

Design of multidimensional finite-wordlength FIR and IIR filters by simulated annealing

Jan Radecki, Janusz Konrad and Eric Dubois

`konrad@inrs-telecom.quebec.ca`



Université du Québec

Institut national de la recherche scientifique

INRS-Télécommunications

16 Place du Commerce, Verdun

Québec, Canada, H3E 1H6

*IEEE Transactions on Circuits and Systems –
II: Analog and Digital Signal Processing*
vol. 42, no. 6, pp. 424–431, June 1995

Design of multidimensional finite-wordlength FIR and IIR filters by simulated annealing

Jan Radecki, Janusz Konrad and Eric Dubois

INRS-Télécommunications
Institut National de la Recherche Scientifique
16 Place du Commerce, Verdun, Québec, H3E 1H6, Canada

Abstract

This paper describes a new approach to the design of multidimensional (M-D) finite-wordlength digital filters with specifications in the frequency and spatial domains. The approach is based on stochastic optimization and extends previous work on finite impulse response (FIR) filters in two ways: by inclusion of spatial constraints and by application to the case of infinite impulse response (IIR) filters. The formulation proposed is based on a multiple-term objective function that, in addition to magnitude constraints, also includes step response, group delay and stability constraints. Our attention to these characteristics stems from the application of such filters to video processing that we are actively pursuing. Since filter coefficients are of finite precision and since the objective function is multivariable, non-differentiable and likely to have multiple minima, we use *simulated annealing* for optimization. We show numerous examples of the design of practical filters such as channel and luminance/chrominance separation filters used in the NTSC system. We demonstrate the impact of coefficient precision as well as of group delay and step response constraints on filter parameters.

1 Introduction

The work presented in this paper has been motivated by our research in the area of advanced video systems (NTSC, HDTV) where M-D digital filtering is often used. For example, component signals in a modern NTSC receiver are extracted using separable analog/digital 2-D filtering, often referred to as *comb* filtering. The quality of the displayed image, however, can be improved by employing fully-digital non-separable filtering instead [12]. Since digital filtering will play an essential role in digital transmission systems of the future (HDTV, video-conferencing, videophone), it is important to develop general and flexible design methods for finite-precision filters.

Usually, digital filters are designed with respect to magnitude response only. In this paper, we are interested in a design with respect to filter properties in the frequency and space-time domains. If filter output is to be evaluated by human observers, good spatio-temporal properties, such as small ringing at sharp intensity transitions, are also important. This is due to the fact that ringing is easily perceived by the human visual system (HVS). A related problem is that of phase-linearity; the HVS is sensitive to phase distortions. Thus, a filter should have linear or almost linear phase. Since the designed filters are intended for finite-precision implementation, the additional constraint is that filter coefficients be of finite wordlength. Consequently, for FIR filters, magnitude and step responses need to be considered; exact phase linearity can be easily achieved by constraints on filter coefficients. However, IIR filters need to be additionally optimized with respect to phase response and stability.

There have been numerous methods proposed for the design of M-D digital filters. Typical design methods for FIR filters are based on windowing, spectral transformations or error criterion optimization [13]. M-D IIR filters are usually designed either by spectral transformations of 1-D filters [13] or by error criterion optimization. The error criterion may be established in space-time domain [13], frequency domain [2],[29] or state space [20]. Stability of an IIR filter is usually enforced analytically [2], by spectral factorization [29] or by verification after each coefficient update. Some of the methods optimize only magnitude response while others take into account both magnitude and phase properties [10],[29]. Finally, from the optimization point of view, linear programming [10] or non-linear optimization procedures are used [29],[13]. In order to account for spatio-temporal filter properties, in addition to the magnitude response error, also an error related to step response oscillations (ringing) has been used successfully for 1-D [24] and 2-D FIR filters [7],[23].

Implementation of a filter in fixed-point hardware requires finite-precision coefficients. The simplest approach is to quantize the full-precision coefficients¹. This method, however, produces sub-optimal results since filter response may change substantially after coefficient quantization. Moreover, for IIR filters such quantization may move poles close to or into the instability region, thus making the filter unusable. Thus, local and global design methods have been proposed. Local methods usually start from a quantized full-precision solution and are implemented through a univariate search, random search or branch and bound method. Since these methods are sub-optimal, designs assuring optimality of the solution have been proposed. One example is mixed integer programming [17], while another is *simulated annealing* [11],[8],[6], a method capable of finding the global minimum of complex, multi-variable functions with multiple minima [16]. Simulated annealing has been first ap-

plied to general 1-D linear-phase FIR filters, but later has been extended to some special 1-D digital filters such as *Nyquist* filters [5] or multiplierless filters for video applications [3].

In this paper we propose a comprehensive design method for M-D finite-precision filters. The method is based on a multiple-term cost function that comprises magnitude, phase (group delay), stability and step response errors. In this way several requirements are combined. Their relative importance is controlled by suitable weights. In order to minimize this multi-variable, non-differentiable function with multiple minima we use *simulated annealing* based on the *Metropolis algorithm*.

The paper is organized as follows. In Section 2 a general formulation of the finite-precision filter design problem is given, followed by a description of the solution method used. Then, in Sections 3 and 4 elements of the cost function are described in detail and examples of computer simulations are given for FIR and IIR filters, respectively. In Section 5 the paper is summarized and conclusions are drawn.

2 General problem: formulation and solution

2.1 Cost function

Let $\mathbf{H}(\boldsymbol{\omega}) = H(\boldsymbol{\omega}) \cdot e^{j\Phi(\boldsymbol{\omega})}$ be the (complex) frequency response of an N -dimensional FIR or IIR filter, where $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_N)^T$ is an N -dimensional frequency vector. In the special case of 1-D and 2-D filters, for which examples are given in the next two sections, we have $\boldsymbol{\omega} = \omega$ and $\boldsymbol{\omega} = (\omega_1, \omega_2)^T$, respectively. Clearly, $H(\boldsymbol{\omega})$ and $\Phi(\boldsymbol{\omega})$ are the magnitude and phase responses of the filter. Let $\tau_i(\boldsymbol{\omega})$ be its group delay with respect to ω_i . We assume a general cascaded form for the filter, which includes both the direct form and a cascade of low order sections as special cases. We write the frequency response of such filter as follows

$$\mathbf{H}(\boldsymbol{\omega}) = \frac{\prod_{k=1}^K \sum_{\mathbf{m} \in \mathcal{M}_k} a_k(\mathbf{m}) \cdot e^{-j\boldsymbol{\omega}^T \mathbf{m}}}{\prod_{l=1}^L \sum_{\mathbf{n} \in \mathcal{N}_l} b_l(\mathbf{n}) \cdot e^{-j\boldsymbol{\omega}^T \mathbf{n}}} \quad (1)$$

where K, L denote the number of cascaded sections in the numerator and denominator, and $\mathbf{m} = (m_1, \dots, m_N)^T, \mathbf{n} = (n_1, \dots, n_N)^T$ are N -dimensional vectors of coefficient indices from numerator and denominator masks \mathcal{M}, \mathcal{N} , respectively. Since one coefficient in each denominator polynomial $\mathbf{B}_l(\boldsymbol{\omega})$ is redundant, we assume that $b_l(\mathbf{0})=1$ for $l = 1, \dots, L$. Then, for FIR filters $L=1$ and $b_1(\mathbf{n})=0$ for all \mathbf{n} except $\mathbf{n} = \mathbf{0}$. For the canonical filter representation we have $K = L = 1$.

In this paper we assume that filters are defined on a rectangular sampling lattice. This lattice is most often used in signal processing as it allows separability as well as simple implementation. The formulation presented below, however, is general enough to handle other sampling lattices; filter indices and masks would have to be suitably redefined. We assume that for FIR filters the mask \mathcal{M} has non-causal support, i.e., full axis for 1-D case and four quadrants for 2-D case. For the numerator and denominator masks of an IIR filter, causal support is assumed, i.e., half axis for 1-D case and first quadrant for 2-D case. We do not consider non-symmetric half-plane (NSHP) denominators since they are non-separable; the need for transfer function separability is explained below.

For IIR filters of dimension higher than 1, assurance of filter stability is both non-trivial and computationally expensive [1]. A recent result by Wan and Fahmy [28], however, permits us to greatly simplify the frequency response (1). They have shown (Theorem 7 and Corollary 8 in [28]) that the denominator of a transfer function possesses quadrantal symmetry and satisfies stability constraints *if and only if* it is separable. This result has a significant implication since very often filters characterized by a frequency response with quadrantal symmetry are sought. For example, in image processing horizontal and vertical mirror images of a pattern should be treated in the same way as the original pattern, and thus quadrantal symmetry is appropriate.

We conclude that the desired frequency response with quadrantal symmetry can be approximated by a separable denominator function

$$\prod_{l=1}^L \prod_{i=1}^N \sum_{n \in \mathcal{N}_{l,i}} b'_{l,i}(n) \cdot e^{-j\omega_i n}. \quad (2)$$

With the above factorization, the N -D stability test simplifies to $(N \cdot L)$ simple 1-D tests.

Let $\boldsymbol{\theta} = [\theta_1, \dots, \theta_P]^T$ be a vector that contains all $P = \sum_k |\mathcal{M}_k| + \sum_l \sum_i (|\mathcal{N}_{l,i}| - 1)$ unknown coefficients, where $|\mathcal{M}|$ denotes the number of coefficients in mask \mathcal{M} . Note that since $b'_{l,i}(0) = 1$, each denominator polynomial has $|\mathcal{N}_{l,i}| - 1$ unknown coefficients. Let each coefficient have a fixed-point representation with b bits, and let the maximum absolute value of each coefficient be A . Then, the quantization step is defined as $q = A/2^{b-1}$ and its value belongs to the set $\mathcal{S} = \{\theta : \theta = -A + q \cdot m, m = 1, \dots, 2^b\}$. This could be easily extended to accommodate different A and/or b for each coefficient.

Filters designed by the algorithm proposed in this paper are intended for use in video processing. Since the final evaluation of an image is done by the human observer, two properties of the HVS are critical. First, the HVS is sensitive to phase nonlinearities (i.e., non-constant group delay) [15],[21] introduced at the processing stage. Such nonlinearities manifest themselves as double contours, distortions in regular patterns, etc. Thus, filters with linear or almost-linear phase are required. Secondly, although a digital filter is uniquely specified by its magnitude and phase responses, the spatio-temporal performance of a filter is also important to the HVS. Filters used in video processing must not produce annoying “ghosts” or multiple echoes (called ringing) around sharp intensity transitions; filter response to the unit step must not contain large oscillations or overshoots. Note that optimization of frequency and spatio-temporal properties of a filter cannot be done independently. What we can hope for is to achieve a certain compromise between the different requirements.

Apart from the need for a sharp amplitude transition, it is not clear what spatio-temporal shape a unit step should have. For $N=1$ there is no ambiguity, but already for $N=2$ different spatial shapes can be selected. We will use the corner unit step [7] since it gives more constraints than the diagonal step [23] and simultaneously exhibits both horizontal and vertical ringing. This is a valuable property as images are very often composed of objects with contours in these directions.

To design a digital filter under various constraints we propose to use the following multiple-term cost function:

$$\mathcal{E}(\boldsymbol{\theta}) = \mathcal{E}_H + \lambda_\tau \mathcal{E}_\tau + \lambda_s \mathcal{E}_s + \lambda_\eta \mathcal{E}_\eta, \quad (3)$$

where the λ 's are non-negative weights and the \mathcal{E} 's are error terms (implicitly dependent on θ) measuring the departure of filter magnitude, group delay, stability and step response from some desired properties.

2.1.1 Magnitude error

\mathcal{E}_H is the *magnitude error* due to the departure of $H(\omega)$ from the desired magnitude response $D(\omega)$. To obtain filters with equiripple response we define \mathcal{E}_H as follows:

$$\mathcal{E}_H = \max_{\omega \in \Omega_H} [\gamma_H(\omega) \cdot |H(\omega) - D(\omega)|], \quad (4)$$

where $\gamma_H(\omega)$ is a weighting function and Ω_H is a set of frequencies ω at which \mathcal{E}_H is evaluated. This set usually forms a uniform rectangular sampling grid in $[-0.5, 0.5]^N$ (N -fold Cartesian product)². Non-uniform grids are occasionally used, especially around transition bands. In the literature, control over the transition band is often omitted ($\gamma_H=0.0$). We shape the magnitude response in the transition band as well, although with a reduced emphasis ($\gamma_H=0.5$). We do so in order to control the impact of step response optimization on the transition band properties; experiments have shown that a relatively small change of λ_η may significantly affect the magnitude error in and close to the transition band. The weight γ_H can be also used to get a very close approximation at certain frequencies, e.g., DC ($\omega = \mathbf{0}$), by choosing a very large weight. Note that the error (4) is evaluated only for $\omega \in \Omega_H$, and thus the sampling grid in Ω_H must be chosen dense enough in order to provide a sufficiently close approximation to $D(\omega)$.

2.1.2 Group delay (phase) error

To minimize phase non-linearity in IIR filters, we chose to minimize a measure of variability of the group delay. We do not use the phase $\Phi(\omega)$ directly since phase unwrapping would have to be carried out for each ω . Thus, we calculate analytically group delay(s) $\tau_i(\omega)$, and define the *group delay error* \mathcal{E}_τ as follows

$$\mathcal{E}_\tau = \max \left\{ \max_{\omega \in \Omega_\tau} [\gamma_\tau(\omega) |\tau_1(\omega) - \bar{\tau}_1|], \dots, \max_{\omega \in \Omega_\tau} [\gamma_\tau(\omega) |\tau_N(\omega) - \bar{\tau}_N|] \right\}. \quad (5)$$

Ω_τ is the set of frequencies ω , similar to Ω_H , at which \mathcal{E}_τ is evaluated. $\gamma_\tau(\omega)$ is again a frequency-dependent weighting function. For filters with stop-band attenuation of at least 20 dB, phase linearity is irrelevant in the stop band [9] and $\gamma_\tau(\omega)=0$ may be used. $\bar{\tau}_i$ is a value of the i -th group delay to be as closely approximated by $\tau_i(\omega)$ over Ω_τ as possible. If such a value is known, as sometimes may be the case, then $\bar{\tau}_i$ is fixed. If it is not known and only a group delay constancy is required, then, for example, $\bar{\tau}_i$ may be defined as a γ_τ -weighted average of the i -th group delay

$$\bar{\tau}_i = \frac{1}{|\Omega_\tau|} \sum_{\omega \in \Omega_\tau} \gamma_\tau(\omega) \tau_i(\omega). \quad (6)$$

Note that the error (5) expresses the maximum value of individual ($i = 1, \dots, N$) group delay variations over Ω_τ , i.e., first, the maximum of weighted variation for each $\tau_i(\omega)$ is calculated, and then the largest such variation is selected.

2.1.3 Stability error

In order to assure IIR filter stability, zeroes of each polynomial in (2) must be within the unit circle. The zeroes can be found using any factorization scheme, although in our designs we have used at most second-order sections. If $|z|_{max}$ is the zero with the largest magnitude, then $1 - |z|_{max}$ is the filter's stability margin.

Suitable margin of stability must be assured for filters designed in a floating-point arithmetic and implemented in a fixed-point arithmetic; after quantization the position of poles may change. In this paper, however, we design finite-wordlength filters directly. A control over filter stability margin may be useful for two reasons. First, the stability margin of an IIR filter has a direct impact on the settling time (rate of decay) of filter's impulse response [1]. Consequently, it affects the extent of oscillations in filter step response. Secondly, limit cycles in IIR filters depend on the position of filter poles with respect to the unit circle [24].

In order to control the filter's margin of stability we discourage generation of poles too close to the unit circle by proposing the following *stability error*:

$$\mathcal{E}_s = \begin{cases} \min\{1/((1 - \delta) - |z|_{max}), \zeta\}, & \text{for } |z|_{max} < 1 - \delta, \\ \zeta, & \text{for } |z|_{max} \geq 1 - \delta, \end{cases} \quad (7)$$

where $0 \leq \delta < 1$ is the minimum margin of stability and ζ is a large positive number. Stability error (7) can be interpreted as follows. Each pole of a filter must be located within the circle with radius $1 - \delta$; otherwise a high penalty is assigned to \mathcal{E}_s . Inside this circle, the stability error grows with $|z|_{max}$ approaching the circle; coefficients producing poles farther from the circle are penalized less. Criterion (7) is applicable to any M-D filter.

2.1.4 Step response error

Let $\eta(\mathbf{n})$ be the filter response to the corner unit step. Since it is very difficult, if not impossible, to propose a suitable error in a general case, we make several restricting assumptions. We assume that the filter is low-pass (otherwise no step would appear at the output), has positive DC gain, and that its overshoots do not exceed a fraction ϑ of the input step response value. Let \mathbf{n}_{max} and \mathbf{n}_{min} be coordinates of the maximum and minimum values of $\eta(\mathbf{n})$ over its full region of support. We also assume that the largest oscillations in the step response occur around \mathbf{n}_{min} and \mathbf{n}_{max} . Although occasionally not true, this assumption has held very well for the examples given in this paper. This restriction greatly simplifies the calculations since only one maximum for the causal part and one minimum for the anticausal part have to be considered.

With $\mathcal{K}(\mathbf{n})$ being a neighborhood of \mathbf{n} , we define two sets as follows:

$$\begin{aligned} \mathcal{W}_1(\mathbf{n}) &= \{\mathbf{k} \in \mathcal{K}(\mathbf{n}) : \eta(\mathbf{k}) \geq (1 - \vartheta)D(\mathbf{0}), \Delta_j \eta \leq 0, j = 1, \dots, N\}, \\ \mathcal{W}_2(\mathbf{n}) &= \{\mathbf{k} \in \mathcal{K}(\mathbf{n}) : \eta(\mathbf{k}) \leq \vartheta D(\mathbf{0}), \Delta_j \eta \leq 0, j = 1, \dots, N\}, \end{aligned}$$

where $\Delta_j \eta = \eta(k_1, k_2, \dots, k_j + 1, \dots, k_N) - \eta(k_1, k_2, \dots, k_j, \dots, k_N)$ is the first-order difference of η in the direction k_j . We separately search for the causal overshoot (around \mathbf{n}_{max}) and for the anticausal overshoot (around \mathbf{n}_{min})

$$\mathcal{E}_\eta^+ = \max_{\mathbf{k} \in \mathcal{W}_1(\mathbf{n}_{max})} \eta(\mathbf{n}_{max}) - \eta(\mathbf{k}), \quad \mathcal{E}_\eta^- = \max_{\mathbf{k} \in \mathcal{W}_2(\mathbf{n}_{min})} \eta(\mathbf{k}) - \eta(\mathbf{n}_{min}). \quad (8)$$

Note that due to the definitions of \mathcal{W}_1 and \mathcal{W}_2 , only a search among neighbors with step response value above $(1 - \vartheta)D(\mathbf{0})$ (causal) and below $\vartheta D(\mathbf{0})$ (anticausal) threshold is executed. Moreover, this set is further restricted by requiring that first-order differences (approximation to local gradient components) be negative, i.e., that positions be not on the step rise. We define the *step response error* as the maximum of the above overshoots:

$$\mathcal{E}_\eta = \max(\mathcal{E}_\eta^+, \mathcal{E}_\eta^-). \quad (9)$$

2.2 Minimization method

To design a digital filter with constraints described above we need to minimize the cost function (3) with respect to $\boldsymbol{\theta}$. The function under minimization is characterized by several important properties. First, it is defined over discrete-valued vector $\boldsymbol{\theta}$. Consequently, no gradient-type technique can be used directly, unless quantization follows each coefficient update, which is a suboptimal approach. Secondly, in the general case it is multimodal, i.e., possesses multiple minima. It is also non-differentiable due to the way \mathcal{E}_η is defined; no gradient-based method can be used. Finally, it is usually defined over a few dozen to a few hundred variables.

A method that we propose to carry out this minimization is simulated annealing [16]. This stochastic search algorithm is a computer simulation of the process of *annealing* of solids; the behavior of a solid is simulated by generating sample configurations from a probability distribution with an energy function (such as the one defined in (3)) divided by a “temperature” parameter T . The sample configurations are produced using, e.g., the *Metropolis algorithm* [22]. Initially T is chosen to be sufficiently high to generate a full range of configurations, even the unlikely ones. This assures avoidance of local minima and is equivalent to melting the solid. As the process evolves, T is very slowly reduced. It has been shown [14] that if the reduction of T is sufficiently slow, then the system attains (in a limit) the state of minimal energy.

We use the Metropolis algorithm to generate samples from a suitable discrete-time Markov chain. Let $\boldsymbol{\theta}_s$ be the vector of coefficients at (discrete) time s , and let $\boldsymbol{\theta}_s^*$ be another vector of coefficients that is identical to $\boldsymbol{\theta}_s$ except the coefficient θ_{i_s} . Note that only one coefficient is modified in one trial. Let T_0 and T_{min} denote the initial and final temperatures, and let μ be the number of trials for which the temperature T stays constant. Every μ trials the temperature is modified according to the schedule $T_{j+1} = \rho T_j$. Let the energy increment, resulting from an attempted change at time s , be $\Delta\mathcal{E}_s = \mathcal{E}(\boldsymbol{\theta}_s^*) - \mathcal{E}(\boldsymbol{\theta}_s)$. The Metropolis-based simulated annealing algorithm has been implemented as follows:

- 1) start with an initial temperature T_0 and an initial state $\boldsymbol{\theta}_0$; $j=0, s=0$,
- 2) randomly select a new state $\boldsymbol{\theta}_s^*$ in two steps:
 - a. generate a new location i_s (uniformly distributed) of a coefficient to be updated,
 - b. generate a new value (uniformly distributed) from \mathcal{S}_{i_s} for coefficient θ_{i_s} ,
- 3) compute the energy increment $\Delta\mathcal{E}_s$ and make the following decision:
 - a. if $\Delta\mathcal{E}_s \leq 0$, accept the new state unconditionally: $\boldsymbol{\theta}_{s+1} = \boldsymbol{\theta}_s^*$,
 - b. if $\Delta\mathcal{E}_s > 0$, accept the new state with probability $p = \exp(-\Delta\mathcal{E}_s/T_j)$,

- 4) $s = s + 1$; if $s \equiv 0 \pmod{\mu}$, then $T_{j+1} = \rho T_j$ and $j = j + 1$,
- 5) if $T_j \leq T_{min}$, exit; otherwise return to 2.

In our implementation of simulated annealing, multiplication by $\rho < 1$ causes the reduction of temperature: $T_j = T_0 \rho^j$. T is decreased in larger steps when T is large and in smaller steps when it is small. Optimal annealing schedules, such as the logarithmic one [14], change the temperature very little for large j (small T). To reduce the decrements for small T we use three different values of ρ as follows

$$T_{j+1} = \begin{cases} \rho_1 T_j, & \text{if } 0 \leq j < J_1, \\ \rho_2 T_j, & \text{if } J_1 \leq j < J_2, \\ \rho_3 T_j, & \text{if } J_2 \leq j, \end{cases} \quad (10)$$

where $\rho_1 < \rho_2 < \rho_3$ and J_1, J_2 indicate when a change of ρ occurs. These parameters as well as μ have been selected experimentally during the design. Typically, we have used $\rho_1=0.97-0.99$, $\rho_2=0.992-0.997$, $\rho_3=0.998-0.999$ switched around $J_1=0.01$ and $J_2=0.005$. We have changed the temperature every $\mu = 10P$ trials.

Another important parameter in simulated annealing is the initial temperature T_0 . The theoretical value of T_0 , in order to assure convergence to the global optimum, is usually impractically high. Thus, sub-optimal values of T_0 are often used. We have used initial temperatures from 0.125 to 0.5 depending on the designed filter. For such temperatures the initial acceptance ratio was typically about 15-20%. We have experimented with higher values of T_0 giving initial acceptance ratios in the 30-50% range, however we encountered significant numerical difficulties. The very unlikely states generated at high temperatures gave filters with 300% magnitude ripples and group delays of the order of several hundred. For such filters the calculation of step response error as defined in Section 2.1.4 fails; for example a high-pass filter instead of a low-pass one is generated. To improve chances of locating the global optimum we first carry out the optimization several times for different seeds and different annealing schedules, and then use the best result as the initial state for further optimization with reduced T_0 .

To speed-up the calculations at low temperatures we have modified the Metropolis algorithm as follows. If a proposed coefficient θ_{i_s} differs from the current coefficient by more than 0.5 for 1-D low-pass filters and by 0.25 for 2-D low-pass filters, then we consider such a change very unlikely and we proceed to the next trial. Skipping the calculation of all errors in (3) lets us substantially reduce the computational complexity of the algorithm, and from our experience does not impair results in any way. This is probably due to the fact that the modification is used only at low temperatures when simulated annealing behaves similarly to the steepest descent algorithm (acceptance ratio below 1%). Other modifications speeding-up the execution are possible [6].

3 Design of FIR filters

We have designed finite-precision FIR filters by minimizing $\mathcal{E}_H + \lambda_\eta \mathcal{E}_\eta$ ($\lambda_\tau = \lambda_s = 0$). We have implemented the method for odd number of coefficients horizontally and vertically.

Since filter coefficients are real, $H(\omega)$ is symmetric and Ω_H needs to cover only the $[0, 0.5]$ range. For 2-D filters, half-plane symmetry of coefficients is sufficient for linear phase. We

use, however, quadrantal coefficient symmetry due to its implementation efficiency; Ω_H is defined over $[0, 0.5] \times [0, 0.5]$. We use Ω_H consisting of 128 uniformly sampled points in the 1-D case and 32×32 points in the 2-D case.

To describe performance of the designed filters we calculate certain parameters over a denser grid Ω_f . For 1-D filters Ω_f consists of 1024 uniformly sampled points, and for 2-D filters it is a uniform 128×128 orthogonal grid. Let Ω_f^p and Ω_f^s be subsets of Ω_f such that all their elements are in the pass and stop bands, respectively. For each band we identify the frequency at which the maximum magnitude ripple occurs:

$$\omega'_H = \arg \max_{\omega \in \Omega_f^p} |H(\omega) - D(\omega)|, \quad \omega''_H = \arg \max_{\omega \in \Omega_f^s} |H(\omega) - D(\omega)|.$$

To measure filter performance we use the maximum pass-band ripple δ_p and the minimum stop-band attenuation α_s defined as follows:

$$\delta_p = 20 \log \frac{H(\omega'_H)}{D(\omega'_H)} \quad [\text{dB}], \quad \alpha_s = 20 \log H(\omega''_H) \quad [\text{dB}]. \quad (11)$$

Note that δ_p is the maximum linear deviation of filter response from $D(\omega)$ in the pass band expressed in dB, and not the maximum dB ripple in this band. The two are not necessarily the same since for the same excursion of $H(\omega'_H)$ above and below $D(\omega'_H)$ different values of δ_p result. Since δ_p is signed, the linear deviation $H(\omega'_H) - D(\omega'_H)$ can be always uniquely recovered using the relationship (11). Note that since the optimization is performed over Ω_H and since $\Omega_H \subset \Omega_f$, δ_p and α_s are degraded compared to their values calculated over Ω_H . We also compute step response overshoots \mathcal{E}_η^+ and \mathcal{E}_η^- using (8) with $\vartheta=0.5$, $(-3,3)$ neighborhood \mathcal{K} in the 1-D case and $(-3,3) \times (-3,3)$ neighborhood in the 2-D case. Note that for 1-D FIR filters with linear phase we have $\mathcal{E}_\eta^- = \mathcal{E}_\eta^+$ because of coefficient symmetry.

3.1 1-D FIR filters

We have experimented with low-pass NTSC channel filters with the pass band from 0 to $\omega_p=0.2933$ and the stop band from $\omega_s=0.3344$ to 0.5 [4]. Note that the transition band is narrow and is expected to result in significant step response overshoots. We have defined the desired magnitude response to be $D(\omega) = \Psi(1, 0.5(1 + \cos \pi(\omega - \omega_p)/(\omega_s - \omega_p)), 0)$, where

$$\Psi(a_1, a_2, a_3) = \begin{cases} a_1, & \text{for } 0 \leq \omega \leq \omega_p, \\ a_2, & \text{for } \omega_p < \omega < \omega_s, \\ a_3, & \text{for } \omega_s \leq \omega \leq 0.5. \end{cases} \quad (12)$$

We have chosen $K=1$, $|\mathcal{M}_1|=49$, $A=1.0$, $b=6, 8$ or 10 , and $\gamma_H(\omega) = \Psi(1, 0.5, 1)$. To assure a specific gain at frequency ω (e.g., $H(0)=1.0$), we set $\gamma_H(\omega)$ to a high value.

Table 1 shows the comparison between SA-designed filters and filters designed by the N -step Newton minimization [19] ($\lambda_\eta=0$). The reference design is the full-precision³ implementation of the N -step Newton minimization. The finite-precision version of this method has been implemented by in-loop coefficient quantization [25]. Clearly, coefficient precision has a significant impact on magnitude ripples of FIR filters. As we shall see later this effect is less severe for IIR filters. SA-designed filters have the same or better pass-band ripple and/or

increased stop-band attenuation, while step response oscillations are almost unchanged. For 6-bit filters, however, we could not obtain any improvement by using simulated annealing, which may be due to the four times smaller state space for each coefficient (compared with 8-bit design) that is easier to explore.

The bottom part of Table 1 shows parameters of filters designed with the step response constraint. As expected, the maximum overshoot is noticeably reduced, although at the cost of increased magnitude ripples. The compromise between the magnitude ripples and the overshoot amplitude is due to the fact that no other degree of freedom is left (phase linearity is enforced analytically). We will see in Section 4.1 that in the case of IIR filters the compromise also involves the group delay. Magnitude and step responses of one of the designed filters are shown in Fig. 1 (jointly with an IIR filter).

3.2 2-D FIR filters

We have designed 2-D low-pass FIR filters with diamond-shaped pass band, that are used in enhanced NTSC coding [12]. We have chosen $K=1$, $|\mathcal{M}_1|=165$ (15 coefficients per line, 11 lines), $A=0.25$, $b=8$ or 6. Due to the quadrantal symmetry of the filter mask, there are 48 independent coefficients to be optimized. We chose a smaller vertical dimension of the filter because it is intended for video processing where a line delay is much more costly than a pixel delay. The desired magnitude response of the filter is derived from the 1-D version; values are calculated along straight lines through the origin. The 2-D weighting function $\gamma_H(\boldsymbol{\omega})$ is also derived from the 1-D version. As in the 1-D case, values of γ_H are increased at $\boldsymbol{\omega} = \mathbf{0}$ to assure the unit gain at DC.

We have compared the proposed method with N -step Newton minimization [25] for the case without step response constraint ($\lambda_\eta=0$). Table 2 shows parameters of filters designed by full- and finite-precision N -step Newton minimization and by finite-precision SA. Note a very significant improvement in all parameters of SA-designed filters in comparison with the Newton-designed ones. Also, note a significant increase in the magnitude error for 6-bit coefficients in comparison with the 8-bit design. As we shall see in Section 4.2 this effect is far less pronounced in IIR filters. We have also tested the impact of step response constraint (bottom of Table 2). Notice that an increase in λ_η causes, as expected, a reduction of step response overshoots while at the same time increasing the magnitude errors. Again, a direct trade-off between magnitude ripples and step response oscillations can be observed.

Fig. 2 shows the magnitude and step responses of the 2-D 8-bit diamond-shaped FIR filter designed with the step response constraint ($\lambda_\eta=0.5$). Note the overshoot ridge at the rise of the step response and its fast decay without oscillations, a typical characteristic of a low-pass filter with a wide transition band. For more results please consult [25].

4 Design of IIR filters

To control phase non-linearity and stability of an IIR filter, we use the full four-term cost function (3). No symmetry constraints are applied to the numerator or denominator coefficients, although, if needed, such constraints could be included.

Since IIR coefficients are real, $H(\boldsymbol{\omega})$ possesses half-plane symmetry. Consequently, the sets Ω_H and Ω_τ are as before for the 1-D case, and are defined over the range $[-0.5, 0.5] \times [0, 0.5]$ for the 2-D case. In all experiments with IIR filters, the same $\gamma_H(\boldsymbol{\omega})$ is used as for FIR filters (Section 3), but with a suitably extended support in the 2-D case.

The same parameters as those in Section 3 are used here to describe filter performance. Additionally, the stability margin $1 - |z|_{max}$ and the i -th group delay spread $\Delta\tau_i$ in the pass band

$$\Delta\tau_i = \tau_i(\boldsymbol{\omega}'_{\tau_i}) - \tau_i(\boldsymbol{\omega}''_{\tau_i}) \quad (13)$$

are used with $\boldsymbol{\omega}'_{\tau_i}$ and $\boldsymbol{\omega}''_{\tau_i}$ defined as follows

$$\boldsymbol{\omega}'_{\tau_i} = \arg \max_{\boldsymbol{\omega} \in \Omega_f^p} \tau_i(\boldsymbol{\omega}), \quad \boldsymbol{\omega}''_{\tau_i} = \arg \min_{\boldsymbol{\omega} \in \Omega_f^p} \tau_i(\boldsymbol{\omega}).$$

4.1 1-D IIR filters

We have designed 1-D IIR filters with the same frequency specifications as in Section 3.1. We have chosen a cascade of $K=L=5$ sections in both numerator and denominator with $|\mathcal{M}_i|=3$, $|\mathcal{N}_i|=3$ for $i = 1, \dots, 5$ (25 coefficients in total), and $\gamma_\tau(\omega) = \Psi(1, 0.5^2(1 + \cos \pi(\omega - \omega_p))/(\omega_s - \omega_p)^2, 0)$ to insure rapid roll-off in the transition band.

In Table 3, parameters of 1-D IIR filters designed using the proposed method are shown. As expected, magnitude response ripples increase with the reduction of coefficient precision. Note, however, that there is only 0.07 dB increase in δ_p and a 0.9 dB reduction in α_s when the precision is reduced from 10 to 8 bits. For FIR filters designed by the proposed algorithm, these numbers were 0.18 dB and 4.97 dB, respectively. Similarly, for the change from 8- to 6-bit precision, the IIR filter loses 0.28 dB in the pass band and 2.94 dB in the stop band, while the corresponding FIR filter loses 0.38 dB and 14.03 dB, respectively. This relationship, which we have observed in other cases as well, perhaps stems from the fact that IIR filters have rational representation.

To verify the impact of group delay and step response constraints, we have designed 1-D IIR filters with various weights λ . The corresponding filter parameters are shown in the bottom part of Table 3. For $\lambda_\tau=0$ no control over group delay is exercised and its maximum spread $\Delta\tau$ is increased more than twofold in comparison with the filter designed with $\lambda_\tau=0.0125$. For $\lambda_\tau=0.05$, insuring better group delay constancy, $\Delta\tau$ is clearly reduced, however at the cost of increased magnitude ripples. On the other hand, if λ_γ is set to 0, we lose control over step response overshoots (an increase to 0.191).

In order to verify the impact of the stability margin on the extent of step response oscillations, we have designed filters using different values of λ_s (another way would be to change the minimum stability margin δ). To describe the decay rate of step response oscillations, a parameter called *settling time* t_s is often used. This parameter gives the time needed for filter output to settle to a fraction σ of the unit impulse applied at the input. t_s depends on filter stability margin: $t_s = \text{int}[-\ln \sigma / \ln(1 + (1 - |z|_{max}))]$ [1]. Table 4 shows filters with various stability margins designed using the proposed method for different values of λ_s . Note that the smaller the weight λ_s , the smaller the stability margin and the larger the settling time t_s ($\sigma=0.01$). t'_s in Table 4 is the actual settling time evaluated from filter impulse response for a given σ . The disparity between t_s and t'_s is small for large stability margins. For small margins, however, there is a substantial discrepancy.

In Fig. 1 magnitude response, step response and group delay of an IIR filter (jointly with an FIR filter) are shown. Examples of the design with other specifications and a comparison with analytic Chebyshev and elliptic filters can be found in [26],[18].

4.2 2-D IIR filters

We have designed 2-D low-pass IIR filters with the same frequency specifications as used in Section 3.2. We have chosen a cascade of $K=2$ sections in the numerator with $|\mathcal{M}_1|=25$ and $|\mathcal{M}_2|=9$, and of $L=6$ sections in the denominator with $|\mathcal{N}_i|=3$ for $i = 1, \dots, 6$ (46 coefficients in total). The same γ_H as in Section 3.2 has been used and γ_τ has been derived from the one dimensional version along straight lines through the origin.

Using the proposed method we have designed 6- and 8-bit 2-D low-pass IIR filters with diamond-shaped pass band (Table 5). As expected, magnitude response ripples increase with the reduction of coefficient precision. However, there is only a 0.05 dB increase in δ_p and a 1.33 dB reduction in α_s when the precision is reduced from 8 to 6 bits. For FIR filters these numbers were 0.52 dB and 9.05dB, respectively.

To verify the impact of the group delay and step response constraints, different combinations of weights have been used (bottom of Table 5). As expected, for $\lambda_\tau=0$ the group delay spread is about three times larger while for $\lambda_\tau=0.1$ it is about three times smaller than for $\lambda_\tau=0.02$. Unlike in the 1-D case, this improvement came at little cost to the magnitude ripples, which is perhaps due to the wide transition band. By setting λ_η to 0, the control over step response overshoots was lost and resulted in an increase of \mathcal{E}_η to 0.154 (from 0.1).

In Fig. 3 magnitude response, step response and group delays are shown for one of the designed filters. Other examples of the design of video 2-D IIR filters can be found in [27].

5 Summary and conclusions

We have presented a new approach to the design of M-D digital filters with finite-precision coefficients. The approach presented rests on stochastic minimization of a multiple-term objective function incorporating various constraints on filter properties in both frequency and spatial domains. Several conclusions can be drawn from the results obtained. Firstly, the proposed algorithm is an effective method of designing filters with several competing constraints, although it is computationally intensive (typically it took about 1 hour for 1-D IIR and 6-8 hours for 2-D IIR filters on DECstation 5000/120). A drawback of the method is the need to establish practical annealing parameters. Secondly, IIR filters are less sensitive to the precision of coefficients than FIR filters; for the same performance simpler arithmetic and thus simpler hardware can be used. Thirdly, IIR filters with relatively small group delay variation can be designed using the proposed method. This variation depends on the desired magnitude response, however. For 1-D filters with narrow transition band we have obtained filters with group delay variation of about 2 pixels – too high a value for image filtering. For 2-D low-pass luminance filters (wide transition band, gentle roll-off), however, we have obtained filters with group delay variations as low as 0.1-0.2 pixels, a value sufficiently low for practical application to image filtering (an application to NTSC coding will be published in a forthcoming paper). Fourthly, FIR filters cannot be easily optimized for minimal step

response overshoots; any reduction in such overshoots is reflected in dramatically increased errors in the magnitude response. In fact, we have obtained several 1-D and 2-D FIR filters with small overshoots (0.07 to 0.09) only to discover unacceptably high errors in the magnitude response. This trade-off is less severe for IIR filters since any reduction in step response overshoots is reflected jointly in magnitude and group delay errors. Thus, depending on the filter application, a suitable use of weights λ may place the unwanted error in a less crucial property of the filter. Finally, the step response error used in this paper may not be the most appropriate one if ringing visibility at sharp image transitions is a concern. We have tested the designed filters on several images and we have concluded that the largest overshoot is not always the most objectionable artifact. Often the dominant role is played by the second largest overshoot and by the extent of ringing. Thus, the step response error needs to be further investigated to better match properties of the HVS.

References

- [1] P. Agathoklis, E. Jury, and M. Mansour, "The margin of stability of 2-D linear discrete systems," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-30, pp. 869–873, Dec. 1982.
- [2] S. Aly and M. Fahmy, "Design of two-dimensional recursive digital filters with specified magnitude and group delay characteristics," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 908–916, Nov. 1978.
- [3] L. Banzato, N. Benvenuto, and G. Cortelazzo, "A design technique for two-dimensional multiplierless FIR filters for video applications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 2, pp. 273–284, Sept. 1992.
- [4] K. Benson, *Television Engineering Handbook*, ch. 21. McGraw-Hill, Inc., 1986.
- [5] N. Benvenuto, M. Marcesi, and A. Uncini, "Applications of simulated annealing for the design of special digital filters," *IEEE Trans. Signal Process.*, vol. 40, pp. 323–332, Feb. 1992.
- [6] N. Benvenuto and M. Marchesi, "Digital filter design by simulated annealing," *IEEE Trans. Circuits Syst.*, vol. CAS-36, pp. 459–460, Mar. 1989.
- [7] A. Biasiolo, G. Cortelazzo, and G. Mian, "Computer aided design of multidimensional FIR filters for video applications," *IEEE Trans. Consum. Electron.*, vol. 35, pp. 290–295, Aug. 1989.
- [8] F. Catthoor and H. de Man, "Simulated-annealing-based optimization of coefficient and data word-lengths in digital filters," *Int. J. Circuit Theory and Appl.*, vol. 16, pp. 371–390, 1988.
- [9] J.-B. Chartier, "Traitement numérique des signaux pour la transmission des séquences vidéo: impact de la phase non-linéaire," Tech. Rep. 93–01, INRS-Télécommunications, Jan. 1993.
- [10] A. Chottera and G. Jullien, "Design of two-dimensional recursive digital filters using linear programming," *IEEE Trans. Circuits Syst.*, vol. CAS-29, pp. 817–826, Dec. 1982.
- [11] E. Diethorn and D. Munson, "Finite word length FIR digital filter design using simulated annealing," in *Proc. IEEE Int. Symp. Circuits and Systems*, pp. 217–220, May 1986.
- [12] E. Dubois and W. Schreiber, "Improvements to NTSC by multidimensional filtering," *SMPTE J.*, vol. 97, pp. 446–463, June 1988.
- [13] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*. Prentice Hall, 1984.

- [14] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721–741, Nov. 1984.
- [15] T. Huang, J. Burnett, and A. Deczky, "The importance of phase in image processing filters," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-23, pp. 529–542, Dec. 1975.
- [16] S. Kirkpatrick, C. Gelatt Jr., and M. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, May 1983.
- [17] D. Kodek, "Design of optimal finite wordlength FIR digital filters using integer programming techniques," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-28, pp. 304–308, June 1980.
- [18] J. Konrad, J. Radecki, and E. Dubois, "On the design of finite wordlength IIR filters for video applications," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, pp. IV.341–IV.344, Mar. 1992.
- [19] G. Lampropoulos and M. Fahmy, "A new technique for the design of two-dimensional FIR and IIR filters," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-33, pp. 268–280, Feb. 1985.
- [20] T. Lin, M. Kawamata, and T. Higuchi, "Design of 2-D digital filters with an arbitrary response and no overflow oscillations based on a new stability condition," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 113–126, Feb. 1987.
- [21] M. Manry and J. Aggarwal, "The measurement of phase distortion due to filtering in digital pictures," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-25, pp. 534–541, Dec. 1977.
- [22] N. Metropolis, A. Rosenbluth, M. Rosenbluth, H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.*, vol. 21, pp. 1087–1092, June 1953.
- [23] V. Ouvrard and P. Siohan, "Design of two-dimensional video filters with spatial constraints," in *Signal Process. VI: Theories and Applications (Proc. Sixth European Signal Process. Conf.)*, pp. 1001–1004, Aug. 1992.
- [24] L. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*. Prentice-Hall, 1975.
- [25] J. Radecki, J. Konrad, and E. Dubois, "A comparison of simulated annealing and N-step Newton methods for designing 1-D and 2-D finite wordlength FIR filters," in *Proc. Canadian Conf. Electr. Comp. Eng.*, pp. 53.3.1–53.3.4, Sept. 1990.
- [26] J. Radecki, J. Konrad, and E. Dubois, "Design of finite wordlength IIR filters with prescribed magnitude, group delay and stability properties using simulated annealing," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, pp. 1637–1640, May 1991.
- [27] J. Radecki, J. Konrad, and E. Dubois, "Design of finite wordlength 2-D IIR filters using simulated annealing," in *Signal Process. VI: Theories and Applications (Proc. Sixth European Signal Process. Conf.)*, pp. 953–956, Aug. 1992.
- [28] Y. Wan and M. Fahmy, "N-dimensional symmetries and their applications in digital filters," *Signal Process.*, vol. 19, pp. 103–117, 1990.
- [29] J. Woods, J.-H. Lee, and I. Paul, "Two-dimensional IIR filter design with magnitude and phase error criteria," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-31, pp. 886–893, Aug. 1983.

Financial support:

This work was supported by the Natural Sciences and Engineering Research Council of Canada under Operating Grant OGP0121619.

List of footnotes

1. Coefficients of a computer-designed filter are always expressed by a finite number of bits. However, a 32- or 64-bit floating point number can be considered as a full precision value when compared, for example, with its 8-bit fixed-point approximation.
2. Frequency ω is considered normalized with respect to the sampling frequency.
3. By a full-precision design we mean optimization using double-precision (64 bits) floating-point arithmetic.

List of table captions

Table 1 1-D low-pass linear-phase FIR filters ($K=1$, $|\mathcal{M}|=49$) designed using SA and N -step Newton minimization. Note that for 1-D linear-phase FIR filters $\mathcal{E}_\eta^- = \mathcal{E}_\eta^+$.

Table 2 2-D low-pass linear-phase FIR filters with diamond-shaped pass band ($K=1$, $|\mathcal{M}|=165$; 15 coefficients per line, 11 lines) designed using SA and N -step Newton minimization.

Table 3 1-D low-pass IIR filters ($K=L=5$, $|\mathcal{M}_i|=3$, $|\mathcal{N}_i|=2$, $i = 1, \dots, 5$) designed using SA with $\lambda_s=0.2$.

Table 4 Impact of stability margin on parameters of 1-D 8-bit IIR filter designed using SA with $\lambda_\tau=0.125$, $\lambda_\eta=0.25$, and $\sigma=0.01$.

Table 5 2-D low-pass IIR filters with diamond-shaped pass band ($K=2$, $|\mathcal{M}_1|=25$, $|\mathcal{M}_2|=9$; $L=6$, $|\mathcal{N}_i|=2$, $i = 1, \dots, 6$) designed using SA with $\lambda_s=0.02$.

List of figure captions

Fig. 1 Response of SA-designed 1-D 8-bit FIR ($\lambda_\eta=0.5$) and IIR filters ($\lambda_\tau=0.125$, $\lambda_s=0.2$, $\lambda_\eta=0.25$): (a) $H(\omega)$; (b) $\tau(\omega)$; (c) $\eta(n)$.

Fig. 2 Response of SA-designed 2-D 8-bit low-pass FIR filter with diamond-shaped pass band ($\lambda_\eta=0.0$): (a) $H(\omega_1, \omega_2)$; (b) $\eta(n_1, n_2)$.

Fig. 3 Response of SA-designed 2-D 8-bit low-pass IIR filter with diamond-shaped pass band ($\lambda_\tau=0.02$, $\lambda_s=0.02$, $\lambda_\eta=0.2$): (a) $H(\omega_1, \omega_2)$; (b) $\eta(n_1, n_2)$; (c) $\tau_1(\omega_1, \omega_2)$; (d) $\tau_2(\omega_1, \omega_2)$.

method	b	λ_η	δ_p [dB]	α_s [dB]	\mathcal{E}_η^+
Newton	full prec.	0.0	-0.09	39.87	0.151
Newton	10	0.0	+0.15	34.65	0.150
Newton	8	0.0	+0.36	28.14	0.156
Newton	6	0.0	+0.71	17.36	0.156
SA	10	0.0	-0.15	36.36	0.152
SA	8	0.0	-0.33	31.39	0.156
SA	6	0.0	+0.71	17.36	0.156
SA	8	0.3	+0.44	26.26	0.141
SA	8	0.5	+0.68	21.98	0.117

Table 1:

method	b	λ_η	δ_p [dB]	α_s [dB]	\mathcal{E}_η^+	\mathcal{E}_η^-
Newton	full prec.	0.0	+0.16	33.86	0.134	0.093
Newton	8	0.0	+0.41	25.42	0.147	0.104
Newton	6	0.0	+1.24	17.15	0.156	0.086
SA	8	0.0	+0.25	29.56	0.135	0.092
SA	6	0.0	-0.77	20.49	0.133	0.078
SA	8	0.3	-0.42	24.50	0.100	0.098
SA	8	0.5	-0.76	22.47	0.088	0.078

Table 2:

b	λ_τ	λ_η	δ_p [dB]	α_s [dB]	$\Delta\tau$	$1 - z _{max}$	\mathcal{E}_η^+	\mathcal{E}_η^-
10	0.00125	0.25	-0.29	26.07	2.28	0.071	0.129	0.129
8	0.00125	0.25	-0.36	25.17	2.63	0.109	0.124	0.122
6	0.00125	0.25	-0.64	22.23	3.69	0.206	0.139	0.118
8	0.0	0.25	+0.22	29.11	5.85	0.109	0.135	0.134
8	0.05	0.25	-0.97	18.12	1.78	0.265	0.133	0.100
8	0.0125	0.0	-0.35	24.63	2.29	0.099	0.095	0.191

Table 3:

λ_s	δ_p [dB]	α_s [dB]	$\Delta\tau$	$1 - z _{max}$	\mathcal{E}_η^+	\mathcal{E}_η^-	t_s	t'_s
0.02	-0.36	25.35	4.18	0.051	0.142	0.140	93	43
0.2	-0.36	25.17	2.63	0.109	0.124	0.122	44	29
0.3	-0.26	26.50	4.93	0.161	0.141	0.141	30	27
0.5	-0.38	24.52	5.76	0.199	0.154	0.133	25	26

Table 4:

b	λ_τ	λ_η	δ_p [dB]	α_s [dB]	$\Delta\tau_1$	$\Delta\tau_2$	$1 - z _{max}$	\mathcal{E}_η^+	\mathcal{E}_η^-
8	0.02	0.2	+0.49	21.91	0.410	0.386	0.517	0.099	0.100
6	0.02	0.2	+0.54	20.58	0.571	0.480	0.386	0.092	0.119
8	0.0	0.2	+0.78	19.91	1.283	0.910	0.600	0.092	0.089
8	0.1	0.2	+0.63	22.59	0.129	0.141	0.812	0.136	0.098
8	0.02	0.0	+0.70	21.55	0.695	0.707	0.746	0.154	0.111

Table 5:

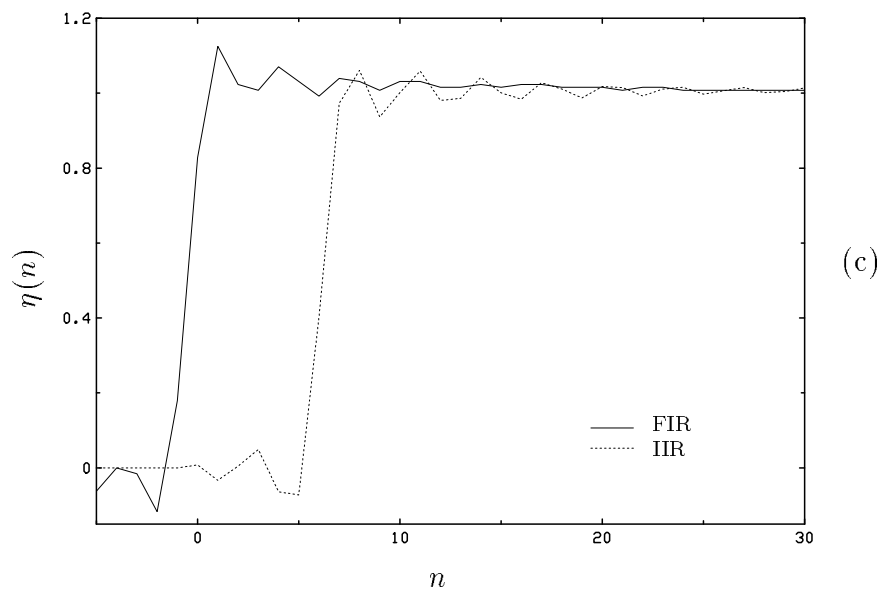
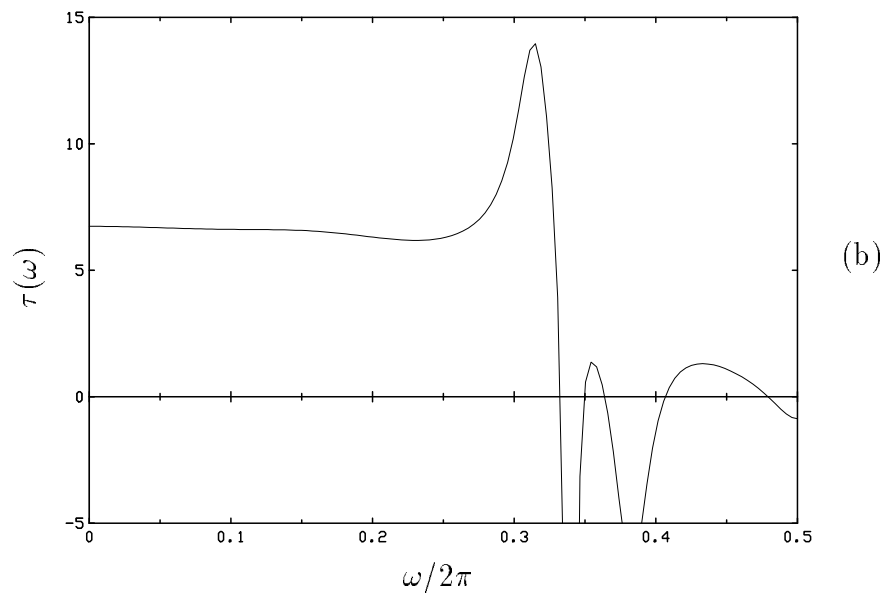
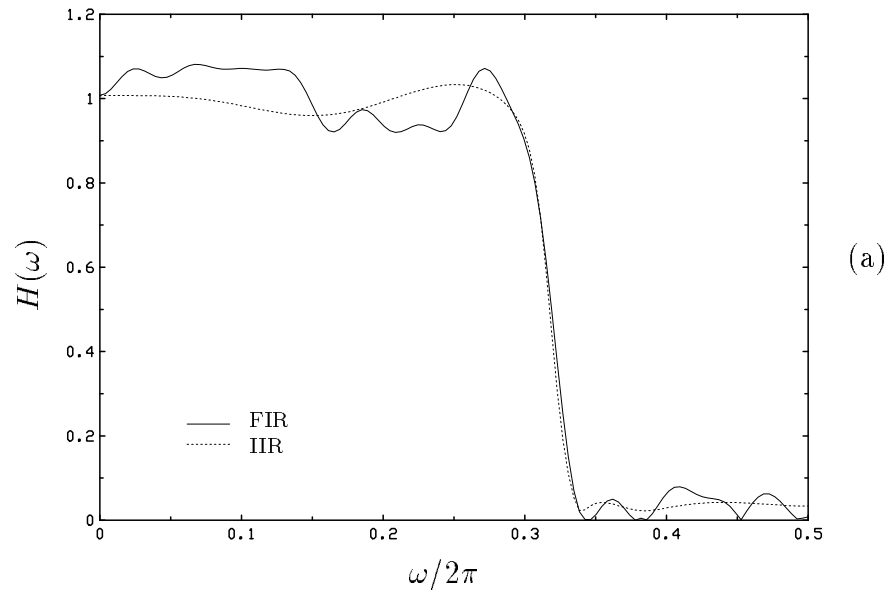
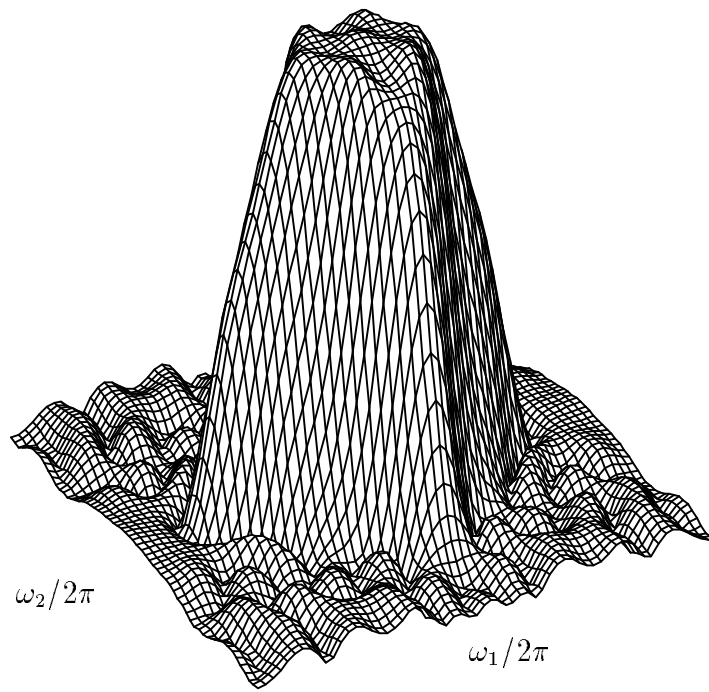
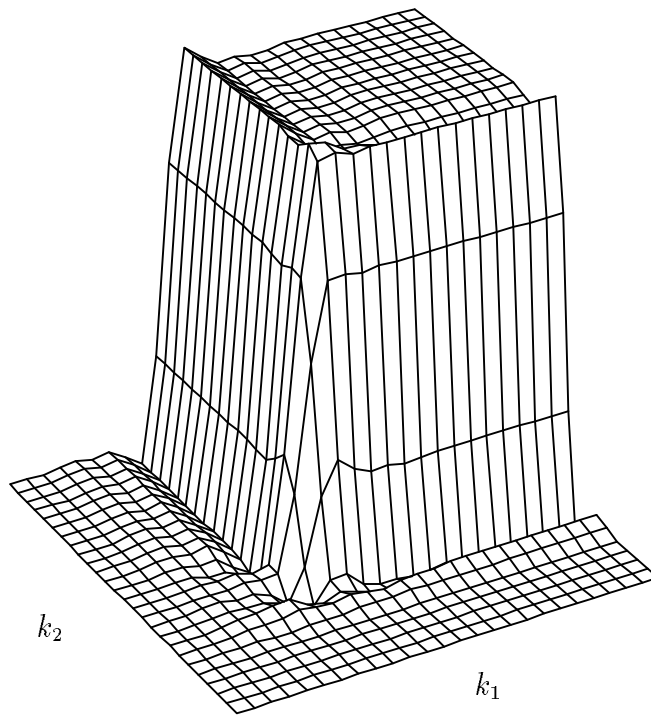


Figure 1:

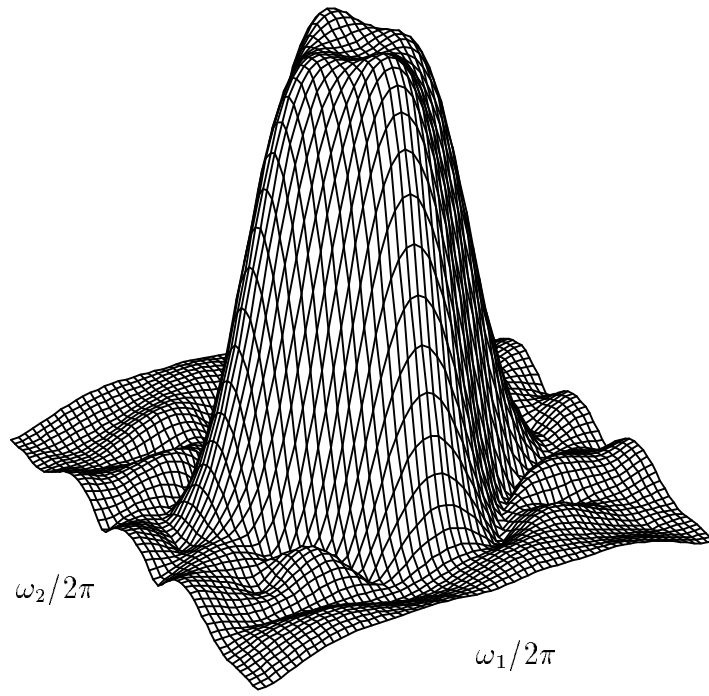


(a)

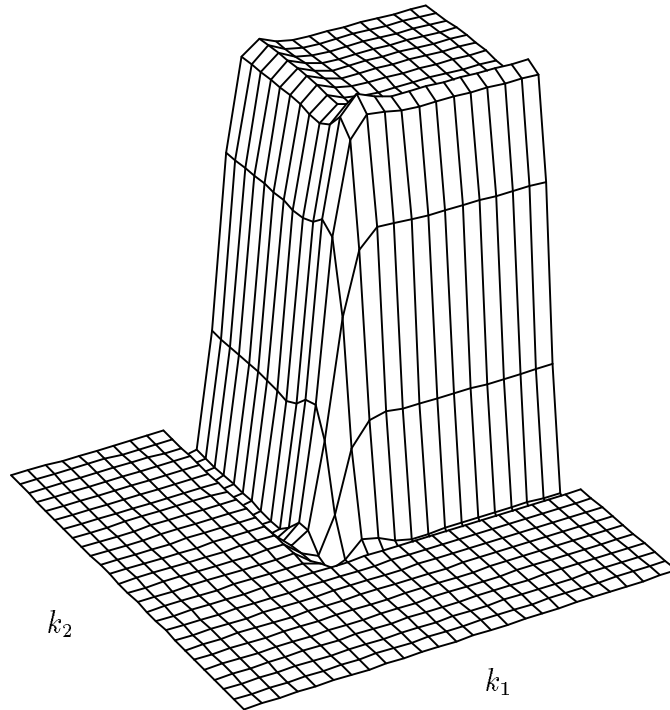


(b)

Figure 2:

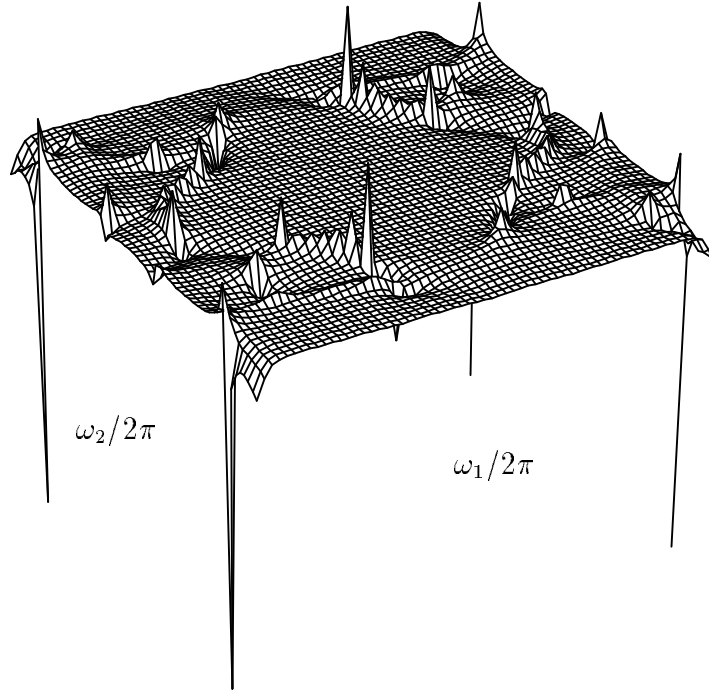


(a)

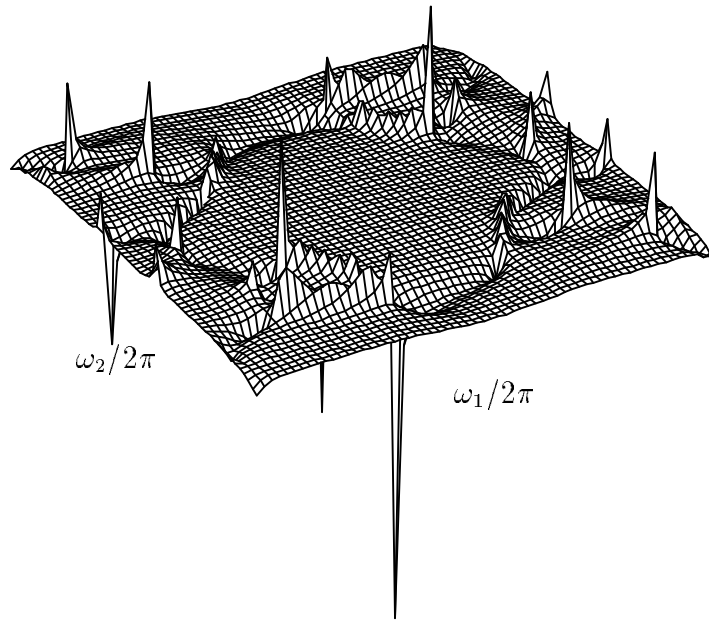


(b)

Figure 3:



(c)



(d)

Figure 3: