed for most filters in order to achieve a reasonable design. Hence, the overall complexity of the algorithm is, at worst, no larger than that of the optimal non-separable. For filters having some special symmetry features, the number of independent parameters is reduced by almost half. In particular, quadrantally symmetric and antisymmetric filters are designed using only N_1+N_2+2

free parameters, as against $2N_1N_2 + N_1 + N_2 + 1$ in the non-separable case. Most half-plane symmetric filters require only $N_1 + N_2 + 1$ independent coefficients for each term on an average, since they are considered in pairs. This results in a large reduction in computational complexity for such filters, when designed using this algorithm.

Chapter 6

Conclusions

The design of fast-acting $2-D\ FIR$ digital filters was the objective of this research effort. The aim was to formulate an algorithm to design a *separable 2-D FIR filter that is especially suited to real time processing applications, and lends itself readily to a parallel hardware implementation.* The attempt was successful in that such a rigorous algorithm was devised after a complete mathematical analysis, and it accommodates a parallel architecture. In addition, this new technique permits more flexible designs and, according to the simulation results, one can get acceptable performance with less number of parameters.

At first, an extensive study of some common filter types was done, and some general observations on the behaviour of the individual filter components were made. These were used later, in the actual design and simulation, to simplify and speed up the process. An example of this is the reduction of the *weight map* $\mathcal{W}$ using the idea of the truncation of the SVD matrices outlined in Chapter 2, to reduce computational effort. Another example is in the design of filters with certain special kinds of symmetry, discussed in Chapter 2. The study sheds light on the conditions under which constraints can be incorporated in certain filter configurations to speed up the design process, and provide for realizable designs. The constraint of real coefficients, and that of considering coefficients as sums of conjugate pairs are some examples.

The simulation results demonstrate the significant gains to be achieved, when using this algorithm to design $2-D$ separable FIR filters, especially in cases of symmetrical filter shapes. The examples of the half - plane symmetric, almost separable one quadrant fan filter, and quadrantally symmetric elliptical filter, show that excellent quality designs are achievable with very little computational effort. Also, these designs are perfect for a fast, and simple implementation owing to very few coefficients, and a small number of channels. For more complicated shapes, which are almost completely nonseparable (*e.g.* rotated elliptical filter), better designs are obtained with this approach than equivalent designs obtained with the nonseparable design algorithm.

For shapes which are neither symmetric nor almost separable, however, a suitable decomposition of the ideal response needs to be done before designing, in order to avoid excessive computation time, and the SVD of the final coefficient matrix is used to obtain the separable filter coefficients. This entire process involves the solving of the necessary conditions for *each* component obtained from the decomposition, and the final results may not be the most optimal achievable. Also, it has not been conclusively proved that the optimal *n term* approximation obtained using this approach of successive approximation, is equivalent to the optimal *n term* approximation, obtained as a whole without adding terms successively. Again, there are several possible decompositions which result in separable components, different from the SVD used here. Examples of these are the LU decomposition, and the decomposition using *Interlocking Factors*. A comparative study of the use of such alternate decompositions as against the use of the SVD, might lead to better, or more efficient, designs. The above issues need to be considered in greater depth, and a more comprehensive computational complexity analysis of this algorithm should be done, to complete this research effort. Possible extensions to the design of $2-D$ separable IIR filters is also a topic for further research in this area.

To conclude, an algorithm to design $2-D$ separable FIR filters has been formulated. The significant features of this algorithm are :

• It offers great flexibility in design since the length of each $1-D$ filter can be selected independently, as also the number of such filters to be used.

• Simpler and faster filtering action implementation owing to the parallel structure.

• Greatly reduced computational complexity for filters having some special symmetry properties.

• This approach thus integrates $2-D$ filter design and parallel implementations to create filters with high throughput.

Bibliography

[1] J.L. Aravena,"Triple Matrix Product Architectures for Fast Signal Processing," *IEEE Trans. Circ. Syst.,* Vol. 35, No. 1, pp 119-122, Jan. 1988.

[2] M.O. Ahmad, and J.D. Wang, "An analytic least mean square solution to the design problem of two-dimensional FIR filters with quadrantally symmetric or antisymmetric frequency response," *IEEE Trans. Circuits and Systems*, Vol CAS-36, pp 968-979, July 1989.

[3] G. Gu, and J.L. Aravena, "Weighted Least Mean Square Design of 2-D FIR Digital Filter," *IEEE Trans. Sig. Proc.*, (to appear in Nov. 94).

[4] J.L. Aravena, and G. Gu, "Weighted Least Mean Square Design of 2-D FIR Digital Filter : General Weighting Function," *Proc. 27th csilomar Conf. Syst. Sig. & Comp.*, Asilomar, CA, Nov. 1993, pp 711-714.

[5] Wu-Sheng Lu, and Andreas Antoniou, "Two-Dimensional Digital Filters," Marcel Dekker,Inc., New York, 1992.

[6] 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.

[7] W.S. Lu, H.P. Wang, and A. Antoniou, "Design of two-dimensional FIR Digital Filters by using the Singular-Value Decomposition," *IEEE Trans. Circuits and Systems*, Vol. 37, No. 1, pp 35-46, Jan. 1990.

[8] W.S. Lu, H.P. Wang, and A. Antoniou, "Design of two-dimensional Digital Filters using Singular-Value Decomposition and Balanced Approximation Method," *IEEE Trans. on Signal Processing*, Vol. 39, No. 10, pp 2253-2262, Oct. 1991.

[9] S. Treitel, and J.L. Shanks, "The design of multistage separable planar filters," *IEEE Trans. Geosci. Electron.*, Vol. GE-9, pp 10-27, Jan. 1971.

[10] P. Karivaratharajan, and M.N.S. Swamy, "Realization of a 2-dimensional FIR digital filter using separable filters," *Electron. Lett.*, Vol. 14, No. 8, pp 244-251, April 1978.

[11] B.R. Suresh, and B.A. Shenoi, "Exact realization of 2-dimensional digital filters by separable filters," *Electron. Lett.*, Vol. 12, No. 10, pp 242-244, 1976.

[12] 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.

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

[14] T.S. Huang, "Two-dimensional windows," *IEEE Trans. Audio Electroacoust.*, Vol. AU-20, pp 88-89, March 1972.

[15] A. Antoniou, and W.S. Lu, "Design of 2-D nonrecursive filters using window method," *IEE Proc.*, Vol. 137, pt. G, pp 247-250, Aug. 1990.

[16] 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 Process.*, Vol. ASSP-29, pp 125-127, Feb. 1981.

[17] J.F. Kaiser, "Nonrecursive digital filter design using the $I_0 - sinh$ window function," *Proc. 1974 IEEE Int. Symp. Circuit Theory*, pp 20-23.

[18] W. Chen, C.H. Smith, and S.C. Fralick, "A fast computational algorithm for the discrete cosine transform," *IEEE Trans. Commun.*, Vol. COM-25, pp 1004-1009, Sep. 1977.

[19] J. Makhoul, "A fast cosine transform in one and two dimensions," *IEEE Trans. Acoust., Speech, and Signal Process.*, Vol. ASSP-28, pp 27-34, Feb. 1980.

[20] A. Gupta, and K.R. Rao, "A fast Recursive Algorithm for the Discrete Sine Transform," *IEEE Trans. Acoust., Speech, Signal Process.*, Vol. ASSP-38, No. 3, pp 553, March 1990.

[21] A. Gupta, and K.R. Rao, "An Efficient Algorithm Based on the Discrete Sine Transform," *IEEE Trans. on Signal Processing*, Vol. 39, No. 2, pp 486, February 1991.

[22] J.H. McClellan, "The design of two-dimensional filters by transformations," *Proc. 7 th Annual Princeton Conf. Information Sciences and Systems*, pp 247-251, 1973.

[23] R.M. Mersereau, W.F.G. Mecklenbräuker, 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.

[24] Y. Kamp, and J.P. Thiran, "Chebyshev approximations for two-dimensional nonrecursive digital filters," *IEEE Trans. Circuits Syst.*, Vol. CAS-22, pp 208-218, March 1975.

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

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

[27] L.R. Rabiner, J.H. McClellan, and T.W. Parks, "FIR digital filter design techniques using weighted Chebyshev approximation," *Proc. IEEE*, Vol. 63, pp 595-610, April 1975.

[28] C. Charalambous, "A unified review of optimization," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-22, pp 289-300, March 1974.

[29] C. Charalambous, "Acceleration of the least p th algorithm for minimax optimization with engineering applications," *Mathematical Programming*, Vol. 17, pp 270-297, 1979.

[30] C. Charalambous, "The performance of an algorithm for minimax design of two-dimensional linear phase FIR digital filters," *IEEE Trans. Circuits Syst*, Vol. CAS-32, pp 1016-1028, Oct. 1985.

[31] D.J. Shpak and A. Antoniou, "A generalized Remez method for the design of FIR digital filters," *IEEE Trans. Circuits Syst.*, Vol. CAS-37, pp 161-174, Feb. 1990.

Appendix A

Theorem : *Suppose* $D \in E^{M_1 \times M_2}$, $p \in E^{M_1 \times 1}$, $q \in E^{M_2 \times 1}$, *and suppose* $J = tr\{(D - pq^*)(D - pq^*)^*\}$, *defines the merit index, then the minimizing solution to this is obtained from the SVD of D.*
Proof :

$$J = tr\{(D - pq^*)(D - pq^*)^*\} \quad (A.1)$$

$$= tr\{(D - pq^*)(D^* - qp^*)\} \quad (A.2)$$

Let q be fixed at q_0. Then,

$$\delta J_p = tr\{(D - pq_0^*)(-q_0)\delta p^*\} + tr\{(-\delta p)q_0^*(D^* - q_0 p^*)\}$$

Equating δJ_p to 0, we obtain,

$$tr\{(D - pq_0^*)(-q_0)\delta p^*\} + tr\{(-\delta p)q_0^*(D^* - q_0 p^*)\} = 0$$

$$\Rightarrow tr\{(D - pq_0^*)(-q_0)\delta p^*\} + tr\{((D - pq_0^*)(-q_0)\delta p^*)^*\} = 0$$

$$\Rightarrow 2tr\{(D - pq_0^*)(-q_0)\delta p^*\} = 0$$

$$\Rightarrow tr\{(Dq_0 - pq_0^* q_0)\}\delta p^* = 0 \quad (A.3)$$

$\Rightarrow \{(Dq_0 - pq_0^* q_0)\}$ is orthogonal to δp^*.
Since δp is arbitrary,

$$\Rightarrow (Dq_0 - pq_0^* q_0) = 0$$

$$\Rightarrow (Dq_0 = pq_0^* q_0)$$

Now, $q_0^* q_0$ is a scalar $= \| q_0 \|^2$

$$Dq_0 = p \| q_0 \|^2$$

Hence,

$$p_{opt} = \frac{Dq_0}{\| q_0 \|^2} \quad (A.4)$$

From eqs. A.1, A.2 and A.4,

$$\hat{J} = tr\{(D - p_{opt}q^*)(D^* - qp_{opt}^*)\}$$

$$= tr\{(D - p_{opt}q^*)D^* - (D - p_{opt}q^*)qp_{opt}^*)\}$$

From eq. A.3, $tr\{(Dq_0 - pq_0^* q_0)\}\delta p^* = 0$ when $p = p_{opt}$. Hence,

$$\hat{J} = tr\{(D - p_{opt}q^*)D^*\} \quad (A.5)$$

$$= tr\{(DD^* - p_{opt}q^* D^*)\}$$

The minimum of this is :

$$J_{min} = max(tr\{p_{opt}q^*D^*\}) \tag{A.6}$$

$$= max(tr\{\frac{Dq(Dq)^*}{\| q \|^2}\})$$

$$= max(tr\{\frac{(Dq)^*Dq}{\| q \|^2}\}),$$

which is a scalar. Hence,

$$J_{min} = \frac{(Dq)^*Dq}{\| q \|^2} = \frac{q^*D^*Dq}{\| q \|^2} \tag{A.7}$$

Eq. A.7 is the representation of the SVD of D. Thus, the minimizing solution to the merit index defined in eq. A.1 is obtained from the SVD of D.

If D represents the ideal filter frequency response matrix, then eq. A.1 represents the cost function to be minimized in the least mean squares sense without any constraints. The vectors p, q represent the $1 - D$ FIR filters in the w_1 and the w_2 directions, and are optimized using the SVD of D.

Appendix B

Theorem : *The system of equations defined by eqs. 3.39 and 3.40 has a limiting solution if and only if there exists a unique optimal non-separable FIR filter.*
Proof :
Suppose the original general non-separable $2-D\ FIR$ filter minimization problem with cost function defined by

$$J(A) = tr\{(W \bullet (D - \Omega_1 A \Omega_2^T))(D - \Omega_1 A \Omega_2^T)^*\}$$

has a unique solution, given by $A = A_0$. Following the development in Chapter 3, A_0 must satisfy the following necessary condition for the minimum cost function.

$$\mathcal{F}^*\mathcal{W}(D) = \mathcal{F}^*\mathcal{W}\mathcal{F}(A_0) \tag{B.1}$$

Hence, A_0 exists and is unique, if and only if $\mathcal{F}^*\mathcal{W}\mathcal{F}$ is an invertible map. Since it is clearly self-adjoint,

$$\mathcal{F}^*\mathcal{W}\mathcal{F} > 0 \tag{B.2}$$

In particular, *the matrix representation of $\mathcal{F}^*\mathcal{W}\mathcal{F}$ is positive definite.*
The separable case considered here, is characterized by the condition that the A in the minimization problem is replaced by ab^T. The cost function can therefore be written as :

$$J(a,b) = \left\langle \mathcal{W}(D - \mathcal{F}(ab^T)), (D - \mathcal{F}(ab^T)) \right\rangle \tag{B.3}$$

$$= \langle \mathcal{W}(D), D\rangle - \left\langle ab^T, \mathcal{F}^*\mathcal{W}(D)\right\rangle - \left\langle \mathcal{F}^*\mathcal{W}(D), ab^T\right\rangle$$

$$+ \left\langle \mathcal{F}^*\mathcal{W}(\mathcal{F}(ab^T)), ab^T\right\rangle$$

For a fixed vector $a \in E^{N_1}$, this is a minimization with respect to $b \in E^{N_2}$. Using the notation defined in Chapter 3, one can explicitly state this problem as a conventional quadratic minimization problem. Let $J_0 = \langle \mathcal{W}(D), D\rangle$. From eqs. 3.11 and 3.13,

$$J(a,b) = J_0 - \langle \mathcal{S}^*(a \otimes b), \mathcal{F}^*\mathcal{W}(D)\rangle - \langle \mathcal{F}^*\mathcal{W}(D), \mathcal{S}^*(a \otimes b)\rangle$$

$$+ \langle \mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*(a \otimes b), \mathcal{S}^*(a \otimes b)\rangle \tag{B.4}$$

From eq. 3.15,

$$J(a,b) = J_0 - \langle \mathcal{S}^*((a \otimes I_{N_2})b), \mathcal{F}^*\mathcal{W}(D)\rangle - \langle \mathcal{F}^*\mathcal{W}(D), \mathcal{S}^*((a \otimes I_{N_2})b)\rangle$$

$$+ \langle \mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*((a \otimes I_{N_2})b), \mathcal{S}^*((a \otimes I_{N_2})b)\rangle$$

$$J(a,b) = J_0 - \langle ((a \otimes I_{N_2})b), \mathcal{S}\mathcal{F}^*\mathcal{W}(D)\rangle - \langle \mathcal{S}\mathcal{F}^*\mathcal{W}(D), ((a \otimes I_{N_2})b)\rangle$$

$$+ \langle \mathcal{S}\mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*((a \otimes I_{N_2})b), ((a \otimes I_{N_2})b)\rangle$$

$$J(a,b) = J_0 - \langle b, (a^* \otimes I_{N_2})\mathcal{S}\mathcal{F}^*\mathcal{W}(D)\rangle - \langle (a^* \otimes I_{N_2})\mathcal{S}\mathcal{F}^*\mathcal{W}(D), b\rangle$$
$$+ \langle (a^* \otimes I_{N_2})\mathcal{S}\mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*((a \otimes I_{N_2})b), b\rangle \tag{B.5}$$

For fixed $a_0 \in E^{N_1}$, this is a standard minimization problem with respect to $b \in E^{N_2}$ and it is in quadratic form. The solution b_1 exists if and only if

$$(a_0^* \otimes I_{N_2})\mathcal{S}\mathcal{F}^*\mathcal{W}(D) = (a_0^* \otimes I_{N_2})\mathcal{S}\mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*(a_0 \otimes I_{N_2})b_1 \tag{B.6}$$

Let

$$Q_a = (a_0^* \otimes I_{N_2})\mathcal{S}\mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*(a_0 \otimes I_{N_2})$$

Therefore, b_1 exists if and only if Q_a is invertible. Now, for $a_0 \neq 0, b_1 \neq 0$, Q_a is invertible if and only if $\mathcal{F}^*\mathcal{W}\mathcal{F}$ is invertible, since,

$$\langle b, Q_a b\rangle = \langle b, (a_0^* \otimes I_{N_2})\mathcal{S}\mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*(a_0 \otimes I_{N_2})b\rangle$$
$$= \left\langle ab^T, \mathcal{F}^*\mathcal{W}\mathcal{F}(ab^T)\right\rangle$$

$\mathcal{F}^*\mathcal{W}\mathcal{F}$ is positive definite by eq. B.2. Then

$$\left\langle ab^T, \mathcal{F}^*\mathcal{W}\mathcal{F}(ab^T)\right\rangle \quad > 0$$

and this implies, $Q_a > 0$. Hence, *i.e. the separable minimization problem has a solution with respect to b for fixed a.*
Now consider the minimization problem with respect to a for b fixed at b_1. The cost function, in this case, can be obtained from eqs. 3.12 and B.4 as :

$$J(a,b) = J_0 - \langle \mathcal{S}^*\mathcal{T}(b \otimes a), \mathcal{F}^*\mathcal{W}(D)\rangle - \langle \mathcal{F}^*\mathcal{W}(D), \mathcal{S}^*\mathcal{T}(b \otimes a)\rangle$$
$$+ \langle \mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*\mathcal{T}(b \otimes a), \mathcal{S}^*\mathcal{T}(b \otimes a)\rangle \tag{B.7}$$
$$= J_0 - \langle (b \otimes I_{N_1})a, \mathcal{T}^*\mathcal{S}\mathcal{F}^*\mathcal{W}(D)\rangle - \langle \mathcal{T}^*\mathcal{S}\mathcal{F}^*\mathcal{W}(D), (b \otimes I_{N_1})a\rangle$$
$$+ \langle \mathcal{T}^*\mathcal{S}\mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*\mathcal{T}(b \otimes I_{N_1})a, (b \otimes I_{N_1})a\rangle$$
$$= J_0 - \langle a, (b^* \otimes I_{N_1})\mathcal{T}^*\mathcal{S}\mathcal{F}^*\mathcal{W}(D)\rangle - \langle (b^* \otimes I_{N_1})\mathcal{T}^*\mathcal{S}\mathcal{F}^*\mathcal{W}(D), a\rangle$$
$$+ \langle (b^* \otimes I_{N_1})\mathcal{T}^*\mathcal{S}\mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*\mathcal{T}(b \otimes I_{N_1})a, a\rangle$$

This is also a standard minimization problem in quadratic form, but with respect to $a \in E^{N_1}$ with b fixed at $b_1 \in E^{N_2}$. From the earlier analysis, the solution a_1 to this problem exists if $(b_1^* \otimes I_{N_1})\mathcal{T}^*\mathcal{S}\mathcal{F}^*\mathcal{W}\mathcal{F}\mathcal{S}^*\mathcal{T}(b_1 \otimes I_{N_1})$ is invertible. As in the previous case, this is equivalent to $\mathcal{F}^*\mathcal{W}\mathcal{F}$ being invertible, which is true from eq. B.2. Hence, a_1 exists; *i.e. the separable minimization problem has a solution with respect to a for fixed b.*
Once it has been established that the system of equations defined by eqs. 3.21 and 3.22 has a solution for each iteration, one needs to establish next that the

solution is a limiting one i.e. the solution converges towards the optimal in each successive iteration. The approach adopted here is to show that the system defines an optimizing sequence of the cost function on the compact set given by $\| a \|= 1$. From eq. B.2, $\mathcal{F}^*\mathcal{W}\mathcal{F} > 0$ which implies,

$$\exists \mu > 0 \; such \;\; that \; \left\langle ab^T, \mathcal{F}^*\mathcal{W}\mathcal{F}(ab^T) \right\rangle \;\; \geq \mu \parallel ab^T \parallel^2 = \mu \parallel a \parallel^2 \parallel b \parallel^2$$

$$\Rightarrow \langle b, Q_a b \rangle \;\; \geq \mu \parallel a \parallel^2 \parallel b \parallel^2 \tag{B.8}$$

Let μ be the $min_k \sigma_k(\mathcal{F}^*\mathcal{W}\mathcal{F})$. For fixed $a = a_0$, b must satisfy eq. B.6. Hence,

$$b(a_0) = Q_{a_0}^{-1}(a_0^* \otimes I_{N_2})\mathcal{S}\mathcal{F}^*\mathcal{W}(D)$$

$$\Rightarrow \parallel b(a_0) \parallel^2 \leq \frac{1}{\mu^2 \parallel a_0 \parallel^4} \parallel (a_0^* \otimes I_{N_2}) \parallel^2 \parallel \mathcal{S}\mathcal{F}^*\mathcal{W}(D) \parallel^2$$

$$\leq \frac{c}{\mu^2 \parallel a_0 \parallel^2},$$

where c is a constant. This implies that b is bounded above and is therefore finite for any arbitrarily fixed a.

Now, suppose, $\parallel a_0 \parallel = 1$. One needs to minimize the cost function J for this fixed a over all possible b. Let $J_1^* = min_b J_{a_0}(b)$. Suppose the solution to this problem is $b = \hat{b_0}$, i.e.,

$$J(a_0, \hat{b_0}) = J_1^*$$

Now, b is fixed at $b = \hat{b_0}$, and the necessary condition for minimization of J over all possible a is solved. Let the solution to this problem be $a = \hat{a_0}$. Obviously,

$$J(\hat{a_0}, \hat{b_0}) = min_a J(a, \hat{b_0}) \leq J(a_0, \hat{b_0})$$

Set

$$a_1 = \frac{\hat{a_0}}{\parallel \hat{a_0} \parallel}, \;\; and \;\; b_1 = \hat{b_0} \parallel \hat{a_0} \parallel$$

Then $J(a_1, b_1) = J(\hat{a_0}, \hat{b_0}) \leq J(a_0, \hat{b_0})$. This process is applied iteratively on the equation set. Noting that $\parallel a_0 \parallel = 1$ and $\parallel a_1 \parallel = 1$, i.e. they are both on the unit sphere, we conclude, from the above analysis, that the procedure defines a sequence of vectors $\{a_k\}$ on the unit sphere and a sequence of monotonically decreasing costs. Since the unit ball is a compact set, we can extract a subsequence $\{a_{k_0}\}$ which is convergent, converging to a limit a^0. For this limit, the minimization with respect to b defined by $min_b J(a^0, b)$ must then yield a minimum cost for some $b = b^0$. The optimal solution yielding the minimum cost is thus (a^0, b^0), and this gives the best separable FIR approximation to the desired response in the weighted least mean squares sense.

Appendix C

Theorem : *Suppose that $\mathcal{W}(D)$ is real and half-plane symmetric, if (a, b) is an optimal solution to the cost function defined in eq. 4.1, then (a^c, b^c) is also an optimal solution.*
Proof :
Suppose the ideal response D is half-plane symmetric. Choose W such that W is also half - plane symmetric and real. Then,

$$\mathcal{W}(D) \ \ is \ \ real \ \ and \ \ half-plane \ \ symmetric \tag{C.1}$$

Suppose (a, b) are one set of optimal solutions for the approximation problem defined in eq. 4.1. Then, (a, b) must satisfy eqs. 3.39 and 3.40. From eqs. 3.36, 3.37, 3.38, 3.39, and 3.40,

$$Gb^c = \sum_{k=1}^{q} b^T R_k b^c L_k a = X(b)a \tag{C.2}$$

$$G^T a^c = \sum_{k=1}^{q} a^T L_k a^c R_k b = Y(a)b, \tag{C.3}$$

Taking conjugates of eqs. C.2 and C.3, we get,

$$G^c b = \sum_{k=1}^{q} b^* R_k^c b L_k^c a^c \tag{C.4}$$

$$G^* a = \sum_{k=1}^{q} a^* L_k^c a R_k^c b^c, \tag{C.5}$$

Let $\alpha = a^c, \beta = b^c$. Then, eqs. C.4 and C.5 can be written as :

$$G^c \beta^c = \sum_{k=1}^{q} \beta^T R_k^c \beta^c L_k^c \alpha = X(\beta)\alpha \tag{C.6}$$

$$G^* \alpha^c = \sum_{k=1}^{q} \alpha^T L_k^c \alpha^c R_k^c \beta = Y(\alpha)\beta, \tag{C.7}$$

The following steps show that $L_k^c = \pm L_k$, $R_k^c = \pm R_k$, and G is always real under the assumptions of this theorem. From eqs. 3.31, $L_k = \Omega_1^* L_{w_k} \Omega_1, \ \ R_k = \Omega_2^T R_{w_k} \Omega_2^c$. Consider any element of L_k, say $L_k(l_1, l_2)$.

$$L_k(l_1, l_2) = \sum_{k=-M_1}^{M_1-1} e^{\jmath\pi \frac{(2k+1)}{2M_1} l_1} L_{w_k}(k, k) e^{-\jmath\pi \frac{(2k+1)}{2M_1} l_2} \tag{C.8}$$

$$L_k(l_1, l_2) = \sum_{k=-M_1}^{M_1-1} e^{\jmath\pi\frac{(2k+1)}{2M_1}(l_1-l_2)} L_{w_k}(k, k) \tag{C.9}$$

From eq. 3.34

$$L_k^c(l_1, l_2) = L_k^T(l_1, l_2) = L_k(l_2, l_1)$$

Hence,

$$L_k^c(l_1, l_2) = \sum_{k=-M_1}^{M_1-1} e^{\jmath\pi\frac{(2k+1)}{2M_1}(l_2-l_1)} L_{w_k}(k, k) \tag{C.10}$$

Let $k' = -(k+1)$. Applying a change of variables, the above becomes,

$$L_k^c(l_1, l_2) = \sum_{k'=M_1-1}^{-M_1} e^{\jmath\pi\frac{(-2k'-2+1)}{2M_1}(l_2-l_1)} L_{w_k}(-k'-1, -k'-1) \tag{C.11}$$

$$L_k^c(l_1, l_2) = \sum_{k'=-M_1}^{M_1-1} e^{\jmath\pi\frac{(2k'+1)}{2M_1}(l_1-l_2)} L_{w_k}(-k'-1, -k'-1) \tag{C.12}$$

Since W is chosen to be real and half-plane symmetric, the matrices U_{w_k} and V_{w_k} resulting from the SVD of W are columnwise mirror image symmetric or antisymmetric. Therefore,

$$L_{w_k}(-k'-1, -k'-1) = \pm L_{w_k}(k', k')$$

Eq. C.12 then reduces to

$$L_k^c(l_1, l_2) = \pm \sum_{k'=-M_1}^{M_1-1} e^{\jmath\pi\frac{(2k'+1)}{2M_1}(l_1-l_2)} L_{w_k}(k', k') = \pm L_k(l_1, l_2) \tag{C.13}$$

by which one can conclude that $L_k^c = \pm L_k$. By a similar reasoning, $R_k^c = \pm R_k$. Therefore, eqs. C.6 and C.7 are equivalent to

$$G^c\beta^c = \sum_{k=1}^{q} \beta^T R_k \beta^c L_k \alpha = X(\beta)\alpha \tag{C.14}$$

$$G^*\alpha^c = \sum_{k=1}^{q} \alpha^T L_k \alpha^c R_k \beta = Y(\alpha)\beta, \tag{C.15}$$

Now, from eq. 3.36, $G = \Omega_1^* \sum_{k=1}^{q} L_{w_k} D R_{w_k} \Omega_2^c$. Consider any element of G, say $G(l_1, l_2)$.

$$G(l_1, l_2) = \sum_{k_2=-M_2}^{M_2-1} \sum_{k_1=-M_1}^{M_1-1} e^{\jmath\pi\frac{(2k_1+1)}{2M_1}l_1} (W \bullet D)(k_1, k_2) e^{\jmath\pi\frac{(2k_2+1)}{2M_2}l_2} \tag{C.16}$$

$$= \sum_{k_2=-M_2}^{M_2-1} \sum_{k_1=-M_1}^{M_1-1} e^{\jmath\pi(\frac{(2k_1+1)}{2M_1}l_1+\frac{(2k_2+1)}{2M_2}l_2)}(W \bullet D)(k_1,k_2) \tag{C.17}$$

Now consider the corresponding element of G^c.

$$G^c(l_1,l_2) = \sum_{k_2=-M_2}^{M_2-1} \sum_{k_1=-M_1}^{M_1-1} e^{-\jmath\pi\frac{(2k_1+1)}{2M_1}l_1}(W \bullet D)^c(k_1,k_2)e^{-\jmath\pi\frac{(2k_2+1)}{2M_2}l_2} \tag{C.18}$$

$$= \sum_{k_2=-M_2}^{M_2-1} \sum_{k_1=-M_1}^{M_1-1} e^{-\jmath\pi(\frac{(2k_1+1)}{2M_1}l_1+\frac{(2k_2+1)}{2M_2}l_2)}(W \bullet D)^c(k_1,k_2) \tag{C.19}$$

Let $k_1' = -(k_1+1)$, $k_2' = -(k_2+1)$. Applying a change of variables, one gets,

$$G^c(l_1,l_2) = \sum_{k_2'=M_2-1}^{-M_2} \sum_{k_1'=M_1-1}^{-M_1} e^{-\jmath\pi(\frac{(-2k_1'-2+1)}{2M_1}l_1+\frac{(-2k_2'-2+1)}{2M_2}l_2)}(W \bullet D)^c(-k_1'-1,-k_2'-1) \tag{C.20}$$

$$= \sum_{k_2'=-M_2}^{M_2-1} \sum_{k_1'=-M_1}^{M_1-1} e^{\jmath\pi(\frac{(2k_1'+1)}{2M_1}l_1+\frac{(2k_2'+1)}{2M_2}l_2)}(W \bullet D)^c(-k_1'-1,-k_2'-1) \tag{C.21}$$

Since $\mathcal{W}(D)$ is real and half - plane symmetric from eq. C.1,

$$(W \bullet D)^c(-k_1'-1,-k_2'-1) = (W \bullet D)(k_1',k_2')$$

Hence,

$$G^c(l_1,l_2) = \sum_{k_2'=-M_2}^{M_2-1} \sum_{k_1'=-M_1}^{M_1-1} e^{\jmath\pi(\frac{(2k_1'+1)}{2M_1}l_1+\frac{(2k_2'+1)}{2M_2}l_2)}(W \bullet D)(k_1',k_2') \tag{C.22}$$

From eqs. C.17 and C.22,

$$G(l_1,l_2) = G^c(l_1,l_2)$$

Therefore, G *is real.*
Eqs. C.14 and C.15 can then be written as :

$$G\beta^c = \sum_{k=1}^{q} \beta^T R_k \beta^c L_k \alpha = X(\beta)\alpha \tag{C.23}$$

$$G^T\alpha^c = \sum_{k=1}^{q} \alpha^T L_k \alpha^c R_k \beta = Y(\alpha)\beta, \tag{C.24}$$

The above represent the necessary and sufficient conditions for optimality, obtained using the vectors (α,β) i.e. (a^c,b^c). Hence, if (a,b) are solutions to the optimization problem, (a^c,b^c) are also solutions to the problem.

The above result has an interesting and significant application to the separable design algorithm, which is stated in the following corollary.

Corollary : *Under the assumptions of the above theorem, if* $W \bullet D = USV^*$ *is the* SVD *of* $W \bullet D$*, the structure of the* U*,* S*,* V *matrices can be used to obtain real coefficients in the design process.*
Let $W \bullet D = USV^*$ be the SVD of $W \bullet D$. Then,

$$W \bullet D = U(:, 1 : m)S(1 : m, 1 : m)V^*(:, 1 : m), \quad m = rank(S)$$

Now,

$$(W \bullet D)(k_1, k_2) = \sum_{i=1}^{m} U(k_1, i)S(i, i)V^*(i, k_2),$$

$$(W \bullet D)(-k_1 - 1, -k_2 - 1) = \sum_{i=1}^{m} U(-k_1 - 1, i)S(i, i)V^*(i, -k_2 - 1),$$

and they are equal as proved in the theorem above.
For filters with frequency responses such that the U and V matrices are columnwise mirror image symmetric or antisymmetric, $U(k_1, i) = \pm U(-k_1 - 1, i), \quad V^*(i, k_2) = \pm V^*(i, -k_2 - 1) \quad for \;\; each \;\; i$. The optimal solutions in these cases are real, and this constraint can be included in the problem construction by modifying the necessary conditions suitably using eqs. 4.2 and 4.3.
There are filters with frequency responses such that the U, S and V matrices have the property that $U(k_1', i) = \pm U(-k_1 - 1, i + 1), \quad V^*(i, k_2) = \pm V^*(i + 1, -k_2 - 1), \quad S(i, i) = S(i + 1, i + 1) \quad for \;\; each \;\; odd \;\; valued \;\; i$, and m is always even, $i.e.$ the matrices U and V are pairwise, columnwise mirror image symmetric or antisymmetric, with the singular values also varying in pairs. In these cases, the optimal solutions always exist as sums of conjugate pairs $(a + a^c, b + b^c)$. This implies that if (a, b) is an optimal solution, then (a^c, b^c) is also an optimal solution by the theorem proved above, and in addition, in such cases, $(\Re(a), \Re(b))$ *is also an optimal solution*, thus ensuring that the optimal coefficients are real. It is to be noted that the above conditions hold for frequency responses where the frequency sample points are taken as $w_{ik_i} = \frac{\pi(k_i)}{M_i}; \;\; k_i = -M_i, \ldots, M_i; \;\; i = 1, 2.$

Appendix D

The following is a listing, together with a description, of the programs used in the design of separable filters using the *successive approximation algorithm* developed in the preceding chapters.

(i) **symm.m**

This is a MATLAB program to design a $2-D$ separable, FIR digital filter which has an ideal frequency response that is either quadrantally symmetric or antisymmetric. The program computes the optimal filter coefficients by the method of recursion, with a randomly chosen starting point. Only the first quadrant of the ideal response is needed for the design, and the coefficient vectors are constrained suitably, to account for the symmetry in the ideal response.

Parameters :

D is the matrix representing the first quadrant of the ideal frequency response, of size $m_1 \times m_2$.

a, b are the optimal coefficient vectors, of sizes $(n_1 + 1) \times 1$, and $(n_2 + 1) \times 1$, respectively, n_1, n_2 being the filter orders in the w_1, w_2 directions.

$\mathcal{W}$ is the matrix representing the weighting function in the first quadrant, of size $m_1 \times m_2$, and having the symmetry of D.

Ω_1, and Ω_2 are the discrete frequency matrices, of sizes $m_1 \times (n_1+1)$ and $m_2 \times (n_2+1)$ respectively. They are obtained from the function *omega.m*, and are suitably modified by multiplication with the matrices S_1, S_2 defined in eq. 4.5.

apr is the matrix representing a separable approximation to D, of size $m_1 \times m_2$. It is initially set to zero.

fil is the matrix representing the designed separable filter frequency response, of size $m_1 \times m_2$. It is also initially set to zero, and is obtained as the summation of each separable term, given here by apr.

$coeff$ is the matrix representing the summation of the optimal filter coefficients of each separable term given by ab^T, and is of size $max(n_1 + 1) \times max(n_2 + 1)$. It is initially set to zero.

L_k, R_k, are matrices of sizes $(n_1+1)\times(n_1+1)$, and $(n_2+1)\times(n_2+1)$ respectively.

G, X, Y, are matrices of sizes $(n_1+1)\times(n_2+1)$, $(n_1+1)\times(n_1+1)$, $(n_2+1)\times(n_2+1)$, respectively.

count represents the number of separable terms used in the design.

iter is the vector containing the number of iterations required for convergence to the optimal solution, for each separable term.

Algorithm :

1) Find the SVD of $\mathcal{W}$, and truncate based on the Frobenius norm, as explained in Chapter 2, using the function *reduc.m*.

2) Set count to 1.

3) Determine the L_k, and R_k matrices, using eq. 3.31.

4) Set $D = D - apr$

5) Determine G from eq. 3.36, and extract the real part.

6) Set $iter(count) = 0$.

7) Start the loop for recursion, with a, b set to random real numbers. a_0, b_0 are used as dummy variables in the recursive process. A flag *done* is used to indicate convergence. Increment $iter(count)$ for each iteration.

8) Determine X using eq. 3.37, and extract the real part. Determine a_0 using eq. 3.39. If $abs(max(a_0 - a)) < 0.00001$, solution is assumed to have converged.

9) Set $a = a_0$.

10) Determine Y using eq. 3.38, and extract the real part. Determine b_0 using eq. 3.40. If $abs(max(b_0 - b)) < 0.00001$, solution is assumed to have converged.

11) Set $b = b_0$.

12) Repeat 8-11 until both a and b converge, when flag *done* is set to 1.

13) Determine $apr = \Omega_1 a b^T \Omega_2^T$.

14) Determine $coeff = coeff + ab^T$, with suitable padding of the vectors $a,\ b$.

15) Determine $fil = fil + apr$

16) Set count to 2, and repeat 3-15 to obtain the *2-term* approximation.

17) Add terms, proceeding in a similar fashion, until the maximum ripple of the designed filter is within acceptable bounds.

18) Determine the SVD of $coeff$.

19) Truncate $coeff$ using as many singular values as necessary to obtain the same maximum ripple as obtained before.

20) Determine the final filter coefficients from the truncated $coeff$, and also the final designed filter using these coefficients.

ii) **conj.m**

This is a MATLAB program to design a $2-D$ separable, FIR digital filter which has an ideal frequency response that is half plane symmetric, and has properties such that the optimal solutions can be considered as sums of conjugate pairs. The parameters are the same as in *symm.m*, except for the following :

Parameters :

D is the matrix representing the complete ideal frequency response, of size $m_1 \times m_2$

$a,\ b$ are the optimal coefficient vectors, of sizes $(2n_1+1) \times 1$, and $(2n_2+1) \times 1$, respectively, $n_1,\ n_2$ being the filter orders in the $w_1,\ w_2$ directions.

$\mathcal{W}$ is the matrix representing the weighting function in all four quadrants, of size $m_1 \times m_2$, and having the symmetry of D.

Ω_1, and Ω_2 are the discrete frequency matrices, of sizes $m_1 \times (2n_1+1)$ and $m_2 \times (2n_2+1)$ respectively. They are obtained from the function *omega.m*.

$coeff$ is the matrix representing the summation of the optimal filter coefficients of each separable term given by ab^T, and is of size $max(2n_1+1) \times max(2n_2+1)$. It is initially set to zero.

$L_k, \quad R_k$, are matrices of sizes $(2n_1+1)\times(2n_1+1)$, and $(2n_2+1)\times(2n_2+1)$ respectively.

$G, \quad X, \quad Y$, are matrices of sizes $(2n_1+1)\times(2n_2+1)$, $(2n_1+1)\times(2n_1+1)$, $(2n_2+1)\times(2n_2+1)$, respectively.

Algorithm :

1) Perform 1-6 as in the *symm.m*.

2) Perform 7 as before, except that a, b are set to random complex numbers.

3) Perform 8-12 as before, except that X and Y are not constrained to be real.

4) Determine $apr = \Omega_1(a+a^c)(b+b^c)^T\Omega_2^T$.

5) Determine $coeff = coeff + (a+a^c)(b+b^c)^T$, with suitable padding of the vectors a, b.

6) Determine $fil = fil + apr$

7) Proceed as before to obtain the designed filter, satisfying a given specification.

iii) **real.m**

This is a MATLAB program to design a $2-D$ separable, FIR digital filter which has an ideal frequency response such that the optimal coefficients are always constrained to be real. The quadrantally symmetric or antisymmetric responses are special examples of this case. Hence, this program is almost entirely similar to *symm.m*, except that two or all four quadrants of the ideal response need to be used for the design, and the discrete frequency matrices are not modified. Also, the coefficient vectors have no symmetry constraints on them.

iv) **map.m**

This is a MATLAB program to obtain the map of the performance variation with number of coefficients. Here, the number of independent coefficients is fixed at some value. For this value, the number of separable terms is fixed at several values, and the separable filter orders are determined for each of these values. The programs (i,ii, or iii) are then run for each set of values to determine the optimal separable filter, from which the cost function can be calculated. The reduction using the SVD on the $coeff$ matrix is not performed. This entire process is repeated for a new fixed value of coefficients. The performance map is thus obtained.

Vita

Vidya Venkatachalam was born in Madras, India. She received her Bachelor of Engineering degree in Electrical and Electronics Engineering from Anna University, Madras, India, in June 1992. She joined the Graduate program in Louisiana State University in August, 1992, where she is currently a candidate for the Master of Science in Electrical Engineering degree.

Abstract

Venkatachalam, Vidya, B.E., Anna University, 1992
Master of Science in Electrical Engineering, Spring Commencement, 1995
Major: Electrical Engineering
Parallel and Separable 2-D FIR Digital Filter Design
Thesis directed by Associate Professor Jorge L. Aravena
Pages in thesis, 94. Words in abstract, 132.

ABSTRACT

This thesis presents the results of a research which develops a technique to design $2-D$ filters by approximating an ideal frequency response with sums of separable FIR components. The technique is independent of the nature of the ideal response, and can accommodate the inclusion of a weighting function. This approach gives the designer flexibility in selecting the $1-D$ filter orders and the number of separable filters to be used for best results. The problem is solved for the weighted least mean squares case, and a rigorous mathematical analysis is used to formulate the separable design algorithm. This work includes a brief analysis of the computational complexity of the formulated technique, and simulation results demonstrating the effectiveness of the design algorithm. It also offers suggestions for further research in this area.