Two-Dimensional IFIR Structures Using Generalized Factorizable Filters

Roberto Manduchi

Abstract—In this paper we extend the idea of interpolated FIR (IFIR) filters to the two-dimensional (2-D) case. IFIR filters make for the reduction of the computational weight, in the onedimensional (1-D) case as well as in the 2-D case. In the 1-D case, the justification to such a performance advantage rests upon the relationship between filter order, transition bandwidth and minimax errors for equiripple linear-phase filters. Even though no similar relation is known for minimax optimal multidimensional filters, a qualitatively parallel behavior is shared by a class of suboptimal filters ("Generalized Factorizable") recently introduced by Chen and Vaidyanathan, for which an efficient implementation exists. In our scheme, we use Generalized Factorizable filters for both the stages of the IFIR structure. An interesting problem peculiar to the multidimensional case is the choice of the sublattice which represents the definition support of the first-stage (shaping) filter. We present a strategy to choose (given the spectral support of the desired frequency response) the optimal sublattice, and to design the second-stage (interpolator) filter in order to achieve low overall computational weight.

Index Terms—Multidimensional filters, sampling lattices.

I. Introduction

A. Background

THE IDEA idea of interpolated finite impulse response (IFIR) filters has been introduced by Neuvo *et al.* [16]. In its simplest form, an IFIR structure is an FIR filter whose transfer function can be written as

$$H_{\rm IFIR}(z) = H^{(0)}(z)\hat{H}^{(1)}(z^L)$$
 (1)

where $H^{(0)}(z)$ and $\hat{H}^{(1)}(z)$ are the transfer functions of two FIR filters $h^{(0)}(n)$ and $\hat{h}^{(1)}(n)$ and L is some integer. In other words, an IFIR filter is the cascade of two FIR filters $(h^{(1)}(n)$, the *shaping* filter, and $h^{(0)}(n)$, the *interpolator* filter), with $h^{(1)}(n) \neq 0$ only if n is a multiple of L (then $\hat{h}^{(1)}(n) = h^{(1)}(Ln)$). The interpolator removes the undesired spectral repetitions of the shaping filter (see Fig. 1).

IFIR filters are interesting because they can be implemented efficiently. The number of elementary operations (multiplications or sums) per input sample (called OPS) required to implement an FIR filter is approximately equal to the number

Manuscript received March 14, 1995; revised February 22, 1996. The work of the author was supported by a fellowship from the Associazione Elettrotecnica Italiana (A.E.I), Italy. This paper was recommended by Associate Editor W.-S. Lu.

The author was with the Electrical Engineering and Computer Science Department, University of California, Berkeley, CA 94720 USA. He is now with the Interactive Media Group, Apple Computer, Inc., Cupertino, CA 95014 USA.

Publisher Item Identifier S 1057-7130(97)03644-6.

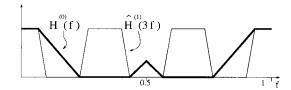


Fig. 1. A 1-D IFIR filter with L=3: frequency responses of $\hat{H}^{(1)}(z^3)$ (the shaping filter) and of $H^{(0)}(z)$ (the interpolator).

N of non-null coefficients of its impulse response¹. Hence, if $N^{(0)}$ is the length of $h^{(0)}(n)$, and $\hat{N}^{(1)}$ is the length of $\hat{h}^{(1)}(n)$, approximately $N_{\rm IFIR} = N^{(0)} + \hat{N}^{(1)}$ OPS's are required to implement the IFIR structure (in spite of the fact that the length of the overall impulse response $h^{(0)} * h^{(1)}(n)$ is $N^{(0)} + L (\hat{N}^{(1)} - 1) > N_{\text{IFIR}}$ if L > 1 and $\hat{N}^{(1)} > 2$. The minimum length N_h of an FIR filter h(n) to meet some prescribed specifics is, for certain classes of filters (e.g., equiripple with narrow passband and transition band), larger than the length $N_{
m IFIR}$ relative to a suitable IFIR structure satisfying the same specifications. Hence, IFIR filters allow for computational savings in such cases. However, the number of memory cells required for the implementation of the IFIR structure is $N^{(0)} + L(\hat{N}^{(1)} - 1) - 1$, which is typically slightly larger than $N_h - 1$ (the number of memory cells required to realize h(n)).

An interesting aspect of the IFIR filter idea is the connection with the theory of multistage implementation of interpolators and decimators [6], [7]. One can easily show that the interpolation or decimation using an IFIR filter can be directly implemented with the multistage (multirate) scheme of Crochiere and Rabiner. The two structures are formally equivalent: the theory developed for the multistage sampling structure conversion [6], [7] can be used to design IFIR filters, and vice versa.

In the literature, the design of IFIR filters approximating ideal low-pass or band-pass frequency response in a minimax sense has been considered [16], [19]. The impulse response of an IFIR filter can be regarded to as the interpolated version of a "decimated" one. Because the impulse response of a selective band optimal FIR filter is typically highly correlated, it is intuitive that a "simple" interpolator should fulfill the purpose. A quantitative analysis of such a notion may be carried out by exploiting the (approximate) analytical relation which holds among the parameters of a minimax filter (filter length N,

 $^{^{1}}$ Actually, in a direct form realization, when no symmetry is present, the number of multiplications per input sample is N, while the number of sums is N-1.

passband and stopband frequencies f_p and f_s , passband and stopband ripples δ_p and δ_s) for small ripples δ_p and δ_s [12],

$$N \propto \frac{-10 \log(\delta_p \delta_s) - 13}{f_s - f_p}.$$
 (2)

From relation (2) we have that, for sufficiently large N, the transition bandwidth $(f_s - f_p)$ is approximately inversely proportional to N (for fixed product $\delta_p \delta_s$).

In the original simple design technique (proposed, for multistage interpolation-decimation schemes, in [6], and for IFIR filters in [16]), the shaping filter and the interpolator are designed independently of each other. The amount of the passband and of the stopband ripples of the overall filter depends on the relative positions of the oscillations of the frequency responses of $h^{(0)}(n)$ and $h^{(1)}(n)$, which are unknown in general. Hence, only upper bounds for the resulting δ_p and δ_s are predictable. If $\delta_p^{(1)}$ and $\delta_s^{(1)}$ are the passband and stopband ripples of the shaping filter, and $\delta_p^{(0)}$ and $\delta_s^{(0)}$ those of the interpolator, from (1) one has that

$$\delta_{p} \leq \delta_{p}^{(1)} + \delta_{p}^{(0)} + \delta_{p}^{(1)} \delta_{p}^{(0)} \simeq \delta_{p}^{(1)} + \delta_{p}^{(0)}$$
(3)
$$\delta_{s} \leq \max(\delta_{s}^{(1)}, \delta_{s}^{(0)} + \delta_{p}^{(1)} \delta_{s}^{(0)}) \simeq \max(\delta_{s}^{(1)}, \delta_{s}^{(0)}).$$
(4)

$$\delta_s \le \max(\delta_s^{(1)}, \delta_s^{(0)} + \delta_p^{(1)} \delta_s^{(0)}) \simeq \max(\delta_s^{(1)}, \delta_s^{(0)}).$$
 (4)

Such worst-case relations may be used for the choice of the specifics of the shaping filter and of the interpolator, to achieve desired maximum values of δ_p and δ_s , as in [6]. If optimal minimax filters are used, the required filter orders may be obtained using the relations described in [12], [11], and [18]. Note that, although in their original work [16] Neuvo et al. suggested the use of a simple first-order or second-order interpolator, higher order interpolators may be used profitably, as in [19].

Several improvements to this simple design procedure have been proposed. Crochiere and Rabiner early realized that adopting multiple stopband (instead of single stopband) interpolators, provides fairly significant filter order reduction [8]. Such an idea was generalized by Saramäki et al. [19] to obtain equiripple behavior of the overall IFIR filter frequency response. They proposed a procedure to iteratively design $h^{(0)}(n)$ and $\hat{h}^{(1)}(n)$ using the Remez exchange algorithm. Their method enables to design optimal (in a minimax sense) IFIR filters; however, it is not clear how to find a multidimensional version of such a technique.

The theory of multidimensional (M-D) multistage sampling structure conversion has been first proposed by Ansari and Lee [1] and by Chen and Vaidyanathan [4], and then developed to some extent by Manduchi et al. [15]. Also in the multidimensional case, the theory of the multistage sampling structure conversion and of IFIR filters are equivalent, and we will deal only with IFIR filters hereinafter. While in [1] and in [4] the necessary conditions (in terms of sampling lattices and spectral support determination) for a multistage scheme—IFIR structure are stated, and in [15] a simple design example is given, no serious attempt to produce efficient two-dimensional (2-D) IFIR filters defined on a given sampling lattice has been proposed in the literature.

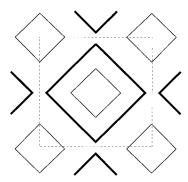


Fig. 2. A 2-D IFIR filter: passband curve of the shaping filter (thin line) and of the interpolator (thick line). The interpolator cancels the undesired spectral repetitions of the shaping filter.

B. Problem Statement

The purpose of this paper is to provide a framework to design 2-D IFIR filters, for a certain class of frequency response shapes widely used in video technology. For the sake of simplicity, only 2-D domains will be studied in this paper. The results can be extended to higher dimensions without major problems.

Let Λ be the lattice of definition of the filter. The simplest IFIR scheme is composed by the cascade of a shaping filter, whose coefficient are not null only on a sublattice Λ_1 (L times less dense than Λ), and of an interpolator. As in the 1-D case, the purpose of the interpolator is to cancel the undesired spectral repetitions of the shaping filter (see Fig. 2).

The following issues must be taken care of in the M-D case:

- 1) M-D sampling lattices admit more than just one sublattice for a given decimation ratio [9], [5]. Each sublattice induces a different geometry of the spectral repetitions of the shaping filter. In this work, we show that, given the desired frequency response mask, certain sublattices make for the easy interpolation of the samples of the shaping filter, while other ones are unsuitable. In our algorithm, all feasible sublattices of definition of the shaping filter are tested. Note that the "feasible" sublattices (i.e., such that the the repetitions of the spectral support of the shaping filter do not overlap) are in a finite number. In fact, increasing the subsampling ratio, the density of spectral repetitions increases; corresponding to some "critical" subsampling ratio, spectral overlap cannot be avoided.
- 2) The frequency response of an M-D filter cannot be easily characterized as in the 1-D case; in other words, a filter's passband or stopband region can exhibit any shape (while in the case of low-pass 1-D, they are bound to be segments). Devising a general technique for the IFIR system design seems an overwhelming task. Fortunately, for a large variety of applications, only certain subclasses are of interest. For example, frequency responses with passband in the shape of a parallelogram (typically a diamond) are suitable for the sampling structure conversion of video signals [9], [20], [15], as well as for many other applications. We will concentrate in this work on such a class of spectral masks.

We design the shaping filter and the interpolator separately, in the spirit of the early work of Neuvo $et\ al.$ [16]. Among the variety of 2-D filter design techniques available in the literature [10], we have chosen the one recently proposed by Chen and Vaidyanathan in [2], [4], [3], and studied later by Manduchi [14]. We will call such filters "Generalized Factorizable" (GF) [14]. GF filters are designed starting from two 1-D filters, taking the tensor product of their impulse responses, subsampling it on a suitable sublattices of Z^2 , and finally reordering the samples on the desired definition lattice. The main advantages of GF filters are as follows.

- 1) They "naturally" provide spectral supports in the shape of parallelograms.
- 2) The design procedure is very fast (the computational burden is due to the design of two 1-D filters) and simple, lending itself to use in automatic design systems (CAD).
- They admit an efficient "generalized factorizable" implementation, which can reduce effectively the computational weight.
- 4) It is possible to control the size and some characteristics of the impulse response.

In addition, GF filters exhibit the interesting property that approximate (worst case) relations can be found among the filter parameters of interest (filter size, transition region, passband and stopband ripples). Exploiting such relations it is possible to predict lower bounds on the performance attainable by 2-D IFIR filters.

A key point in the 2-D IFIR design is the design of the interpolator. While the spectral support of the shaping filter is constrained by the shape of the desired overall frequency response $D(\mathbf{f})$, the interpolator $H^{(0)}(\mathbf{f})$ is only required to be unitary on the passband region of $D(\mathbf{f})$, and vanishing on the undesired spectral repetitions of the shaping filter. The criterion for the choice of the interpolator's spectral mask is the minimization of the computational weight. It turns out that, in general, several feasible solutions can be devised. We introduce a computational geometry algorithm to derive all the "optimal" spectral masks for the interpolator.

In the experimental section, we consider both the nonfactorizable and the "generalized factorizable" implementation of the filters. In this way, we provide a reasonable estimate of the system's performance when filters other than GF are employed. The results show that, in the nonfactorizable case, improvements in terms of computational weight, comparable to the 1-D case, are achievable, depending on the shape of the passband and of the stopband regions. Interestingly enough, the situation is not quite as simple using the "generalized factorizable" implementation. It is shown that, depending on the geometry of the decimation lattice used in the shaping filter's design (which, in turn, depends on the shape of the desired spectral support), the use of an IFIR structure may or may not reduce the overall computational weight.

The paper is organized as follows. Section II briefly reviews the theory of GF filters. Section III introduces our 2-D IFIR design procedure. Section IV shows the experimental example, and Section V has the conclusions. In order to make the paper self-contained, some nonstandard notions regarding multidimensional sampling lattices are reported in Appendix A. In Appendix B we describe a computational geometry algorithm to design the interpolator's spectral mask.

We conclude this Introduction with the nomenclature used throughout the paper.

R, Z Set of real, integer numbers.

 We denote vectors by lower case boldface letters and matrices by upper case boldface letters.
 Their entries are named after the following example:

$$\boldsymbol{a} \stackrel{\Delta}{=} (a_1, a_2)^T; \boldsymbol{A} \stackrel{\Delta}{=} (\boldsymbol{a}_1 | \boldsymbol{a}_2) \stackrel{\Delta}{=} \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$$
(5)

where symbol "T" means vector/matrix transposition.

 $\mathbf{A}^{-T} = (\mathbf{A}^{-1})^T.$

I Identity matrix.

• Par(V) (where $V = (v_1|v_2|\cdots|v_M)$) = the parallelepiped centered at the origin $\Sigma_{i=1}^{M} \alpha_i v_i, -1 \leq \alpha_i \leq 1$.

 $\mathcal{R}(\boldsymbol{u})$ Parallelepiped centered at the origin with edges pairwise parallel to the axes $\{\boldsymbol{a}: |a_i| \leq |u_i|\} = \operatorname{Par}(\operatorname{diag}(u_1, u_2, \cdots, u_M))$

 $A \setminus B$ Difference between sets A and B (i.e., the set of elements of A that do not belong to B).

LAT(C) Lattice with basis C.

 Λ^* Dual of lattice $\Lambda = LAT(\mathbf{C}^{-T})$.

We denote a signal $h(\cdot)$ defined on an M-dimensional lattice $\Lambda = \text{LAT}(C)$ with h(a) (where a is assumed to belong to Λ) or with h(Cn) (where $n \in Z^M$). The two notations are interchanged liberally.

 We use term "filter" meaning the filter's impulse response (denoted by lower case letter) and its frequency response (denoted by upper case letters)

• Conventional filter size = the number of filter coefficients not forced to zero.

OPS Elementary operation per input sample.

 $LDFL(\Lambda)$ Least dense factorizable lattice containing lattice Λ .

DFS(Λ) Densest factorizable sublattice contained in lattice Λ .

II. GENERALIZED FACTORIZABLE FILTERS

In recent papers [2]–[4], Chen and Vaidyanathan introduced a class of M-D filters, which are designed starting from M 1-D low-pass prototypes. The resulting frequency response has passband in the shape of a parallelogram. Although an M-D filter designed this way is not factorizable, its polyphase components are "generalized factorizable", in the sense that they can be written as the tensor product of M 1-D filters oriented along suitable directions.

Chen and Vaidyanathan's algorithm is actually the only known technique to design *Generalized Factorizable* (GF) filters [14]. GF filters admit a "generalized factorizable" imple-

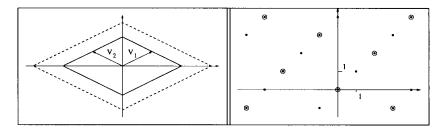


Fig. 3. Left: spectral mask of the filter in the example of Section II. Solid line: passband curve. Dashed line: stopband curve. Right: decimation lattices LAT(A) (dots) and LAT(AH) (circles).

mentation. If N_i is the length of the ith 1-D filter used in the design algorithm, only $\sum_{i=1}^{M} N_i$ OPS's are required to realize the GF filter using such an implementation. On the other side, if the direct form realization is adopted, the number of OPS's is equal to the "conventional size" of the M-D filter, i.e., to the number samples of its impulse response not constrained to zero. The conventional filter size is approximately equal to $(\prod_{i=1}^{M} N_i)/|\det(A)|$, where A is the decimation matrix used in the design, as explained later in this section.

In the following, we briefly summarize some important notions relative to GF filters, which are instrumental to our theory. More details, as well as the proofs, can be found in [2]–[4], [14].

A. Chen and Vaidyanathan's Design Algorithm

Consider a sampling lattice $\Lambda = LAT(\mathbf{C})$. Let

$$Par(\mathbf{P}) = \left\{ \sum_{i=1}^{M} \alpha_i \mathbf{p}_i, -1 \le \alpha_i \le 1 \right\}$$
 (6)

be a parallelepiped (representing the desired passband region), characterized by matrix $P = (p_1|p_2|\cdots|p_M)$. We assume that P has only rational entries.

The core of Chen and Vaidyanathan's algorithm can be summarized as follows. Let $\{q_1(n), q_2(n), \cdots, q_M(n)\}$ be ideal 1-D filters with frequency responses $Q_i(f)$ such that

$$Q_i(f) = \begin{cases} 1, & |f| \le f_{p_i} \\ 0, & f_{p_i} < |f| \le 0.5. \end{cases}$$
 (7)

The factorizable filter

$$\overline{h}(\mathbf{n}) = \left(\prod_{i=1}^{M} q_i(n_i)\right) \cdot |\det(\mathbf{A})| \tag{8}$$

has spectral support in $\mathcal{R}(\boldsymbol{f}_p),$ where $\boldsymbol{f}_p=(f_{p_1},\ f_{p_2},\ \cdots,f_{p_M}).$ The filter

$$h(\mathbf{C}\mathbf{n}) = \overline{h}(\mathbf{A}\mathbf{n}) \tag{9}$$

obtained subsampling $\overline{h}(n)$ on LAT(A) and reordering the samples on LAT(C), has spectral support in $C^{-T}A^T\mathcal{R}(f_p)$.

Our purpose is to find a matrix \boldsymbol{A} and passband frequencies $\{f_{p_i}\}$ such that

$$\mathbf{C}^{-T}\mathbf{A}^{T}\mathcal{R}(\mathbf{f}_{p}) = \operatorname{Par}(\mathbf{P}). \tag{10}$$

Let

$$\overline{A} = P^T C \operatorname{den}(P^T C) = DA,$$

$$D = \operatorname{diag}(D_1, D_2, \dots, D_M)$$
(11)

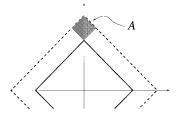


Fig. 4. Example of "corner region" of a diamond-shaped spectral mask. Solid line: passband curve. Dashed line: stopband curve.

where $den(P^TC)$ is the least positive integer such that $den(P^TC)P^TC$ is integral, and D_i is the greatest common divisor of the entries of the i-th row of \overline{A} . Note that A is integral.

Now let $f_{p_i} = D_i/\text{den}(\mathbf{P}^T\mathbf{C})$. It is easy to prove that, with this choice of \mathbf{A} and $\{f_{p_i}\}$, relation (10) is satisfied.

In practice, we will assume that the passband and stopband regions of the M-D filter are shaped like parallelograms with pairwise parallel edges $Par(\mathbf{P}_p)$ and $Par(\mathbf{P}_s)$, where $\mathbf{P}_s = \alpha \mathbf{P}_p$ for some $\alpha > 1$, and will use zero-phase minimax low-pass FIR filters $\{q_i(n)\}$ [14].

As an example, consider the following design parameters: C = I (i.e., $\Lambda = Z^2$), $\mathbf{v}_1 = (1/20, 1/40)^T$, $\mathbf{v}_2 = (-1/20, 1/40)^T$, $\alpha = 3/2$ (see Fig. 3). Using Chen and Vaidyanathan's algorithm, we obtain the sampling matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}$$

and the passband frequencies of the 1-D filters $f_{p_1}=f_{p_2}=1/40$ (the stopband frequencies being $f_{s_1}=\alpha f_{p_1}$ and $f_{s_2}=\alpha f_{p_2}$).

B. Minimax Relationships

It is possible to derive worst-case relations between the product of the passband and of the stopband ripples of the frequency response and some measure of the "transition region", for a fixed conventional size of the GF filter [14]. In general, one can expect that, for a fixed conventional filter size, the product of the passband and of the stopband ripples increases as the area of the "corner region" \mathcal{A} depicted in Fig. 4 decreases. This result, which is reminescent of relation (2) among the minimax parameters of an optimal 1-D filter, should not surprise: a 2-D GF filter is obtained subsampling the tensor product of two 1-D filters, and it inherits their properties. In particular, the "corner region" area is proportional to the product of the transition bands of the two 1-D filters.

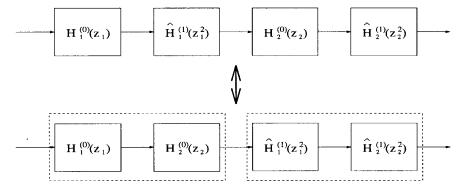


Fig. 5. The cascade of two 1-D IFIR filters (on z_1 and z_2 , respectively) is equivalent to a 2-D IFIR filter.

C. Impulse Response Decimation

Suppose one is given a GF filter $h(\mathbf{a})$ defined on $\Lambda = \text{LAT}(\mathbf{C})$ with spectral support approximating $\text{Par}(\mathbf{P})$. Let $\Lambda_1 = \text{LAT}(\mathbf{CH})$ be a sublattice of Λ , and assume that $\text{Par}(\mathbf{P})$ is contained within some elementary cell of Λ_1^* . In order to design a filter $g(\mathbf{a})$ defined on Λ_1 having the same spectral support of $H(\mathbf{f})$ (within an elementary cell of Λ_1^*), two procedures are available:

- 1) Set $g(\mathbf{CHn}) = |\det(\mathbf{H})|h(\mathbf{CHn})$, i.e. subsample $h(\mathbf{a})$ on Λ_1 and adjust the gain. The conventional size N_g of $g(\mathbf{a})$ is approximately equal to the conventional size N_h of $h(\mathbf{a})$, divided by $|\det(\mathbf{H})|$.
- 2) If the least dense factorizable lattice (LDFL, see Appendix A) containing LAT(AH) is less dense than Z^M , one can compute the new sampling matrix \tilde{A} and the new passband and stopband frequencies of the 1-D filters for the design of $g(\mathbf{a})$. It can be shown [13] that, if the 1-D filters of (7) are forced to exhibit the same passband and stopband ripples as in the design of $h(\mathbf{a})$, the conventional size of $g(\mathbf{a})$ will be again approximately equal to $N_h/|\det(\mathbf{H})|$.

Note that, in general, the passband and stopband ripples of the filter's frequency response increase with the absolute value of the determinant of the decimation matrix used in the filter design [3], [14]. Hence, if the lengths of the 1-D filters used in the design of H(f) and of G(f) are the same (as in the first proposed method), G(f) is likely to exhibit larger ripples than H(f). The second procedure allows to use a decimation matrix \hat{A} with $|\det(\hat{A})| < |\det(AH)|$, and therefore to reduce the ripples of G(f).

D. Generalized Factorizable Implementation Complexity

We consider here, for simplicity's sake, only the 2-D case. If N_1 and N_2 are the lengths of the 1-D filters used in the design of $h(\boldsymbol{a})$ defined on $\Lambda = \text{LAT}(\boldsymbol{C})$, then, adopting the generalized factorizable implementation described in [3], $N_1 + N_2$ OPS's are required (although the conventional filter size is $N_1N_2/|\det(\boldsymbol{A})|^2$. One can determine an approximate relation between the number of OPS's required for the generalized factorizable implementation of $h(\boldsymbol{a})$ and of a filter $g(\boldsymbol{a})$ defined

on a sublattice $\Lambda_1 = LAT(CH)$, designed using the first procedure of point 3 above [14]. Let $\operatorname{diag}(R_1, R_2), R_1, R_2 > 0$ be a basis of LDFL(LAT(AH)) (it can be shown that, using Chen and Vaidyanathan's technique, $LDFL(LAT(A)) = Z^2$. Then the number of OPS's required for the realization of $g(\mathbf{a})$ is $N_1/R_1 + N_2/R_2$. Thus, we can draw the following important observation: the number of OPS's required for the generalized factorizable implementation of a GF filter defined on Λ , and of its version subsampled on a sublattice Λ_1 , may or may not differ. In particular, if LDFL(LAT(A)) = LDFL(LAT(AH)), the number of OPS's required to realize $h(\mathbf{a})$ and $q(\mathbf{a})$ will be the same (note that in this case, the two design procedures presented in point 3 above coincide). Such a result is in contrast with the case of nonfactorizable implementation, where the number of OPS's is reduced by a factor approximately equal to the subsampling ratio $|\det(\boldsymbol{H})|$.

As an example, consider the filter specs for the example of point 1 above, for the quincux sublattice $\Lambda_1 = LAT(CH)$ with

$$H = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$
.

Lattices LAT(\boldsymbol{A}) and LAT($\boldsymbol{A}\boldsymbol{H}$) are represented in Fig. 3. It is easily inferred that LDFL(LAT(\boldsymbol{A})) = LDFL(LAT($\boldsymbol{A}\boldsymbol{H}$)) = Z^2 , hence the same number of OPS's is required for the generalized factorizable implementation of $h(\boldsymbol{a})$ and of $g(\boldsymbol{a})$, although the support of $g(\boldsymbol{a})$ has approximately one half of the samples of the support of $h(\boldsymbol{a})$.

E. Symmetries

We consider here only the 2-D case.

A filter $h(\boldsymbol{Cn})$ designed following Chen and Vaidyanathan's algorithm (with zero-phase $\{q_i(n)\}$) satisfies the property $h(\boldsymbol{Cn}) = h(\boldsymbol{-Cn})$ (so that $H(\boldsymbol{f})$ is zero-phase as well) [3]. In addition, if the upper Hermite normal form matrix \boldsymbol{A}^u associated to the decimation matrix \boldsymbol{A} (see Appendix A) satisfies property $A_{1,1}^u = 2A_{1,2}^u$, or if it is diagonal, then the following symmetry property holds [14]:

$$h(Cn) = h(C\overline{n}), \quad n = (n_1, n_2)^T, \quad \overline{n} = (n_1, -n_2)^T.$$
 (12)

Such quadrantal-like symmetry property can be exploited to reduce the number of multiplications per input sample if the direct form realization (instead of the generalized factorizable one) is used.

 $^{^2}$ A more precise computation of the number of OPS's for the generalized factorizable implementation involves the determination of DFS(LAT(A)), and is discussed in [13].

In the previous example, the upper Hermite normal form associated to \mathbf{A} is $\mathbf{A}^u = \begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix}$. Since $A^u_{1,1} \neq 2A^u_{1,2}$, the quadrantal-like symmetry is not verified here (in spite of the inherent symmetry of the spectral support, see Fig. 3).

III. 2-D IFIR FILTERS

In this section we give a formal definition of a 2-D IFIR structure, and introduce our design algorithm.

IFIR filters can be profitable in the 2-D case as well as in the 1-D case. However, in the evaluation of the achievable computational weight reduction, it is important to consider how the filters are implemented. In particular, the use of the (generalized) factorizable implementation affects dramatically the improvements attainable via an IFIR scheme. We discuss a simple example in Section III-A, which should make such an argument clear. Then, the general 2-D IFIR scheme is introduced in Section III-B. In the design of 2-D IFIR filters, we have basically two degrees of freedom: the choice of the sublattice of definition for the shaping filter, and the determination of the interpolator's spectral support. We adopt GF filters for both the shaping filter and the interpolator. The use of GF filters allows for a simple geometric characterization of the filters' frequency response, which we exploit for the design of the interpolator. In short, we design a set of interpolators which maximize the "corner area" of the transition region. A simple computational geometry algorithm to determine such candidates is described in Appendix B.

Our technique yields a (small) set of feasible IFIR schemes, characterized by the couple subsampling lattice-interpolator. The final choice can be done by direct inspection, evaluating the computational weight and memory requirement for their realization.

A. Factorizable IFIR Filters: A Simple Example

Consider a factorizable filter $h(\mathbf{n})$ defined on Z^2 : $h(\mathbf{n}) = h_1(n_1)h_2(n_2)$. Let N_1 and N_2 be the lengths of $h_1(n_1)$ and $h_2(n_2)$, respectively, so that N_1N_2 is the size of $h(\mathbf{n})$. Implementing $h(\mathbf{n})$ without taking into account the factorability requires N_1N_2 OPS's. If we adopt the factorizable implementation, only $N_1 + N_2$ OPS's are required. Assume now to use a 1-D IFIR structure for both $h_1(n)$ and $h_2(n)$. For instance, we may have

$$H_i(z) = H_i^{(0)}(z)\hat{H}_i^{(1)}(z^2), \qquad i = 1, 2.$$
 (13)

Let $N_i^{(0)}$ and $\hat{N}_i^{(1)}$ be the lengths of filters $h_i^{(0)}(n)$ and $\hat{h}_i^{(1)}(n)$ (without null samples interleaved!) respectively, and assume $N_i^{(0)} + \hat{N}_i^{(1)} = N_i/2$ (for details about actual performances of 1-D IFIR filters, see [16], [19]). Using the factorizable implementation of the 2-D filter, now only $(N_1 + N_2)/2$ OPS's are required. Assume now to "forget" that the filter is factorizable. It is readily seen (see Fig. 5) that the whole structure is equivalent to the cascade of two 2-D filters, the first of whom has non-null samples only on lattice LAT(2I). The number of OPS's required to implement the

cascade of two filters is the sum of their conventional sizes, i.e.

$$N_{\rm IFIR} = N_1^{(0)} N_2^{(0)} + \hat{N}_1^{(1)} \hat{N}_2^{(1)}.$$
 (14)

If $N_1=N_2\stackrel{\Delta}{=}N, N_1^{(0)}=N_2^{(0)}\stackrel{\Delta}{=}N^{(0)}$ and $\hat{N}_1^{(1)}=\hat{N}_2^{(1)}\stackrel{\Delta}{=}\hat{N}^{(1)}$, then $N_{\rm IFIR}=N^2/4-2N^{(0)}\hat{N}^{(1)}$. Hence, using such an IFIR scheme with a nonfactorizable implementation requires less than 25% of OPS's than in the direct case. In general, using 1-D IFIR structures of the form $H_i^{(0)}(z)\hat{H}_i^{(1)}(z^{M_i})$, the overall computational burden can be reduced roughly by a factor $\Sigma_i M_i$ in the factorizable case, and by a factor $\Pi_i M_i$ in the nonfactorizable case.

B. The General Case

Let $\Lambda = \mathrm{LAT}(C)$ be the signal definition lattice, and consider a sequence (descending lattice chain [15]) $\mathcal{L} = (\Lambda_0, \Lambda_1, \cdots, \Lambda_{M-1})$ of lattices such that

$$\Lambda_0 \stackrel{\triangle}{=} \Lambda \supset \Lambda_1 = LAT(\mathbf{C}\mathbf{H}_1) \supset \cdots \supset \Lambda_{M-1}$$

= LAT(\mathbb{C}\mathbf{H}_1\mathbf{H}_2\cdots\mathbf{H}_{M-1}). (15)

We define an IFIR structure on \mathcal{L} as the cascade of M FIR filters $(h^{(0)}(\boldsymbol{a}), h^{(1)}(\boldsymbol{a}), \cdots, h^{(M-1)}(\boldsymbol{a}))$ defined on Λ , such that $h^{(i)}(\boldsymbol{a})$ has non-null coefficients only on Λ_i . Note that $H^{(i)}(\boldsymbol{f})$ is periodic on Λ_i^* .

In this work, we consider only the case M=2, and define $\mathbf{H}=\mathbf{H}_1$. Note that in any elementary cell of Λ^* there are $|\det(\mathbf{H})|$ spectral repetitions of $H^{(1)}(\mathbf{f})$.

Let $D(\mathbf{f})$ be the ideal frequency response to be approximated by $H^{(0)}(\mathbf{f})H^{(1)}(\mathbf{f})$. We will consider here $D(\mathbf{f})$ of the form

$$D(\mathbf{f}) = \begin{cases} 1, & \mathbf{f} \in \text{Par}(\mathbf{v}_1 | \mathbf{v}_2) \\ 0, & \mathbf{f} \in \mathcal{V} \setminus \text{Par}(\alpha \mathbf{v}_1 | \alpha \mathbf{v}_2), \alpha > 1 \end{cases}$$
(16)

where v_1, v_2 are two noncollinear vectors, α is a real number larger than 1, and \mathcal{V} is a suitable elementary cell of Λ^* such that $\operatorname{Par}(\alpha v_1 | \alpha v_2) \subset \mathcal{V}$. Note that we allow for a transition region $\operatorname{Par}(\alpha v_1 | \alpha v_2) \backslash \operatorname{Par}(v_1 | v_2)$. We also put the condition that $\operatorname{Par}(\alpha v_1 | \alpha v_2)$ is contained within some elementary cell \mathcal{V}_1 of $\Lambda_1^* = \operatorname{LAT}((CH)^{-T})$.

- 1) The Shaping Filter: We design the shaping filter $h^{(1)}(\mathbf{a})$ defined on Λ_1 such that, within \mathcal{V}_1 , its frequency response approximates $D(\mathbf{f})$. Then, we "interpolate" it on Λ by adding zero-valued samples on $\Lambda \setminus \Lambda_1$.
- 2) The Interpolator: The purpose of the interpolator is to cancel the $|\det(\boldsymbol{H})| 1$ undesired spectral repetitions of $H^{(1)}(\boldsymbol{f})$ on $\Lambda_1^* \backslash \Lambda^*$. The requirements for the interpolator, therefore, are

$$\mathcal{P}^{(0)} \supset \operatorname{Par}(\boldsymbol{v}_1 | \boldsymbol{v}_2),$$

$$\mathcal{S}^{(0)} \supset \left\{ \operatorname{Par}(\alpha \boldsymbol{v}_1 | \alpha \boldsymbol{v}_2) + \boldsymbol{s}, \boldsymbol{s} \in (\Lambda_1^* \cap \mathcal{V}) \setminus \{ \boldsymbol{0} \} \right\}$$
(17)

where $\mathcal{P}^{(0)}$ and $\mathcal{S}^{(0)}$ are the passband and stopband regions of $H^{(0)}(\boldsymbol{f})$.

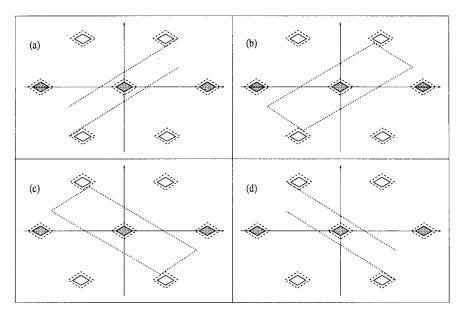


Fig. 6. Spectral repetitions of $H^{(1)}(\mathbf{f})$ on the points of Λ_1^* . Solid line: passband curve; dashed line: stopband curve; filled diamonds: spectral repetitions on the points of Λ^* ; dotted line: feasible stopband curves $\mathcal{S}^{(0)}$ of the interpolator corresponding to the maximal rectangles. Note that case (a) and (d) (corresponding to rectangles R_1 and R_4 in Fig. 11) collapse into generalized 1-D filters.

Note that (17) does not constrain the frequency response of $H^{(0)}(f)$ in points that belong neither to the repetitions of $\Pr(v_1|v_2)$ on Λ^* (where it should be unitary) nor to the repetitions of $\Pr(\alpha v_1|\alpha v_2)$ on $\Lambda_1^*\backslash\Lambda^*$ (where it should be vanishing). Hence, the set of admittable interpolator's frequency response shapes is in general very wide, and, in order to devise a simple design algorithm, we should constrain our choice to some subclass. In our procedure, we design the interpolator as a GF filter, with passband and stopband regions' sides pairwise parallel to those of $\Pr(v_1|v_2)$, satisfying (17). This choice can be justified considering that the sides of the passband region of $H^{(1)}(f)$ and of its copies on $\Lambda_1^*\backslash\Lambda^*$ to be cancelled are pairwise parallel.

Within the chosen filter class, we should pick the filter with the least associated computational weight, such that the overall frequency response requirements are met. As recalled in Section II, in general the minimum filter size to achieve some minimax specifications decreases with the corner area of the transition region. Therefore, our rule for the choice of the passband and stopband regions of the interpolator will be maximizing such a corner area, while satisfying (17).

It is straightforward that the passband region $\mathcal{P}^{(0)}$ should coincide with $\operatorname{Par}(\boldsymbol{v}_1|\boldsymbol{v}_2)$. The choice of the "optimal" stopband curve may be seen as a computational geometry problem. In particular, the solution is not necessarily unique: more than one parallelogram representing $\mathcal{S}^{(0)}$, characterized by the same maximal corner area, may exist.

Our algorithm is effectively simplified if one considers the "transformed" signals (see (9))

$$\overline{h}^{(0)}(\mathbf{A}\mathbf{n}) = h^{(0)}(\mathbf{C}\mathbf{n}), \quad \overline{h}^{(1)}(\mathbf{A}\mathbf{n}) = h^{(1)}(\mathbf{C}\mathbf{n}).$$
 (18)

Now all the passband and stopband regions are rectangular with sides pairwise parallel to the cartesian axes. The passband region $Par(v_1|v_2)$ is transformed into rectangle $Par(A^{-T}C^T(v_1|v_2)) \stackrel{\triangle}{=} \mathcal{R}(u)$ for some u, the stopband re-

gion (of $H^{(1)}(\mathbf{f})$)Par($\alpha \mathbf{v}_1 | \alpha \mathbf{v}_2$) is transformed into rectangle $\mathcal{R}(\alpha \mathbf{u})$, and the frequency repetition lattices Λ^* and Λ_1^* are transformed into $\Gamma^* \stackrel{\triangle}{=} \mathrm{LAT}(\boldsymbol{A}^{-T})$ and $\Gamma_1^* \stackrel{\triangle}{=} \mathrm{LAT}((\boldsymbol{AH})^{-T})$, respectively. The stopband region $\mathcal{S}^{(0)}$ of $H^{(0)}(\boldsymbol{f})$ is transformed into rectangle $\mathcal{R}(b)$ for some b. The larger the area of $\mathcal{R}(\boldsymbol{b})$, the larger the corner area of the corresponding transition region. Hence, we should find the "widest" rectangles contained in some (rectangular) elementary cell of Γ^* , which do not overlap any repetition of $\mathcal{R}(\alpha \mathbf{u})$ on the points of Γ_1^* . We will call such rectangles the maximal rectangles. More precisely, a rectangle centered at the origin with sides pairwise parallel to the axes is maximal if it cannot be expanded along any direction, without bumping into some repetition of itself on the points of Γ^* , or into some repetition of $\mathcal{R}(\alpha u)$ on the points of Γ_1^* . Maximal rectangles are the transformed versions of the candidates from which the support of the stopband region $\mathcal{S}^{(0)}$ of $H^{(0)}(\mathbf{f})$ is chosen.

Once the maximal rectangles are determined, one can verify by direct inspection which interpolator requires the lowest computational weight to attain the desired overall minimax specifics. Note that typically there exist very few maximal rectangles, so that not many tests are necessary.

A simple algorithm to find all maximal rectangles, given $\Gamma^*, \Gamma_1^*, \boldsymbol{u}$ and α , is described in Appendix B.

3) A Study Case: Consider the design parameters introduced in the examples of Section II:

$$C = I, H = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}, v_1 = (1/20, 1/40)^T$$

 $v_2 = (-1/20, 1/40)^T, \alpha = 3/2.$

The repetitions of $\operatorname{Par}(v_1|v_2)$ on the points of Λ_1^* are shown in Fig. 6. The repetitions centered at the points of Λ^* are represented by filled diamonds. The repetitions centered at the points of $\Lambda_1^* \backslash \Lambda^*$ (empty diamonds) are to be cancelled by $H^{(0)}(f)$.

 $TABLE\ I$ Enumeration of the Upper Hermite Normal Form Matrices with Determinant Ranging from 2 to 5

$$\mathbf{H}_{2} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{H}_{3} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{H}_{4} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad \mathbf{H}_{5} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{H}_{6} = \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{H}_{7} = \begin{pmatrix} 3 & 2 \\ 0 & 2 \end{pmatrix} \qquad \mathbf{H}_{8} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \qquad \mathbf{H}_{9} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{H}_{10} = \begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{H}_{11} = \begin{pmatrix} 4 & 2 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{H}_{12} = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{H}_{13} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\mathbf{H}_{14} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \qquad \mathbf{H}_{15} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \qquad \mathbf{H}_{16} = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{H}_{17} = \begin{pmatrix} 5 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{H}_{18} = \begin{pmatrix} 5 & 2 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{H}_{19} = \begin{pmatrix} 5 & 3 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{H}_{20} = \begin{pmatrix} 5 & 4 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{H}_{21} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$$

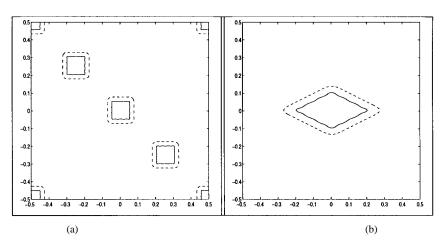


Fig. 7. Contour plot of (a) $\overline{H}(f)$ and of (b) H(f) within an elementary cell of Z^2 . Solid line: passband curve; dashed line: stopband curve.

Let us recall from Section II that, following Chen and Vaidyanathan's algorithm, we obtain the sampling matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}$$

and the passband frequencies of the 1-D filters $f_{p_1}=u_1=f_{p_2}=u_2=1/40.$ Thus

$$A^{-T} = \begin{pmatrix} 1/4 & 1/2 \\ -1/4 & 1/2 \end{pmatrix}$$
 and $(AH)^{-T} = \begin{pmatrix} -1/8 & 1/2 \\ -3/8 & 1/2 \end{pmatrix}$.

The spectral support of the "transformed" filter $h(\mathbf{An})$ is shown in Fig. 11, together with the four maximal rectangles found using the algorithm of Appendix B. These rectangles map into the stopband curves depicted in Fig. 6. Due to the symmetry of the spectral mask of $H^{(1)}(\mathbf{f})$ and of the

repetition lattice, there are two symmetric couples of such curves (namely, couple (a) and (d) and couple (b) and (c) in Fig. 6.

Note that rectangles R_1 and R_4 in Fig. 11 are extremal (see Appendix B): their periodic repetitions on the points of Γ^* form horizontal (vertical) stripes in R^2 . Hence, the corresponding stopband curves are oriented stripes (see Fig. 6). In such a case, the interpolator $H^{(0)}(f)$ can be profitably reduced to a generalized 1-D filter (i.e., a filter with non-null coefficients only along some line in R^2 [14]). In fact, we can set $\overline{H}^{(0)}(f) = \overline{H}_2(f_2)$ where $\overline{H}_2(f)$ is a 1-D filter characterized by passband and stopband frequencies $f_p = u_2, f_s = a_2$ (Fig. 11(a)), or $\overline{H}^{(0)}(f) = \overline{H}_1(f_1)$ with $\overline{H}_1(f)$ characterized by $f_p = u_1, f_s = a_1$ (Fig. 11(d)). Note in passing that other generalized 1-D filters, oriented along different directions, can be found, satisfying (17). The determination of such filters, however, is beyond the scope of the present paper.

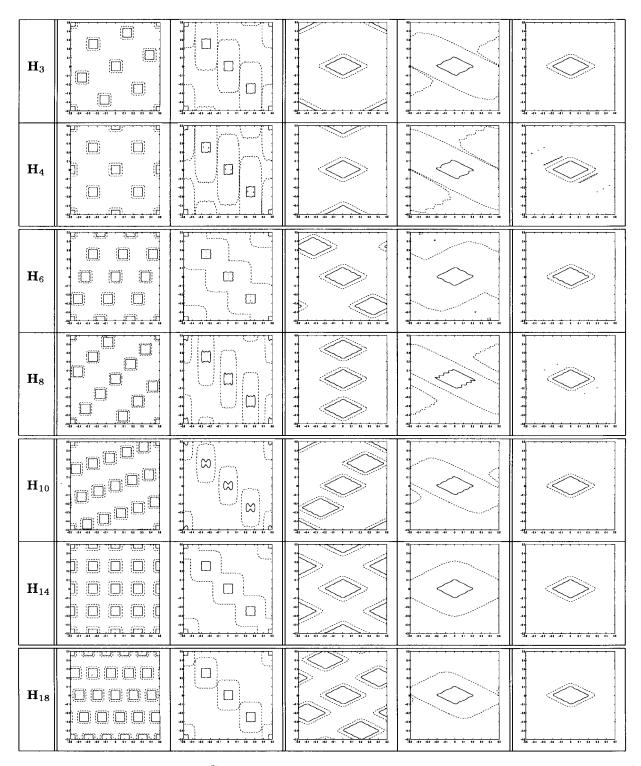


Fig. 8. Contour plot (within an elementary cell of Z^2) of the filters considered in the experimental test. Each row is related to a sublattice $\Lambda_i = \text{LAT}(\boldsymbol{C}\boldsymbol{H}_i)$. Second and fourth column: filters $\overline{H}^{(1)}(\boldsymbol{f})$ and $H^{(1)}(\boldsymbol{f})$. Third and fifth column: filters $\overline{H}^{(0)}(\boldsymbol{f})$ and $H^{(0)}(\boldsymbol{f})$. Sixth column: overall frequency response of the IFIR structure. Solid line: passband curve; dashed line: stopband curve.

IV. EXPERIMENTAL RESULTS

In this section we show an example of use of our algorithm for the design of 2-D IFIR filters. For given filter specifics (in terms of the spectral mask and of the definition lattice $\Lambda = \mathrm{LAT}(\mathbf{C})$), we first design the "direct" GF filter $H(\mathbf{f})$ on Λ for comparison. Then we consider a number of sublattices

 $\{\Lambda_i = \text{LAT}(\boldsymbol{CH}_i)\}\$ of Λ , and for each of them we design the related shaping filter and interpolators.

Both the generalized factorizable and the nonfactorizable implementation are considered. This way, we provide a reasonable guess of the performance attainable by IFIR schemes using filters other than GF.

TABLE II

Parameters of the IFIR Structures Considered in the Experimental Test. $P^u = \text{Upper Hermite Normal Form Matrix Associated to } AH_i; D = \text{Basis of LDFL}(\text{LAT}(AH_i)); N_1^{(0)}, N_2^{(0)} = \text{Lengths of the 1-D Filters Used in the Design of } h^{(0)}(a); \text{MNF, SNF} = \text{Overall Number of Multiplications and Sums Per Input Sample Required for the Non-Factorizable Implementation; MGF, SGF} = \text{Overall Number of Multiplications and Sums Per Input Sample Required for the Generalized Factorizable Implementation}$

	\mathbf{P}^u	D	$N_1^{(0)}$	$N_2^{(0)}$	MNF	SNF	MGF	SGF	δ_p	δ_s
\mathbf{H}_3	$ \left(\begin{array}{cc} 8 & 5 \\ 0 & 1 \end{array}\right) $	$ \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right) $	21	7	250	499	129	137	0.07	0.04
\mathbf{H}_{4}	$\left(\begin{array}{cc} 4 & 2 \\ 0 & 2 \end{array}\right)$	$\left[\begin{array}{cc} 2 & 0 \\ 0 & 2 \end{array}\right)$	21	5	129	489	66	79	0.06	0.02
\mathbf{H}_{6}	$\left(\begin{array}{cc} 12 & 9 \\ 0 & 1 \end{array}\right)$	$\left(\begin{array}{cc} 3 & 0 \\ 0 & 1 \end{array}\right)$	9	15	171	341	83	86	0.08	0.07
\mathbf{H}_8	$\left[\begin{array}{cc} 12 & 5 \\ 0 & 1 \end{array}\right]$	$ \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right) $	39	7	189	376	147	150	0.09	0.02
H ₁₀	$ \left(\begin{array}{cc} 16 & 5 \\ 0 & 1 \end{array}\right) $	$\left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right)$	25	9	144	286	142	134	0.13	0.06
H ₁₄	$\left(\begin{array}{cc} 4 & 0 \\ 0 & 4 \end{array}\right)$	$ \left(\begin{array}{cc} 4 & 0 \\ 0 & 4 \end{array}\right) $	15	15	86	256	47	55	0.07	0.02
H ₁₈	$\left \begin{array}{cc} 20 & 5 \\ 0 & 1 \end{array}\right)$	$\left(\begin{array}{cc} 5 & 0 \\ 0 & 1 \end{array}\right)$	21	15	132	262	98	99	0.11	0.05

In our example, we have tested exhaustively the IFIR systems relative to all the sublattices Λ_i with index in Λ ranging from 2 to 5, corresponding to bases CH_i with H_i in upper Hermite normal form (enumerated in Table I). For each sublattice Λ_i , the "optimal" stopband curves of the interpolator (corresponding to the maximal rectangles) are found. It is seen that some times more than one "maximal" stopband curve exists, while in other cases no one can be found (i.e., there is no elementary cell of Λ_1^* which contains $Par(\alpha v_1 | \alpha v_2)$). These latter sublattices cannot be used in a IFIR structure. An algorithm to check for such an occurrence may be easily devised: one just needs to determine whether $\mathcal{R}(\alpha u)$ is contained within any of the two elementary cells of Γ_1^* centered at the origin and with sides pairwise parallel to the axes [13].

In some instances (like in the case of Fig. 6), there can be couples of symmetric stopband curves, and we will consider just one interpolator for each couple.

For our experiments, we have chosen minimax 1-D filters with equal stopband and passband ripples. In order to provide some homogeneity in the results, the order of the 1-D filters in the design of the interpolator $H^{(0)}(\mathbf{f})$ have been chosen so as to obtain the same ripples exhibited by the 1-D filters in the design of $H^{(1)}(\mathbf{f})$.

Finally, note that all the figures of this section represent the actual passband and stopband curves (as defined in Appendix A) of the filters within the square $\mathcal{R}(0.5, 0.5)$

A. The Test

We consider here the definition lattice $\Lambda = Z^2$ and the diamond-shaped spectral mask characterized by: passband curve = $Par(\mathbf{v}_1|\mathbf{v}_2)$, with $\mathbf{v}_1 = (1/10, 1/20)^T, \mathbf{v}_2 = 1/10$

 $(-1/10, 1/20)^T$, stopband curve = $Par(\alpha v_1 | \alpha v_2)$ with $\alpha = 3/2$. Note that this spectral mask is the same one described in Section II, enlarged by a factor 2 (see Fig. 7). Hence, the decimation lattice is again

$$A = \begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}.$$

The 1-D kernels in the design of the "direct" filter $h(\mathbf{a})$ are both of length N=61; the passband and stopband ripples are $\delta_p=\delta_s=0.05$. Using the nonfactorizable implementation, 465 multiplications and 929 sums per input sample are required to realize $h(\mathbf{a})$ (in fact, the size of $h(\mathbf{a})$ is $N^2/|\det(\mathbf{A})|=930$). Using the generalized factorizable implementation, such values are reduced to 92 and to 117, respectively. Note that the condition for the quadrantal-like symmetry discussed in Section II, point 5, is not satisfied here.

The frequency responses of the considered filters are shown in Fig. 8 (all the relevant cases are present). The quantitative results are summarized in Table II.

In order to provide a better understanding of the results, the upper Hermite normal form matrix P^u associated to each AH_i has been computed and reported in Table II. Whenever either $P_{1,1}^u = 2P_{1,2}^u$, or P^u is diagonal, the condition for the quadrantal-like symmetry of the coefficients of $h^{(1)}(\mathbf{a})$ is satisfied (see Section II); this fact can be exploited for the reduction of the number of multiplications in the non-factorizable implementation. Such a profitable contingency is verified in the cases of H_4 and H_{14} . In particular, with respect to the "direct" filter $h(\mathbf{a})$, in the case of H_{14} , the number of multiplications and of sums per input sample using the nonfactorizable implementation are reduced approximately by a factor 5 and 3, respectively, by using the IFIR structure, with comparable frequency response ripples (see Fig. 9).

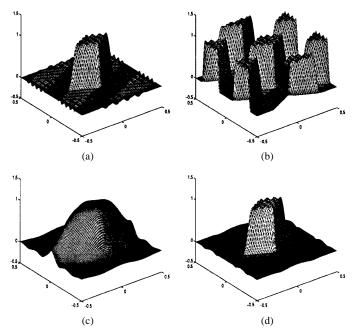


Fig. 9. Frequency responses (within an elementary cell of \mathbb{Z}^2) of the filters corresponding to the sixth row of Fig. 8. (a) $H(\mathbf{f})$. (b) $H^{(1)}(\mathbf{f})$. (c) $H^{(0)}(\mathbf{f})$. (d) $H^{(1)}(\mathbf{f})H^{(0)}(\mathbf{f})$.

In Table II we have also reported matrix D, the diagonal basis of LDFL(LAT(AH_i)). As described in Section II, decimating the impulse response $h(\mathbf{a})$ on LAT(\mathbf{CH}_i) yields lower computational weight (using the generalized factorizable implementation) only if LDFL(LAT(AH_i)) is less dense that LDFL(LAT(A)) (in this case Z^2). In fact, in the cases of H_3, H_8 and H_{10} such a condition is not verified, and the number of multiplications and of sums per input sample using the generalized factorizable implementation is higher than in the case of the direct implementation. Hence, in such cases, the IFIR scheme is not profitable. On the other side, when $|\det(\mathbf{D})| > 1$, the IFIR scheme makes for the reduction of the computational weight also when the generalized factorizable implementation is used.

It can be noticed from Table II that, corresponding to $|\det(\mathbf{H})| = 4$, we have a "saturation" phenomenon (similar to the 1-D case): increasing the index of the sublattice leads to increase the number of OPS's. Beyond such a limit, the spectral repetitions of the shaping filter are too dense to be separated by a "simple" interpolator, and the IFIR scheme is not efficient.

V. CONCLUSIONS

In this paper we have proposed an extension to the 2-D case of the idea of Interpolated FIR filters. We have considered only spectral supports in the shape of parallelograms, for which Generalized Factorizable filters stand as a profitable choice. The existence of approximate relationships among the filter parameters for GF filters enabled us to approach the problem under a geometric framework (i.e., dealing only with the filters' passband and stopband regions).

Several issues that do not have a counterpart in the 1-D case, such as the choice of the sublattice of definition of the shaping filter, and of the shape of the interpolator's spectral support, have been dealt with.

The experiments have been performed considering both the generalized factorizable and the nonfactorizable implementation of the filters. The results show that with nonfactorizable filters, good gains (in terms of computational weight reduction) are achievable. We maintain that, in such a case, the reduction of the number of OPS's corresponds roughly to the index of the sublattice of definition of the shaping filter (so long as such index stays below a saturation level).

Using the generalized factorizable implementation, the situation is more complex. Depending on the spectral support shape, IFIR schemes may lead to the reduction of the computational burden in certain cases. The computational gain, however, is typically lower than in the case of the nonfactorizable implementation.

Future work will consider the use of multi-band interpolators, and the joint design of the the couple shaping filter-intepolator.

APPENDIX A LATTICE THEORY BASICS

In this Appendix we report some notions of lattice theory that are used extensively throughout the paper. For the proofs, as well as for more details, the reader is addressed to [17], [9], [5], and [14].

We deal always with square full-rank matrices in this paper. For the purpose of this Appendix, we assume that the size of the considered matrices is fixed to M.

Any integral matrix U such that U^{-1} is still integral (or, equivalently, such that $|\det(U)| = 1$ is called *unimodular*. Two integral matrices A_1, A_2 such that $A_2^{-1}A_1$ is unimodular are called right-equivalent or associated. For each class of associates, there is just one upper Hermite normal form matrix, i.e., a matrix A^u such that

- 1) A^u is upper triangular;
- 2) $A_{i,j}^u \ge 0$; 3) $A_{i,j}^u < A_{i,i}^u$ for $1 \le i < j \le M$; 4) $A_{i,j}^u = 0$ if $A_{i,i}^u = 0$.

A lattice Λ that admits a basis \boldsymbol{A} is denoted by LAT(\boldsymbol{A}). In other words, $\Lambda = LAT(\mathbf{A}) = \{\mathbf{An} : \mathbf{n} \in \mathbb{Z}^M\}$. Matrices \mathbf{A}_1 and A_2 are bases of the same lattice if (and only if) $A_2^{-1}A_1$ is unimodular. When dealing with sampling lattices, we always assume that they are sublattices of Z^M (i.e., they are integral lattices, so that they admit only integral bases). A lattice is said to be factorizable (or separable) if it admits a diagonal basis.

An elementary cell of a lattice Λ is any region \mathcal{C} such that

- 1) $C + \mathbf{a}_i \cap C + \mathbf{a}_j = \emptyset$ for any $\mathbf{a}_i, \mathbf{a}_j \in \Lambda, \mathbf{a}_i \neq \mathbf{a}_j$;
- 2) $\bigcup_{\boldsymbol{a}_i \in \Lambda} \mathcal{C} + \boldsymbol{a}_i = R^M$.

A lattice $\Gamma = LAT(B)$ is a sublattice of $\Lambda = LAT(A)$ if (and only if) $\mathbf{H} = \mathbf{A}^{-1}\mathbf{B}$ is integral. Term $|\det(\mathbf{H})|$ is called the *index* of Γ in Λ (indicated as $(\Lambda:\Gamma)$), and represents the ratio between the density of points of Λ and of Γ .

Let n be an integer. Then the distinct sublattices having index n in LAT(A) are $\{LAT(AH_i)\}$, where $\{H_i\}$ are the upper Hermite normal forms matrices with determinant equal to n.

An important notion is that of the least dense factorizable lattice (LDFL) Ψ containing $\Lambda = LAT(C)$, The LDFL of a lattice Λ is defined as the smallest factorizable superlattice of Λ . Let $\mathbf{R} = \operatorname{diag}(R_1, R_2, \dots, R_M)$ be a basis of Ψ . The entry R_i is the greatest common divisor of the entries of the ith row of matrix A. We also sometimes consider the densest factorizable sublattice (DFS) of Λ [14].

We adopt the following definition for the Fourier transform of a signal $h(\mathbf{a})$ defined on a lattice $\Lambda = LAT(\mathbf{A})$:

$$H(\mathbf{f}) = \sum_{\mathbf{a} \in \Lambda} h(\mathbf{a}) e^{-j2\pi \mathbf{f}^T \mathbf{a}} = \sum_{\mathbf{n} \in Z^M} h(\mathbf{A}\mathbf{n}) e^{-j2\pi \mathbf{f}^T \mathbf{A}\mathbf{n}}.$$
(19)

 $H(\mathbf{f})$ is periodic on the *dual* lattice $\Lambda^* = \text{LAT}(\mathbf{A}^{-T})$, where $\mathbf{A}^{-T} \stackrel{\triangle}{=} (\mathbf{A}^{-1})^T$.

The Fourier transform of the signal $h_s(\mathbf{a})$ defined on the sublattice Λ_1 of Λ , obtained from the signal h(a) defined on Λ as by $h_s(\mathbf{a}) \stackrel{\Delta}{=} h(\mathbf{a}), \mathbf{a} \in \Lambda_1$, is

$$H_s(\mathbf{f}) = \frac{1}{\Lambda : \Lambda_1} \sum_{\mathbf{r} \in \Lambda_1^* \cap \mathcal{V}} H(\mathbf{f} + \mathbf{r})$$
 (20)

where \mathcal{V} is any elementary cell of Λ^* .

Consider a 2-D filter $H(\mathbf{f})$, defined on \mathbb{Z}^2 . We define transition region of H(f) the region of the elementary cell $\mathcal{R}(0.5, 0.5)$ delimited by the passband curve $\underline{\mathcal{P}}$ and the stopband curve \underline{S} (when they are univocally determined), defined as follows:

$$\underline{\mathcal{P}} = \{ \mathbf{f} : H(\mathbf{f}) = 1 - \delta_p, \|\nabla H(\mathbf{f})\|^2 \neq 0 \}$$
 (21)

$$\underline{S} = \{ \mathbf{f} : H(\mathbf{f}) = \delta_s, ||\nabla H(\mathbf{f})||^2 \neq 0 \}$$
 (22)

where ∇ indicates the gradient operator and

$$\delta_p = \max_{\{\boldsymbol{f}_i: ||\nabla H(\boldsymbol{f}_i)||^2 = 0, |H(\boldsymbol{f}_i)| > 0.5\}} |1 - H(\boldsymbol{f}_i)| \quad (23)$$

$$\delta_{p} = \max_{\{\boldsymbol{f}_{i}: ||\nabla H(\boldsymbol{f}_{i})||^{2} = 0, |H(\boldsymbol{f}_{i})| > 0.5\}} |1 - H(\boldsymbol{f}_{i})| \qquad (23)$$

$$\delta_{s} = \max_{\{\boldsymbol{f}_{i}: ||\nabla H(\boldsymbol{f}_{i})||^{2} = 0, |H(\boldsymbol{f}_{i})| < 0.5\}} |H(\boldsymbol{f}_{i})|. \qquad (24)$$

The region \mathcal{P} contained within $\underline{\mathcal{P}}$ is called the *passband region* of $H(\mathbf{f})$, while the region of $\mathcal{R}(0.5, 0.5)$ outside $\underline{\mathcal{S}}$ is called the stopband region of $H(\mathbf{f})$. Our definitions may be easily extended to the case of filters defined on a nonorthogonal lattice. Instead of region $\mathcal{R}(0.5, 0.5)$, some other suitable elementary cell of the frequency repetition lattice centered at the origin may be chosen.

APPENDIX B MAXIMAL RECTANGLES DETERMINATION

In this Appendix we describe an algorithm to find the maximal rectangles (defined in Section III-B2), given $\Gamma^*, \Gamma_1^*, \boldsymbol{u}$ and

Let $\mathcal{B} = \mathcal{R}(p_1, p_2)$, where LAT(diag (p_1, p_2)) is the densest factorizable sublattice (DFS) of Γ^* . One can easily show that the rectangular elementary cells of Γ_1^* centered at the origin with sides pairwise parallel to the axes are contained in \mathcal{B} .

A maximal rectangle $\mathcal{R}(a_1, a_2)$ such that some side of its is contained within a side of \mathcal{B} (i.e., such that $a_1 = p_1$ (case 1) or $a_2 = p_2$ (case 2)) will be termed *extremal*. The periodical

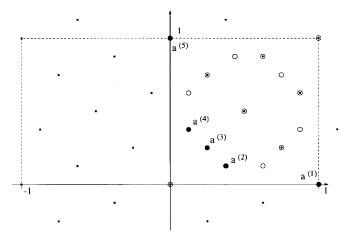


Fig. 10. Lattice Γ_1^* (dots), rectangle \mathcal{B}' (dashed line), set T_1 (circles) and set T_2 (filled circles) for the construction of the maximal rectangles.

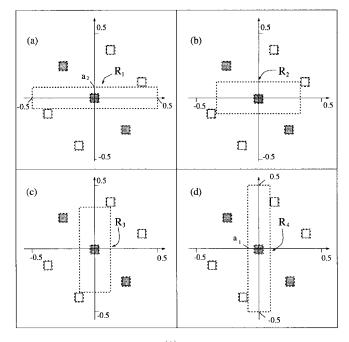


Fig. 11. Spectral repetitions of $\overline{H}^{(1)}(\mathbf{f})$ on the points of Γ_1^* and the four maximal rectangles. Solid line: passband curve; dashed line: stopband curve; filled diamonds: spectral repetitions on the points of Γ^* .

repetitions of extremal maximal rectangles on the points of Γ^* form horizontal (case 1) or vertical (case 2) stripes.

The first step in the algorithm to find the maximal rectangles, is constructing an ordered set T_1 from the points of Γ_1^* contained within the rectangle³ $\mathcal{B}' = \{ \mathbf{f}, \mathbf{f} \in \mathcal{B}, f_2 \ge 0 \}$:

$$T_1 = \{ (|a_1^{(i)}|, a_2^{(i)}) : (a_1^{(i)}, a_2^{(i)}) \in (\Gamma_1^* \cap \mathcal{B}') \setminus \{\mathbf{0}\} \}. \tag{25}$$

Next, we construct T_2 from T_1 by discarding all points $(a_1^{(i)},a_2^{(i)})$ such that some other point $(a_1^{(j)},a_2^{(j)})$ exists in T_1 with $a_1^{(j)} \leq a_1^{(i)}$ and $a_2^{(j)} \leq a_2^{(i)}$. Note that our definition is consistent (i.e. it gives rise to just one set T_2 starting from T_1), and that one can order the elements of T_2 according to the ascending order of their component $a_2^{(i)}$: $T_2 = \{a^{(1)},$ $a^{(2)}, \cdots, a^{(M)}$ }.

³A dual algorithm would interchange the role of f_1 and f_2 .

Finally, the set of maximal rectangles is composed by the M-1 rectangles $\mathcal{R}(\boldsymbol{b}^{(1)}), \mathcal{R}(\boldsymbol{b}^{(2)}), \dots, \mathcal{R}(\boldsymbol{b}^{(M-1)})$, with

$$b_1^{(i)} = \begin{cases} a_1^{(i)}/2, & \boldsymbol{a}^{(i)} \in \Gamma^* \\ a_1^{(i)} - \alpha u_1, & \boldsymbol{a}^{(i)} \notin \Gamma^* \end{cases}$$
(26)

and

$$b_2^{(i)} = \begin{cases} a_2^{(i+1)}/2, & \boldsymbol{a}^{(i+1)} \in \Gamma^* \\ a_2^{(i+1)} - \alpha u_2, & \boldsymbol{a}^{(i+1)} \notin \Gamma^* \end{cases}$$
(27)

A simple example should make our procedure clear. Consider the design parameters of Section III-B3: $\Gamma^* = \text{LAT}(\mathbf{A}^T), \Gamma_1^* = \text{LAT}((\mathbf{A}\mathbf{H})^T)$ with

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix},$$
$$\mathbf{u} = (1/40, 1/40), \quad \alpha = 3/2.$$

One can easily show [14] that DFS(Γ^*) = Z^2 , so that $\mathcal{B} = \mathcal{R}(1,1)$ (see Fig. 10). The points of T_1 are denoted by small circles in Fig. 10. Note that one can write an automatic procedure to obtain the points of Γ_1^* within \mathcal{B} . For example, one can determine LDFL(Γ_1^*), determine (trivially) the points of such lattice within \mathcal{B} , and identify those actually belonging to Γ_1^* . The ordered set T_2 , denoted by filled circles in Fig. 10, is

$$T_2 = \{(1,0), (\frac{3}{8}, \frac{1}{8}), (\frac{1}{4}, \frac{1}{4}), (\frac{1}{8}, \frac{3}{8}), (0,1)\}.$$
 (28)

From T_2 we build the set of maximal rectangles

$$\begin{aligned}
&\{R_1 = \mathcal{R}(\frac{1}{2}, \frac{1}{8} - \frac{3}{80}), R_2 = \mathcal{R}(\frac{3}{8} - \frac{3}{80}, \frac{1}{8}) \\
&R_3 = \mathcal{R}(\frac{1}{8}, \frac{3}{8} - \frac{3}{80}), R_4 = \mathcal{R}(\frac{1}{8} - \frac{3}{80}, \frac{1}{2})\}.
\end{aligned} (29)$$

The maximal rectangles are depicted in Fig. 11. Note that rectangles R_1 and R_4 are extremal.

ACKNOWLEDGMENT

The author would like to acknowledge S. K. Mitra and T. Chen for their useful comments to an earlier draft of this paper. The author also wishes to thank the anonymous referees for their patient reading and valuable comments.

REFERENCES

- R. Ansari and S. Lee. "Two-dimensional nonrectangular interpolation, decimation, and filter banks," in *Proc. IEEE ICASSP 88*, New York, Apr. 1988.
- [2] T. Chen and P. P. Vaidyanathan, "Multidimensional multirate filters derived from one-dimensional filters," *Electron. Lett.*, vol. 27, no. 3, pp. 225–228, Jan. 1991.

- [3] ______, "Multidimensional multirate filters and filter banks derived from one-dimensional filters," *IEEE Trans. Signal Processing*, vol. 41, pp. 1740–1765, May 1993
- [4] 1749–1765, May 1993.
 [4] _____, "Recent developments in multidimensional multirate systems," IEEE Trans. Circuits Syst. Video Technol., vol. 3, pp. 116–137, Apr. 1993
- [5] G. M. Cortelazzo and R. Manduchi, "On the determination of all the sublattices of preassigned index and its application to multidimensional subsampling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, pp. 318–320, Aug. 1993.
- [6] P. E. Crochiere and L. Rabiner, "Optimum FIR digital filter implementation for decimation, interpolation, and narrow-band filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP–23, pp. 444–456, May 1975.
- [7] P. E. Crochiere and L. Rabiner, "Further considerations in the design of decimators and interpolators," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP–24, pp. 296–311, Aug. 1976.
- [8] _______, "Interpolation and decimation of digital signals—A tutorial review," *Proc. IEEE*, vol. 69, pp. 300–331, Mar. 1981.
- [9] E. Dubois, "The sampling and reconstruction of time-varying imagery with applications in video systems," *Proc. IEEE*, vol. 73, pp. 502–522, Apr. 1985
- [10] D. E. Dudgeon and R. M. Mersereau, Multidimensional Digital Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [11] O. Herrmann, L. R. Rabiner, and D. S. K. Chan, "Practical design rules for optimum finite impulse response low-pass digital filters," *Bell Syst. Tech. J.*, vol. 52, no. 6, pp. 769–799, July 1973.
- [12] J. F. Kaiser, "Nonrecursive digital filter design using IO-sinh window function," in *Proc. IEEE ISCAS'74*, Apr. 1974, pp. 20–23.
- [13] R. Manduchi, "Two-dimensional interpolated IFIR filters using generalized factorable filters," Tech. Rep. UCB/CSD/93/785, Univ. California, Berkeley, 1993.
- Berkeley, 1993.
 [14] ______, "Some properties of generalized factorable 2-D FIR filters,"

 IEEE Trans. Signal Processing, vol. 44, pp. 1534–1546, June 1996.
- [15] R. Manduchi, G. M. Cortelazzo, and G. A. Mian, "Multistage sampling structure conversion of video signals," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, pp. 325–340, Oct. 1993.
- [16] Y. Neuvo, D. Cheng-Yu, and S. K. Mitra. Interpolated finite impulse response filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 563–570, June 1984.
- 17] M. Newman. Integral Matrices. New York: Academic, 1972.
- 18] L. R. Rabiner, "Approximate design relationships for low-pass FIR digital filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 456–460, Oct. 1973.
- [19] T. Saramäki, Y. Neuvo, and S. K. Mitra, "Design of computationally efficient interpolated FIR filters," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 70–87, Jan. 1988.
- [20] P. Siohan, "2-D FIR filter design for sampling structure conversion," IEEE Trans. Circuits Syst. Video Technol., vol. 1, pp. 337–350, Dec. 1991.



Roberto Manduchi received the Ph.D. degree in electrical engineering from the University of Padova, Italy.

He worked as a post-Doctoral Fellow at the University of California, Berkeley, at the California Institute of Technology, Pasadena, and at Stanford University, Stanford, CA, conducting research on computer vision and signal processing. He is currently with Apple Computer, Inc., Cupertino,