

Two-Dimensional Orthogonal Wavelets with Vanishing Moments

David Stanhill and Yehoshua Y. Zeevi

Abstract— We investigate a very general subset of 2-D, orthogonal, compactly supported wavelets. This subset includes all the wavelets with a corresponding wavelet (polyphase) matrix that can be factored as a product of factors of degree-1 in one variable. In this paper, we consider, in particular, wavelets with vanishing moments. The number of vanishing moments that can be achieved increases with the increase in the McMillan degrees of the wavelet matrix. We design wavelets with the maximal number of vanishing moments for given McMillan degrees by solving a set of nonlinear constraints on the free parameters defining the wavelet matrix and discuss their relation to regular, smooth wavelets. Design examples are given for two fundamental sampling schemes: the quincunx and the four-band separable sampling. The relation of the wavelets to the well-known 1-D Daubechies wavelets with vanishing moments is discussed.

I. INTRODUCTION

THE M -band 2-D discrete wavelet transform (DWT) is defined by M functions, one scaling function, $\psi^0(\mathbf{x})$, $(M - 1)$ wavelet functions $\psi^p(\mathbf{x})$, and a sampling matrix D . For wavelets associated with multiresolution analysis, the DWT can equally be specified by M wavelet filters $H^p(\omega)$ instead of the M wavelet functions. In this work, we restrict the discussion to compactly supported functions that define an orthogonal DWT. The orthogonal compactly supported wavelet filters are a special case of lossless FIR multirate filter banks. Since a considerable number of studies have been devoted to lossless filter banks, we adopt, in part, the notation and terminology used in this field. Specifically, we limit most of the analysis to the polyphase matrix defined by the wavelet filters (wavelet matrix) rather than to the wavelet functions.

Most current applications of the 2-D DWT use 1-D wavelet functions and filters to construct 2-D separable wavelets [1]. Using separable functions preserves orthogonality but imposes severe limitations on the 2-D wavelets. For example, the commonly used four-band filters cannot have linear phase, except for the Haar wavelet, and their orientation cannot be adjusted for given specifications. Various other methods were proposed for the design of 2-D perfect reconstruction filter banks [2]–[4], which do not require orthogonality (paraunitarity) or finite support. The restrictions imposed on the functions/filters still leave considerable freedom in choosing the wavelet functions/filters.

Manuscript received August 7, 1995; revised April 17, 1996. This work was supported, in part, by the Ollendorff Center of the Department of Electrical Engineering, by the Israel Science Foundation, and by the Fund for Promotion of Research at the Technion. The associate editor coordinating the review of this paper and approving it for publication was Dr. Michael Zervakis.

The authors are with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel.

Publisher Item Identifier S 1053-587X(96)07120-6.

Motivated by the work of Daubechies in one dimension [5], we examine the design of 2-D, orthogonal, compactly supported wavelets with maximal number of vanishing moments. In the 1-D case, these have become the standard wavelets, and they can be considered suboptimal in a large class of applications, such as the representation, processing, and compression of 1-D signals. For detailed experimental results in image compression using separable wavelets showing the optimality of wavelets with vanishing moments, see [6]. The generalization of the 1-D notion of vanishing moments to higher dimensions is straightforward, but its nature and importance has yet to be determined. In 1-D two-band wavelets, the number of vanishing moments is closely related to the flatness of the frequency response of the wavelet filters at the aliasing points. When dealing with more than two bands, even in 1-D, the M functions are no longer uniquely defined by each other, and the vanishing moment constraints ensure flatness of the scaling filter but not of the other wavelet filters. The regularity of the wavelet functions is also closely related to the number of vanishing moments in the 1-D case [7]–[9]. Yet in more than 1-D, no such relation has been rigorously shown, although it was suggested to hold by Daubechies and Lagarias [8]. The work of Villemoes in 2-D (quincunx) [10] introduces a different approach to the general 2-D regularity measurement and is used here to obtain lower bounds of the Hölder exponent.

The factorization method used by Daubechies [5] in 1-D to design orthogonal wavelets with N vanishing moments has no straightforward generalization to more than 1-D since the fundamental theorem of algebra holds only for 1-D. We therefore construct 2-D wavelets by first parameterizing the space of orthogonal wavelets with compact support and then impose a set of constraints on the parameters in order to achieve the desired number of vanishing moments. The shortcoming of this method is that no complete parameterization of the space of 2-D, orthogonal, compactly supported wavelets exists. In most of this work, we limit the design to wavelets with wavelet matrices that are factorable as a product of 1-D degree-1 factors. Such wavelets were proposed by Karlsson and Vetterli [11]. The issue of parameterization of 2-D lossless FIR transfer matrices has been investigated by Venkataraman and Levy [12], [13]. They have shown that the set of lossless FIR transfer matrices (wavelet matrices in our case) that are factorable is not of measure zero in the set of general lossless FIR matrices. Toward the completion of this manuscript, our attention was drawn to a recent work by Kovačević and Vetterli [14]. They use a construction method similar to the one proposed here, and give a few design examples.

The rest of the paper is organized as follows. The first section is devoted to the description of 2-D orthogonal wavelets and filter banks in terms of their polyphase matrix, and the notion of *factorable* wavelets is introduced. In the next section, we give the conditions for obtaining wavelets with L vanishing moments and discuss their relation to regular, smooth wavelets. Specific design results and examples are given in the last section: first for quincunx wavelets and then for four-band, separable sampling wavelets.

A few comments on notation. The term *order* refers to the polynomial order, while degree is used to denote the McMillan degree. We use boldface for 2-D entities like \mathbf{x} , \mathbf{n} . We adopt the notation used in the multidimensional multirate literature; for two vectors $\mathbf{n} = (n_1, n_2)^T$ and $\mathbf{z} = (z_1, z_2)^T$, we have $\mathbf{z}^{\mathbf{n}} \triangleq z_1^{n_1} z_2^{n_2}$, and for a 2×2 matrix D , we define $\mathbf{z}^D \triangleq (z_1^{D_{1,1}} z_2^{D_{2,1}}, z_1^{D_{1,2}} z_2^{D_{2,2}})$. Capital letters are used for functions given in the z and frequency domains, such as $H(\mathbf{z})$ and $H(\boldsymbol{\omega})$. Finally, boldface is used also for matrix-valued functions.

II. 2-D WAVELETS AND MULTIRATE FILTER BANKS

When dealing with multidimensional wavelets and multirate systems, the change in resolution and sampling rate is given by a matrix D . In the 2-D case, D is a 2×2 matrix with both eigenvalues larger than 1 in absolute value, as compared with a single factor larger than 1 in the 1-D case. For a detailed analysis of 2-D subsampling and multirate systems, see Viscito and Allebach [15], and for a more mathematical background in multidimensional wavelets, see Lawton and Resnikoff [16]. Here, we quote a few facts and definitions relevant to the rest of this paper. For a discrete signal given on a lattice Λ , the subsampled signal is defined on the lattice $\Lambda_D = \{D\mathbf{n}: \mathbf{n} \in \Lambda\}$. The density of grid points on Λ_D is smaller by a factor of $M = |\det(D)|$ than the density on Λ . A *coset* of a sublattice is the set of grid points on the lattice that are also on a shifted version of the sublattice. A coset is referred to by a vector \mathbf{k}_i that defines the 2-D shift. There are M distinct cosets that cover the entire lattice. The points $\boldsymbol{\omega}_i = 2\pi D^{-T} \mathbf{k}_i$, where \mathbf{k}_i are coset vectors of D^T , are called points of repeated spectra or "aliasing points." We now define 2-D wavelets.

The discrete 2-D wavelet transform (DWT) of a function $f \in L_2(\mathbb{R}^2)$ is defined as

$$(\mathcal{W}f)_{j,\mathbf{n}}^p \triangleq \int f(\mathbf{x}) \psi_{j,\mathbf{n}}^p(\mathbf{x}) d\mathbf{x} \quad (1)$$

where

$$\psi_{j,\mathbf{n}}^p(\mathbf{x}) \triangleq \sqrt{M^j} \psi^p(D^j \mathbf{x} - \mathbf{n}),$$

$$p = 1, \dots, M-1; \quad j \in \mathbb{Z}; \quad \mathbf{n} \in \mathbb{Z}^2.$$

The $(M-1)$ wavelet functions $\psi^p(x)$ are related to the scaling function $\psi^0(x)$ through the two-scale refinement equations

$$\psi^p(\mathbf{x}) = \sqrt{M} \sum_{\mathbf{n}} h^p(\mathbf{n}) \psi^0(D\mathbf{x} - \mathbf{n}) \quad (2)$$

where $h^p(\mathbf{n})$ is a 2-D sequence. In this work, we focus our effort on real, compactly supported wavelets; hence, the

sequences $h^p(\mathbf{n})$ are real and of finite length. The M sequences $h^p(\mathbf{n})$ and their Fourier transform $H^p(\boldsymbol{\omega})$ can be viewed as the impulse and frequency responses of 2-D FIR filters and are called *wavelet filters*. The scaling filter $H^0(\boldsymbol{\omega})$ is a lowpass filter, and the rest are highpass or bandpass filters.

To assist the analysis of multirate filter banks, we define two matrices: the *polyphase* and *modulation* matrices. In multirate systems, each filter $H^p(\boldsymbol{\omega})$ is decomposed into M polyphase components: one for each coset shift vector \mathbf{k}_j

$$H^p(\boldsymbol{\omega}) = \sum_{j=0}^{M-1} e^{-i\boldsymbol{\omega} \cdot \mathbf{k}_j} H_j^p(D^T \boldsymbol{\omega}) \quad (3)$$

where

$$H_j^p(\boldsymbol{\omega}) \triangleq \sum_{\mathbf{n}} h^p(D\mathbf{n} + \mathbf{k}_j) e^{-i\boldsymbol{\omega} \cdot \mathbf{n}}.$$

The polyphase matrix of a maximally decimated filter bank is a $M \times M$ matrix $\mathbf{H}(\boldsymbol{\omega})$ with one column for each filter and one row for each polyphase component, that is, the p th column is $(H_0^p(\boldsymbol{\omega}), H_1^p(\boldsymbol{\omega}), \dots, H_{M-1}^p(\boldsymbol{\omega}))^T$. Note that the polyphase matrix is sometimes defined as the transpose of the one defined above. In the modulation matrix $\mathbf{H}_M(\boldsymbol{\omega})$, the p th column is built from M shifted versions of the filter $H^p(\boldsymbol{\omega})$: one for each aliasing frequency, i.e., the p th column is

$$\frac{1}{\sqrt{M}} (H^p(\boldsymbol{\omega}), H^p(\boldsymbol{\omega} + 2\pi D^{-T} \mathbf{k}_1), \dots, H^p(\boldsymbol{\omega} + 2\pi D^{-T} \mathbf{k}_{M-1}))^T.$$

The following two propositions, which are well known in 1-D, were proven for the multidimensional case in [16].

Proposition 1: A maximally decimated filter bank is lossless if and only if it obeys one of the following equivalent conditions:

1)

$$\sum_{\mathbf{n}} h^p(\mathbf{n}) h^q(D\mathbf{m} + \mathbf{n}) = \delta_{p,q} \delta_{\mathbf{m},0}.$$

2)

$$\mathbf{H}^\dagger(\boldsymbol{\omega}) \mathbf{H}(\boldsymbol{\omega}) = I \quad \text{or} \quad \sum_{\mathbf{n}} h^\dagger(\mathbf{n} + \mathbf{k}) h(\mathbf{n}) = I \delta_{\mathbf{k},0}.$$

3)

$$\mathbf{H}_M^\dagger(\boldsymbol{\omega}) \mathbf{H}_M(\boldsymbol{\omega}) = I.$$

Proposition 2: A lossless bank of M , 2-D, FIR filters defines a compactly supported, *orthogonal* 2-D wavelet system if and only if it satisfies $\sum_{\mathbf{n}} h^p(\mathbf{n}) = H^p(\boldsymbol{\omega} = 0) = \delta_{p,0} \sqrt{M}$.

The term *orthogonal* applies only to the wavelet filters and not to the wavelet functions. In the z -domain, the polyphase matrix of FIR (causal) filters is a 2-D polynomial matrix in the variables z_1^{-1} and z_2^{-1} . The unitary condition of the polyphase matrix is

$$\mathbf{H}^T(z^{-1}) \mathbf{H}(z) = \mathbf{H}(z) \mathbf{H}^T(z^{-1}) = I. \quad (4)$$

Polyphase matrices that obey the above condition are called paraunitary, i.e., unitary at all points $|z_i| = 1$. We use the term *wavelet matrix* of rank M for the polyphase matrix of an

orthogonal M -band wavelet system. Since the characteristic matrix \mathbf{H}_0 of a paraunitary polynomial (PUP) matrix, which is defined as $\mathbf{H}_0 = \mathbf{H}(z_1 = 1, z_2 = 1)$, is a constant orthogonal matrix, any PUP matrix can always be decomposed as

$$\mathbf{H}(z_1, z_2) = \mathbf{H}_I(z_1, z_2)\mathbf{H}_0$$

where

$$\mathbf{H}_I(z_1 = 1, z_2 = 1) = \mathbf{I}.$$

For wavelet matrices, the characteristic matrix must have a special form to satisfy the constraints in Proposition 2. These matrices are called *Haar wavelet matrices* (HWM's) [17], and the notation $\mathcal{H}^{(M)}$ is used for a HWM of rank M . Most results will be stated for general PUP matrices, but the design examples will be limited to wavelet matrices. The reader unfamiliar with 2-D polynomial matrix notions such as the McMillan degrees should refer to [12]. We will just note that for PUP matrices of McMillan degrees (m, n) , one has $\text{Det}(\mathbf{H}(z_1, z_2)) = z_1^{-m}z_2^{-n}$.

Definition 1: A PUP matrix of McMillan degrees (m, n) in z_1 and z_2 is called *factorable* if it can be factored into a product of McMillan degree-1, 1-D matrices m of them in z_1 and n in z_2 , that is,

$$\mathbf{H}(z_1, z_2) = \prod_{k=1}^{m+n} \mathbf{H}_k(z_k) \quad (5)$$

where $z_k = z_1$ or z_2 and \mathbf{H}_k are 1-D PUP matrices of McMillan degree 1.

Using 1-D results [18], we can decompose any factorable PUP $M \times M$ matrix of McMillan degrees (m, n) as

$$\mathbf{H}(z_1, z_2) = \prod_{k=1}^{m+n} [I + (z_k^{-1} - 1)V_k V_k^T] \mathbf{H}_0 \quad (6)$$

where \mathbf{H}_0 is the characteristic matrix of \mathbf{H} , V_k are unit vectors of size M , and as above, $z_k = z_1$, or z_2 . As mentioned earlier, for wavelet matrices, \mathbf{H}_0 must satisfy an extra constraint of regularity and is denoted by $\mathcal{H}^{(M)}$. The number of free parameters in the above factorization (6) is

$$N_{fp} = \binom{M}{2} + (M-1)(m+n)$$

or, for wavelets,

$$N_{fp} = \binom{M-1}{2} + (M-1)(m+n)$$

where the binomial part comes from the characteristic (Haar wavelet) matrix and the rest from the $(m+n)$ unit vectors of length M . In [12], it is shown that this is equal to the number of degrees of freedom of all PUP matrices of rank M and McMillan degrees (m, n) ; hence, the set of factorable PUP matrices is not a subset of measure zero in the set of all PUP matrices. Furthermore, in the case that one of the polynomial degrees is equal to one, all the PUP matrices are factorable [19]. In [20], we show that separable filter-banks and wavelets are always factorable, and we show how to construct the degree-1 factors given the factorization of the two 1-D polyphase matrices.

III. VANISHING MOMENTS AND REGULARITY

One-dimensional wavelets with vanishing moments have been used extensively in the context of 1-D and 2-D applications of the DWT. In 1-D as well as in 2-D, the number of vanishing moments gives the local order of approximation. The requirement of maximal number of vanishing moments was found to be appropriate for a variety of applications and gave close to optimal results in most cases. This success in 1-D motivates us to study 2-D wavelets with vanishing moments and to apply them in image coding and representation as well as in other 2-D applications.

As in the 1-D case [21], we can characterize the quality of vanishing moments (or L regularity) in several analogous manners.

Theorem 1: The following three conditions are equivalent and can serve as a definition of an orthogonal M -band wavelet system with L vanishing moments.

- 1) All moments up to order $(L-1)$ of the wavelet filters vanish, that is

$$0 = \mu_{k,q}^p \triangleq \sum_{\mathbf{n}} h^p(\mathbf{n}) \mathbf{n}^q \\ = (-i)^{k+q} \left. \frac{\partial^{k+q} H^p(\omega_x, \omega_y)}{\partial \omega_x^k \partial \omega_y^q} \right|_{\omega=0} \quad (7)$$

for all $p = 1, \dots, (M-1)$ and $k+q = 0, 1, \dots, L-1$, where \mathbf{n}^q denotes $(m, n)^{(k,q)} = m^k n^q$.

- 2) All moments up to order $(L-1)$ of the wavelet functions vanish, that is,

$$0 = m_{k,q}^p \triangleq \int \mathbf{x}^q \psi^p(\mathbf{x}) d\mathbf{x} \\ = (-i)^{k+q} \left. \frac{\partial^{k+q} \psi^p(\omega_x, \omega_y)}{\partial \omega_x^k \partial \omega_y^q} \right|_{\omega=\omega_i} \quad (8)$$

for p, k, q as above, and again, \mathbf{x}^q denotes $x^k y^q$.

- 3) The frequency response of the scaling filter has a zero of order L at the $(M-1)$ aliasing frequencies, i.e.,

$$\left. \frac{\partial^{k+q} H^0(\omega_x, \omega_y)}{\partial \omega_x^k \partial \omega_y^q} \right|_{\omega=\omega_i} = 0 \quad (9)$$

for all $(M-1)$ aliasing frequencies ω_i , and $(k+q) = 0, 1, \dots, L-1$.

Proof: 1 \Rightarrow 2: using (2), we can derive the recursion equation

$$m_{k,q}^p = \sqrt{M} \sum_{\mathbf{n}} h^p(\mathbf{n}) \int \mathbf{x}^q \psi^0(D\mathbf{x} - \mathbf{n}) d\mathbf{x} \\ = \sqrt{M} \sum_{\mathbf{n}} h^p(\mathbf{n}) \int (D^{-1}(\xi + \mathbf{n}))^q \psi^0(\xi) \frac{d\xi}{M}. \quad (10)$$

Writing $(D^{-1}(\xi + \mathbf{n}))^q$ explicitly and using the binomial expansion, we arrive at

$$m_{k,q}^p = \frac{1}{\sqrt{M}} \sum_{i=0}^k \sum_{j=0}^q a^i b^{q-i} \binom{k}{i} \binom{q}{j} \\ \cdot \sum_{s=0}^{i+j} \sum_{t=0}^{k+q-i-j} \binom{i+j}{s} \binom{k+q-i-j}{t} \\ \cdot \mu_{s,t}^p m_{(i+j-s), (k+q-i-j-t)}^0 \quad (11)$$

where $D^{-1} = (\mathbf{a}, \mathbf{b})$, and $\mathbf{i} = (i, j)^T$. Hence, if $\mu_{k,q}^p = 0, \forall (k+q) < L$ on the right-hand side of (11), we get $m_{k,q}^p = 0$ on the left-hand side.

2 \Rightarrow 1: We use the recursion relation (11) and prove by way of induction.

- 1) $k = q = 0$: $\mu_{0,0}^p = 0$ must be true for all wavelets (Proposition 2).
- 2) $k + q = l$: Given $m_{k,q}^p = 0 \forall (k+q) \leq l$, we assume that $\mu_{k,q}^p = 0 \forall (k+q) < l$ and will prove that $\mu_{k,q}^p = 0 \forall (k+q) = l$. Using the above assumption, in (11), we are left only with the terms in which $s+t = l$, that is, for $k+q = l$

$$\begin{aligned} 0 &= m_{k,q}^p \\ &= \sum_{i=0}^k \sum_{j=0}^q \mathbf{a}^i \mathbf{b}^{q-i} \binom{k}{i} \binom{q}{j} \mu_{i+j, k+q-i-j}^p m_{0,0}^0 \\ &= \sum_{(i+j)=0}^l \left(\sum_{i-j} \mathbf{a}^i \mathbf{b}^{q-i} \binom{k}{i} \binom{q}{j} \right) \\ &\quad \cdot \mu_{i+j, l-(i+j)}^p m_{0,0}^0 \end{aligned}$$

since $m_{0,0}^0 \neq 0$, for each p , we are left with $l+1$ linear homogeneous equations for $l+1$ unknowns $\mu_{i+j, l-(i+j)}^p$. Since D is not singular, the linear system is nonsingular, and we arrive at $\mu_{i, l-i}^p = 0 \forall i \leq l$.

3 \Rightarrow 1: Once again, we use induction for the proof.

- 1) $k = q = 0$: $\mu_{0,0}^p = 0$ must hold for all wavelets.
- 2) $k + q = l$: Given that

$$0 = \left. \frac{\partial^{k+q} H^0(\omega_x, \omega_y)}{\partial \omega_x^k \partial \omega_y^q} \right|_{\omega=\omega_i}$$

for all aliasing frequencies ω_i and $(k+q) \leq l$, we assume that $\mu_{k,q}^p = 0; \forall p = 1, \dots, (M-1), k+q < l$ and prove that $\mu_{k,q}^p = 0$ for all $k+q = l$. The orthogonality condition in the modulation representation (Proposition 1) gives

$$H^0(\omega) H^p(\omega)^* + \sum_i H^0(\omega + \omega_i) H^p(\omega + \omega_i)^* = M \delta_{p,0}$$

where the sum is over all aliasing frequencies. Differentiating this equation l times, i.e., with $(k+q) = l$, we obtain

$$\begin{aligned} 0 &= \left. \frac{\partial^{k+q}}{\partial \omega_x^k \partial \omega_y^q} [H^0(\omega) H^p(\omega)^*] \right|_{\omega=0} \\ &\quad + \left. \frac{\partial^{k+q}}{\partial \omega_x^k \partial \omega_y^q} \left[\sum_i H^0(\omega + \omega_i) \right. \right. \\ &\quad \left. \left. \cdot H^p(\omega + \omega_i)^* \right] \right|_{\omega=0} \end{aligned}$$

Using condition 3,

$$\begin{aligned} 0 &= \sum_{s=0}^k \sum_{t=0}^q \binom{k}{s} \binom{q}{t} \\ &\quad \cdot \left. \frac{\partial^{s+t} H^0(\omega_x, \omega_y)}{\partial \omega_x^s \partial \omega_y^t} \right|_{\omega=0} \left. \frac{\partial^{k-s+q-t} H^p(\omega_x, \omega_y)^*}{\partial \omega_x^{k-s} \partial \omega_y^{q-t}} \right|_{\omega=0} \end{aligned}$$

and because of the assumption, only the term of $s = t = 0$ will contribute to the result; hence, for all $p = 1, \dots, M-1$, we are left with

$$0 = \sqrt{M} \mu_{k,q}^p = \sqrt{M} \left. \frac{\partial^{k+q} H^p(\omega_x, \omega_y)}{\partial \omega_x^k \partial \omega_y^q} \right|_{\omega=0} \quad (12)$$

1 \Rightarrow 3: This can be shown in a very similar manner by starting from the orthogonality of the rows of the modulation matrix, that is,

$$\sum_j H^j(\omega) H^j(\omega + \omega_k)^* = M \delta_{k,0} \quad (13)$$

and differentiating it. ■

As in 1-D, the fact that a wavelet system has L vanishing moments implies, and is implied by, the fact that all polynomials of degree less than L are included in the space spanned by the scaling function $\{\psi^0(x-n); n \in \mathbb{Z}\}$ [22]. This in return can be considered as a measure of the local approximation order. One main difference between the two-band case and M -band wavelets with $M \neq 2$ is that in the two-band case, the wavelet and scaling functions are fully defined by each other. The requirement of maximum number of vanishing moments, as can be seen from the third alternative in Theorem 1, implies that the scaling filter resembles an ideal low-pass filter, and in the two-band case, this implies that the wavelet filter is close to an ideal highpass filter. When we have more than one wavelet, the requirement of vanishing moments does not give any information about the division of bands between the wavelet filters, and these may overlap extensively.

We now deal with the issue of regularity. In the 1-D case, the relation between the number of vanishing moments of a wavelet function and its regularity has been investigated from several different viewpoints [7]–[9]. In all three papers cited above, it is concluded that a high number of vanishing moments is a necessary condition for obtaining high regularity, where regularity is measured by the Hölder exponent. No such results exist for general N -dimensional wavelets. Furthermore, as suggested in [23] and [24], for a given sampling scheme, the regularity varies considerably by altering the choice of the sampling matrix and the choice of the coset vectors. The generalization of the 1-D derivations is not trivial since they depend, in a crucial way, on the factorability of 1-D polynomials. We face this problem, once again, in our attempts to measure the regularity, in terms of the Hölder exponent, of a given wavelet function/filter. A different method was suggested by Villemoes and was adapted to the 2-D quincunx case [10]. This method works for Hölder exponents $\alpha < 1$, and we use it to get a lower bound for the regularity. In the examples considered in the following section, the regularity is assessed by using a simple generalization of the technique suggested in [9]. We define the approximation of order- j of the scaling function $\psi^0(\mathbf{x})$ at the points $\mathbf{x} = D^{-j} \mathbf{n}$ to be

$$h_j^0(\mathbf{n}) = \sqrt{M} \sum_{\mathbf{m}} h^0(\mathbf{n} - D\mathbf{m}) h_{j-1}^0(\mathbf{m}) \quad (14)$$

where $h_1^0(\mathbf{n}) = h^0(\mathbf{n})$. The regularity of a scaling filter/function is measured by the rate of convergence of the approximation to the scaling function.

Definition 2: The regularity measure of a 2-D scaling function/filter has the value α , where $N - 1 \leq \alpha < N$ if it has at least N vanishing moments and

$$\max_{l \leq N} \max_{\mathbf{n}} |\Delta^{(l, N-l)} h_j^0(\mathbf{n})| \leq CM^{-j(\alpha-N)/2}; \quad \forall j \quad (15)$$

where $\Delta^{(k,q)} h_j^0(\mathbf{n})$ stands for the finite difference of order $(k+q)$

$$\begin{aligned} \Delta^{(k,q)} h_j^0(n_1, n_2) &\triangleq [\Delta^{(k-1,q)} h_j^0(n_1, n_2) - \Delta^{(k-1,q)} h_j^0(n_1-1, n_2)]/M^{-j/2} \\ &= [\Delta^{(k,q-1)} h_j^0(n_1, n_2) - \Delta^{(k,q-1)} h_j^0(n_1, n_2-1)]/M^{-j/2} \end{aligned}$$

with

$$\Delta^{(1,0)} h_j^0(n_1, n_2) = [h_j^0(n_1, n_2) - h_j^0(n_1-1, n_2)]/M^{-j/2}$$

and

$$\Delta^{(0,1)} h_j^0(n_1, n_2) = [h_j^0(n_1, n_2) - h_j^0(n_1, n_2-1)]/M^{-j/2}.$$

As mentioned above, it is very difficult to give general results regarding the regularity of wavelets in more than 1-D. The next theorem, which is a generalization of the 1-D version given in [7] and is given for a certain class of sampling matrices, shows, however, that the number of vanishing moments is related to the regularity in more than 1-D as well. (Although it is stated for orthogonal wavelets with compact support, it has been brought to our attention that it is true also for more general wavelets with sufficient decay that constitute a Riesz basis and for all sampling matrices.)

Theorem 2: In the case of sampling matrices D for which there exists an integer n such that $D^n = M^n I = \beta I$, a compactly supported wavelet function $\psi(\mathbf{x})$ that is orthogonal, i.e.,

$$\int \psi_j(\mathbf{n}(\mathbf{x})) \psi_i(\mathbf{m}(\mathbf{x})) d\mathbf{x} = \delta_{j,i} \delta_{\mathbf{n},\mathbf{m}} \quad (16)$$

is in C^{m+1} only if all its first $(m+1)$ moments vanish, that is, using the notation of Theorem 1

$$m_{k,q} = \int \mathbf{x}^q \psi(\mathbf{x}) d\mathbf{x} = 0, \quad \forall (k+q) = 0, \dots, m. \quad (17)$$

Proof: (This is based on the 1-D proof in [25]). First, we show that the theorem is correct for $m = 0$. From the orthogonality, we have for every $j \geq j_0$

$$\begin{aligned} 0 &= \int \psi(\mathbf{x}) \psi(D^j \mathbf{x} - D^{(j-j_0)} \mathbf{n}_0) d\mathbf{x} \\ &= M^{-j} \int \psi(D^{-j} \xi + D^{-j_0} \mathbf{n}_0) \psi(\xi) d\xi. \end{aligned} \quad (18)$$

If we write $j = pn$ and take the limit of the above expression as p grows to infinity, we have

$$\begin{aligned} 0 &= \lim_{p \rightarrow \infty} \int \psi(D^{-pn} \xi + D^{-j_0} \mathbf{n}_0) \psi(\xi) d\xi \\ &= \int \lim_{p \rightarrow \infty} \psi(\beta^{-p} \xi + D^{-j_0} \mathbf{n}_0) \psi(\xi) d\xi \\ &= \int \psi(D^{-j_0} \mathbf{n}_0) \psi(\xi) d\xi \\ &= \psi(D^{-j_0} \mathbf{n}_0) \int \psi(\xi) d\xi. \end{aligned} \quad (19)$$

Note that we can change the order of the integral and the limit since $\psi(\xi)$ is in C^1 and is compactly supported. This ends the proof for $m = 0$ since we can choose the point $D^{-j_0} \mathbf{n}_0$ so that $\psi(D^{-j_0} \mathbf{n}_0) \neq 0$. We proceed to prove the theorem by induction on l . Given that $\psi(\mathbf{x}) \in C^{l+1}$, with $l \leq m$, we assume that all moments vanish for $(k+q) = i < l$ and prove that all the moments with $(k+q) = l$, equal zero. We use Taylor's Theorem. Since $\psi(\xi)$ is in C^{l+1} , denoting $\xi = (x, y)$, we have

$$\begin{aligned} \psi(D^{-pn} \xi + D^{-j_0} \mathbf{n}_0) &= \sum_{i=0}^l \frac{1}{i!} \sum_{k=0}^i \binom{i}{k} \frac{\partial^i \psi(D^{-j_0} \mathbf{n}_0)}{\partial x^k \partial y^{i-k}} (D^{-pn} \xi)^{(k, i-k)} \\ &\quad + R_l(D^{-pn} \xi + D^{-j_0} \mathbf{n}_0) \end{aligned} \quad (20)$$

where the remanent term R_l is uniformly bounded, and

$$\lim_{\xi \rightarrow 0} \frac{R_l(\xi + D^{-j_0} \mathbf{n}_0)}{|\xi|^l} = 0.$$

Hence, (18) can be written as

$$\begin{aligned} \sum_{i=0}^l \sum_{k=0}^i \binom{i}{k} \frac{\partial^i \psi(D^{-j_0} \mathbf{n}_0)}{i! \partial x^k \partial y^{i-k}} \int (D^{-pn} \xi)^{(k, i-k)} \psi(\xi) d\xi \\ + \int R_l(D^{-pn} \xi + D^{-j_0} \mathbf{n}_0) \psi(\xi) d\xi = 0. \end{aligned} \quad (21)$$

Using $D^n = \beta I$, we have

$$(D^{-pn} \xi)^{(k, i-k)} = \beta^{-pi} x^k y^{i-k}$$

and

$$|D^{-pn} \xi|^i = \beta^{-pi} |\xi|^i.$$

According to our assumption, only the term with $i = l$ is nonzero. Hence, by dividing both terms of (21) by β^{-pl} , we have

$$\begin{aligned} 0 &= \frac{1}{l!} \sum_{k=0}^l \binom{l}{k} \frac{\partial^l \psi(D^{-j_0} \mathbf{n}_0)}{\partial x^k \partial y^{l-k}} \int x^k y^{l-k} \psi(\xi) d\xi \\ &\quad + \int \frac{R_l(\beta^{-p} \xi + D^{-j_0} \mathbf{n}_0)}{|\beta^{-p} \xi|^l} |\xi|^l \psi(\xi) d\xi. \end{aligned} \quad (22)$$

Once more, we take the limit of the above equation as $p \rightarrow \infty$ and change the order of the integration and limit to obtain

$$\begin{aligned} 0 &= \frac{1}{l!} \sum_{k=0}^l \binom{l}{k} \frac{\partial^l \psi(D^{-j_0} \mathbf{n}_0)}{\partial x^k \partial y^{l-k}} \int x^k y^{l-k} \psi(\xi) d\xi \\ &\quad + \int \lim_{p \rightarrow \infty} \frac{R_l(\beta^{-p} \xi + D^{-j_0} \mathbf{n}_0)}{|\beta^{-p} \xi|^l} |\xi|^l \psi(\xi) d\xi \\ &= \frac{1}{l!} \sum_{k=0}^l \binom{l}{k} \frac{\partial^l \psi(D^{-j_0} \mathbf{n}_0)}{\partial x^k \partial y^{l-k}} \int x^k y^{l-k} \psi(\xi) d\xi. \end{aligned} \quad (23)$$

In order to complete the proof, we still have to show that each term in the above sum is zero. Since (23) is true for every point $D^{-j_0} \mathbf{n}_0$, i.e., for every point in a set dense in \mathbb{R}^2 , this can be done by showing that the $l+1$ functions $\partial^l \psi(x, y) / \partial x^k \partial y^{l-k}$

are independent. If

$$0 \equiv \sum_{k=0}^l a_k \frac{\partial^l \psi(x, y)}{\partial x^k \partial y^{l-k}}$$

we can Fourier transform this equation and integrate by parts. Using the fact that all functions vanish at infinity, we have

$$0 \equiv \sum_{k=0}^l a_k \int_{-\infty}^{\infty} \frac{\partial^l \psi(x, y)}{\partial x^k \partial y^{l-k}} e^{-i(x\omega_x + y\omega_y)} dx dy$$

$$0 \equiv \sum_{k=0}^l a_k \omega_x^k \omega_y^{l-k} \int_{-\infty}^{\infty} \psi(x, y) e^{-i(x\omega_x + y\omega_y)} dx dy.$$

ψ cannot be the zero function; hence, we are guaranteed that there exist no neighborhood at which the Fourier transform of ψ is zero since ψ is of compact support. Thus, the 2-D polynomial must be the zero function, i.e., $a_k = 0 \forall k$, thus proving that the $l+1$ functions are independent. ■

It goes without saying that Theorem 2 can be applied to each of the $M-1$ wavelet functions $\psi^p(\mathbf{x})$. Note also that the proof depends on the orthogonality of the functions $\psi_j(\mathbf{n}(\mathbf{x}))$ and does not hold for wavelets that are only orthogonal in the filter bank sense, i.e., that are only lossless. The specific examples given in [24] and [23] indicate that choosing the sampling matrix so that $D^n = \beta I$ results indeed in more regular wavelets.

IV. DESIGNING WAVELETS WITH VANISHING MOMENTS

We now turn to the actual design of orthogonal factorable wavelets. A rank- M wavelet matrix is constructed from paraunitary building blocks of degree one. With the wavelet matrix at hand, the wavelet filters are reconstructed from their polyphase components

$$(H^0(z_1, z_2), H^1(z_1, z_2), \dots, H^{M-1}(z_1, z_2)) \\ = (z^{-\mathbf{k}_0}, z^{-\mathbf{k}_1}, \dots, z^{-\mathbf{k}_{M-1}}) \mathbf{H}_I(z^D) \mathcal{H}^{(M)} \quad (24)$$

where \mathbf{k}_i is the shift vector corresponding to the i th coset. The general form of the HWM is

$$\mathcal{H}^{(M)} = \mathcal{H}_c^{(M)} \begin{pmatrix} 1 & 0 \\ 0 & O_{M-1} \end{pmatrix} \quad (25)$$

where O_{M-1} is an orthogonal $(M-1) \times (M-1)$ matrix, and $\mathcal{H}_c^{(M)}$ is the canonical M -band HWM [17]. The first column of the canonical HWM is $1/\sqrt{M}(1, 1, \dots, 1)^T$; hence, the scaling filter is given by

$$H^0(z_1, z_2) = (z^{-\mathbf{k}_0}, z^{-\mathbf{k}_1}, \dots, z^{-\mathbf{k}_{M-1}}) \mathbf{H}_I(z^D) \frac{1}{\sqrt{M}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \quad (26)$$

Thus, a factorable scaling filter is defined by $(m+n)$ unit vectors, that is, by $(M-1)(m+n)$ free parameters. The other $\binom{M-1}{2}$ parameters that define the HWM do not effect the scaling filter but only the other wavelet filters.

Using any of the three alternatives of Theorem 1, one can see that for L vanishing moments, we have $N_c = (M -$

$1) \binom{L+1}{2} - 1$) constraints; hence, in general, we need wavelets with at least $N_{fp} = N_c$ free parameters. If one uses the third alternative of Theorem 1, that is, (9), it is obvious that only the parameters defining the scaling filter can be taken into account. The other parameters do not appear in (26) and therefore do not appear in (9) either. This decoupling of the Haar wavelet parameters from the vanishing moment constraints is the reason for defining the polyphase (wavelet) matrix the way we did, instead of the more common way, which uses the transpose of the one we use. (Another way of reaching this goal is to use a different factorization in which the characteristic matrix is the first factor on the left and not on the right.) These considerations suggest that in order to obtain wavelets with L vanishing moments, we should use wavelets with McMillan degrees $(m+n) \geq \binom{L+1}{2} - 1$. Furthermore, in both cases dealt with—the two-band quincunx and four-band separable sampling—it is obvious from 1-D results that each polynomial order should be larger or equal to $L-1$. Otherwise, we could find 1-D projections with L vanishing moments and length less than $2L$, contrary to the well-known 1-D results [5]. The rest of the section is devoted to two sampling schemes: the quincunx and four-band separable sampling schemes.

A. Quincunx Wavelets

The quincunx sampling is a 2-D, two-band, nonseparable sampling scheme. Detailed analysis of 2-D multirate systems with quincunx sampling can be found in Kovačević's doctoral thesis [3]. The sampling matrix and HWM we use are (there are several other possibilities)

$$D = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}; \quad \mathcal{H}^{(2)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and choose $\mathbf{k}_i = \{(0, 0)^T, (1, 0)^T\}$ as the coset shift vectors. With this choice of the sampling matrix, we have $D^2 = 2I$, and it has been shown to result in wavelets with higher regularity than with other sampling matrices [3], [24], [23]. In the quincunx case, the McMillan degrees (m, n) of the wavelet matrix are equal to the polynomial orders of the elements of the matrix. Similar to the two-band 1-D case, the wavelet filter is fully determined by the scaling filter. In the two-band factorable case, the vectors V_k in (6) can be parameterized by one parameter $V(\theta_k) = (\cos(\theta_k), \sin(\theta_k))^T$, where $\theta_k \in [0, 2\pi)$. The general form for the vanishing moment constraints on the angles θ_k is developed in Appendix A. We use the third alternative in Theorem 1, i.e., we require the frequency response of the scaling filter to have a zero of order L at the aliasing point $(z_1 = -1, z_2 = -1)$. We get the following $N_c = \binom{L+1}{2} - 1$ constraints:

$$0 = \frac{\partial^{p+l} H^0(z_1, z_2)}{\partial z_1^p \partial z_2^l} \Big|_{z_1=z_2=-1}$$

$$= \sum_{(k+q)\text{odd}} 2\Gamma_{k,q}^{m,n}(p, l) C_{k,q} + p\Gamma_{k,q}^{m,n} \cdot (p-1, l)(S_{k,q} - C_{k,q})$$

$$- \sum_{(k+q)\text{even}} 2\Gamma_{k,q}^{m,n}(p, l) S_{k,q} + p\Gamma_{k,q}^{m,n} \cdot (p-1, l)(S_{k,q} + C_{k,q}) \quad (27)$$

with $(p + l) = 1, 2, \dots, (L - 1)$. The coefficients $\Gamma_{k,q}^{m,n}(p, l)$ were computed and tabulated, and $S_{k,q}$ and $C_{k,q}$ are trigonometric functions of the angles $\{\theta_k\}_{k=1}^{m+n}$. As mentioned earlier, in order to achieve L vanishing moments, one should use, in general, wavelets of degrees $(m + n) \geq N_c$. Since the wavelets with vanishing moments are obtained by solving a set of nonlinear equations, in general, we will have more than one solution with minimal degrees for a given number of vanishing moments. Contrary to the 1-D wavelets with vanishing moments dealt with by Daubechies, different solutions may have completely different spectral characteristics.

We would like to mention the special case of solutions that are actually 1-D filters. A degree $L - 1$ 1-D wavelet filter with L vanishing moments is likewise a factorable quincunx wavelet filter with degrees $(L - 1, L - 1)$ and L vanishing moments. This can easily be demonstrated by factorizing each degree-1 factor of the 1-D polyphase matrix as

$$\begin{aligned} &(I + (z_1^2 - 1)V_i V_i^T) \\ &= (I + (z_1 z_2 - 1)V_i V_i^T)(I + (z_1 z_2^{-1} - 1)V_i V_i^T). \end{aligned}$$

These 1-D quincunx wavelets were suggested by Cohen and Daubechies [24]. Although the wavelet filters are 1-D, due to the nonseparable subsampling, the resulting wavelet functions are 2-D nonseparable functions. Thus, the above arguments concerning the minimal degrees needed in order to obtain L vanishing moments are true in general, but for the special case of 1-D filters, we can satisfy the constraints by using wavelets of degrees $(L - 1, L - 1)$, even though $2(L - 1) < N_c$ for all $L > 2$. We now give a few examples, but first, we note the following:

- The vectors V_i appear in $H(z_1, z_2)$ (see (6)) only in the combination $V_i V_i^T$; hence, since $V(\theta_i) V(\theta_i)^T = V(\pi + \theta_i) V(\pi + \theta_i)^T$, we can limit the search to $\theta_i \in [-(\pi/2), (\pi/2)]$.
- For every solution $\{\theta_i\}$, there exists another solution $\{\pi - \theta_i\}$ that will give rise to a scaling filter (function), which is the reflection of the original filter (function) through its center, i.e., $\tilde{h}^0(n_1, n_2) = h^0(N_1 - n_1, N_2 - n_2)$.
- Given a solution $h^0(n_1, n_2)$ with degrees (m, n) , there exists another solution with degrees (n, m) in which the variable z_1 is interchanged with z_2 in the polyphase matrix. This will cause the transformation $h^i(n_1, n_2) \rightarrow h^i(n_1, -n_2)$.

Example 1.1: The smallest quincunx wavelet filter is the Haar wavelet; its degrees are $(m, n) = (0, 0)$. Thus, the corresponding wavelet matrix is the HWM $\mathcal{H}^{(2)}$. The filters are exactly the 1-D Haar filters, i.e.,

$$h^0(n_1, n_2) := \frac{1}{\sqrt{2}}(1, 1), h^1(n_1, n_2) := \frac{1}{\sqrt{2}}(1, -1)$$

and only the first moment vanishes. The absolute value of the frequency response of the scaling filter, and the scaling function in this case, are plotted in Fig. 1(a) and (b), respectively.

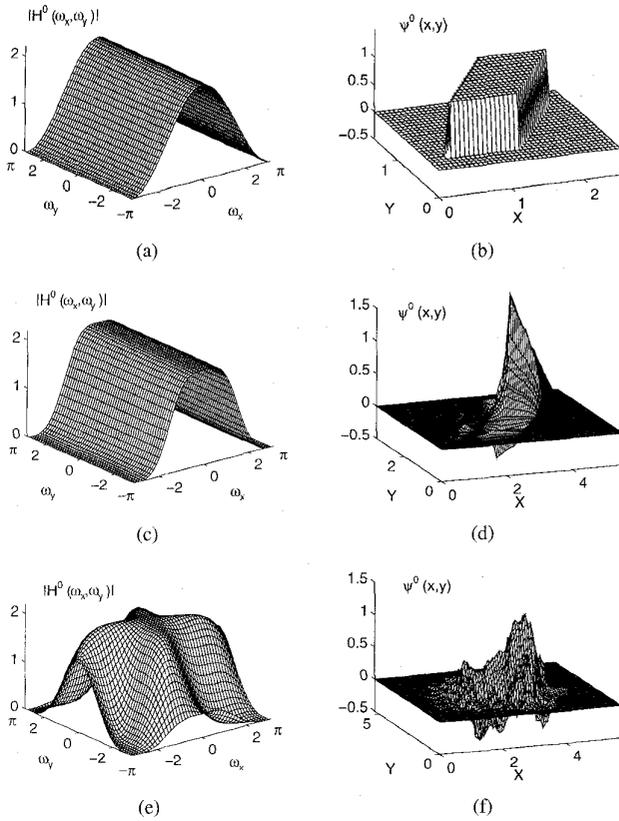


Fig. 1. Quincunx scaling filters and functions. The frequency response of the scaling filter and the scaling function of (a) and (b) the Haar quincunx wavelet (Example 1.1), (c) and (d) wavelet with two vanishing moments, 1-D solution (Example 1.2), (e) and (f) wavelet with two vanishing moments, second solution (Example 1.2).

Example 1.2: We next deal with wavelets with $L = 2$ vanishing moments; in this case, the minimal degrees are $(1, 1)$. The two constraint equations are

$$\begin{aligned} 0 &= 1 - \cos(2\theta_1) - \cos(2\theta_2) \\ 0 &= \cos(2\theta_1) - \cos(2\theta_2). \end{aligned}$$

We have two different solutions (and their reflections):

- 1) $2\theta_1 = 2\theta_2 = \pi/3$: which is the four-tap filter of Daubechies with two vanishing moments

$$\begin{aligned} &h^0(n_1, n_2) \\ &:= \frac{1}{4\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 + \sqrt{3} & 3 + \sqrt{3} & 3 - \sqrt{3} & 1 - \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

- 2) $2\theta_1 = -2\theta_2 = \pi/3$: which produces the scaling filter (which is given in [14])

$$\begin{aligned} &h^0(n_1, n_2) \\ &:= \frac{1}{8\sqrt{2}} \begin{pmatrix} 0 & 3 + 3\sqrt{3} & 3 + \sqrt{3} & 0 \\ -1 - \sqrt{3} & 3 + \sqrt{3} & 3 - \sqrt{3} & -1 + \sqrt{3} \\ 0 & 3 - \sqrt{3} & 3 - 3\sqrt{3} & 0 \end{pmatrix}. \end{aligned}$$

The frequency response of the scaling filter and the scaling function for the first solution are shown in Fig 1(c) and (d), and the corresponding functions for the second solution are depicted in Fig. 1(e) and (f).

For $L > 2$ the constraint equations cannot be solved analytically, and a numerical program given in *Numerical Recipes* [26] was used for the solution of the nonlinear equations. In addition to the $(m+n)$ free parameters defining the wavelet matrix, there are $\binom{m+n}{n}$ alternatives in which the factors of degree 1 can be arranged in (6). Each alternative results in a different set of constraints, thus introducing another discrete parameter into the problem. It turns out that a solution does not exist for all possible alternatives. For example, analyzing the equations for wavelet filters with degrees (3, 2), we found that one cannot obtain three vanishing moments in the two cases in which all the factors of degree 1 in one variable precede the factors in the other variable.

Example 1.3: The minimal degrees for obtaining three vanishing moments (not including 1-D filters) are (3, 2). The functions plotted in Fig. 2(a) and (b) correspond to a solution obtained by choosing the combination $(z_1, z_1, z_2, z_1, z_2)$ and the five parameters as

$$\{2\theta\}_{i=1}^5 = \{-0.89679, -2.93452, 0.78210, 0.54617, 1.78178\}.$$

In Fig. 2(c) and (d), we show another solution; this solution was obtained using the order $(z_1, z_1, z_2, z_2, z_1)$ and the parameters

$$\{2\theta\}_{i=1}^5 = \{0.80821, 1.25886, 0.78210, 1.78178, -2.09172\}.$$

Although both solutions have three vanishing moments, there is a considerable difference between them. The first solution is much closer to an ideal quincunx lowpass filter, with 75% of its energy concentrated in the lowpass band, as opposed to the second solution, which has 60%. The difference becomes very apparent by looking at the scaling functions. The regularity measure, which was introduced in the previous section, gave the value $\alpha = 0.6 \pm 0.1$ for the first solution and $\alpha = 0.05 \pm 0.1$ for the second. Applying the method suggested by Villemoes [10] to the second example, we attempted to lower bound the Hölder exponent α . We were not able to show that the function is even continuous ($\alpha > 0$) when applying the method for resolution levels $j \leq 14$. The 1-D solution corresponding to the Daubechies wavelet with three vanishing moments obtained the highest value of regularity with $\alpha = 1.0 \pm 0.1$, which agrees with the 1-D bound for the Hölder exponent of $\alpha = 1.0878$.

For the sake of completeness, we searched the entire space of wavelet matrices with degrees (3, 2) and three vanishing moments; we did so using the complete characterization of 2-D PUP matrices given in [12]. To obtain compactly supported wavelets with three vanishing moments, one must solve 15 equations for 15 variables. All the solutions found coincide with solutions that were obtained using factorable wavelets, but it is feasible that other solutions, that were not found numerically, do exist.

Finally, we give two examples of quincunx scaling filters and functions with four vanishing moments. The frequency

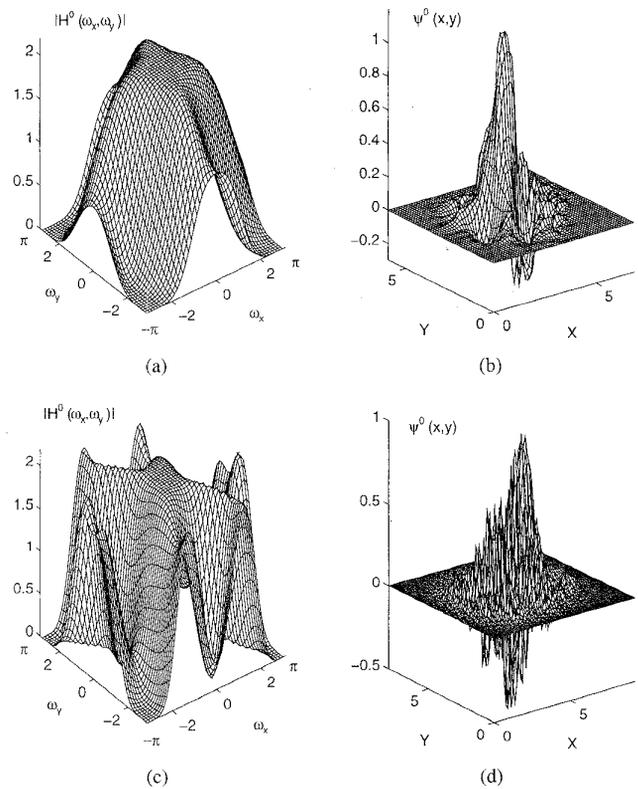


Fig. 2. Quincunx scaling filters and functions with three vanishing moments. The frequency response of the scaling filter and the scaling function of the first ((a) and (b)) and second ((c) and (d)) solutions of Example 1.3. Both scaling functions show the eighth-order approximation, using the subdivision scheme.

response of the scaling filter and the scaling function are plotted in Fig. 3(a) and (b) for one example and in Fig. 3(c) and (d) for the other. The first example has a good frequency response with 80% of the energy in the lowpass band but is very spiky: $\alpha = 0.5 \pm 0.1$. The second example is more regular, with $\alpha = 0.9 \pm 0.1$, but with a slightly inferior frequency response.

B. Four-Band, Separable Sampling Wavelets

This sampling scheme is the most commonly used in all applications of the 2-D DWT since it is the simplest generalization of the two-band 1-D sampling. The sampling matrix is $D = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, and the four-band HWM has the general form

$$\mathcal{H}^{(4)} = \frac{1}{2} \begin{pmatrix} 1 & -\sqrt{3} & 0 & 0 \\ 1 & 1/\sqrt{3} & -2\sqrt{\frac{2}{3}} & 0 \\ 1 & 1/\sqrt{3} & \sqrt{\frac{2}{3}} & -\sqrt{2} \\ 1 & 1/\sqrt{3} & \sqrt{\frac{2}{3}} & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & O_3 \end{pmatrix}. \quad (28)$$

Thus, we have a three-parameter space of HWM. We chose the four coset shift vectors to be $k_i = \{(0, 0)^T, (1, 0)^T, (0, 1)^T, (1, 1)^T\}$; hence, the three points of repeated spectra are $\omega_i = \{(\pi, 0)^T, (0, \pi)^T, (\pi, \pi)^T\}$. With this choice, the expression for a factorable scaling filter of

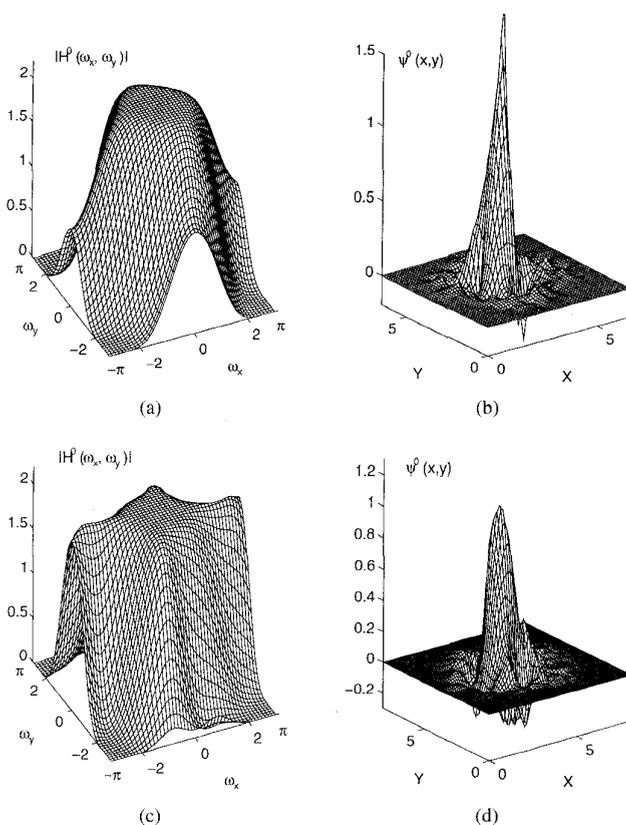


Fig. 3. Quincunx scaling filters and functions with four vanishing moments. The frequency response of the scaling filter and the scaling function of the two examples mentioned in the text.

McMillan degree (m, n) is, according to (26),

$$H^0(z_1, z_2) = \frac{1}{2}(1, z_1, z_2, z_1 z_2) \prod_{k=1}^{m+n} [I + (z_k^{-2} - 1)V_k V_k^T] \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}. \quad (29)$$

Once again, in order to obtain L vanishing moments, we impose $N_c = 3\binom{L+1}{2} - 1$ constraints on the $3(m+n)$ free parameters defining the scaling filter. Just as we did in the quincunx case, we seek solutions with the minimal possible degrees, which, in general, are given by $N_{fp} = N_c$, that is, $(m+n) = \binom{L+1}{2} - 1$. Once we have designed the scaling function by fixing the $3(m+n)$ free parameters, we still have three free parameters defining the HWM.

Note that using 1-D wavelets with L vanishing moments to design separable 2-D wavelets will result in wavelets with McMillan degrees $(2L-2, 2L-2)$ (see [20]). Since in this case $(m+n) = 4(L-1) \neq \binom{L+1}{2} - 1$ for all $L \neq 6$, in the following examples, we will not obtain separable filters as private solutions.

Example 2.1: The Haar wavelet matrix gives rise to the smallest wavelet filters. Each filter has four nonzero coefficients, and the corresponding function is compactly supported

on $[0, 1]^2$. In this case, only the first moment vanishes, and the resulting functions are discontinuous and piecewise constant.

Example 2.2: In order to obtain wavelets with $L = 2$ vanishing moments, we must use wavelets with McMillan degrees $(1, 1)$; thus, the wavelet matrix is given by

$$H(z_1, z_2) = [I + (z_1^{-1} - 1)V_1 V_1^T][I + (z_2^{-1} - 1)V_2 V_2^T]\mathcal{H}^{(4)}.$$

We have six bilinear constraints (two derivatives at three aliasing frequencies) on eight unknowns (two vectors of length 4) defined by six free parameters. The two unit vectors solving this set of equations are

$$V_1 = \frac{\pm 1}{2\sqrt{2}} \begin{pmatrix} \sqrt{3} \\ 1 \\ \sqrt{3} \\ 1 \end{pmatrix}, \quad V_2 = \frac{\pm 1}{2\sqrt{2}} \begin{pmatrix} \sqrt{3} \\ \sqrt{3} \\ 1 \\ 1 \end{pmatrix}.$$

As mentioned previously, the three parameters defining the HWM do not influence the number of vanishing moments but rather control, in some sense, the orientation of the three wavelet filters/functions. In the following example, we used the separable HWM

$$\mathcal{H}^{(4)} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

This is one way to guarantee that each wavelet filter has a maxima at one aliasing point and a zero at all others.

Once again, for $L > 2$, only numerical solutions can be obtained. In Fig. 4, an example of wavelet filters and functions with four vanishing moments are depicted. Their McMillan degrees are $(5, 4)$, the filters are of length $(12, 10)$, and the regularity of the scaling function was measured to be $\alpha = 1.25 \pm 0.1$.

V. DISCUSSION AND CONCLUSIONS

Several methods have been suggested for constructing 2-D orthogonal wavelet filters and functions [1], [24]. These methods are based on some sort of transformation from 1-D wavelets. Hence, the resulting wavelets in each case constitute only a small restricted subset of all possible orthogonal wavelets. Here, we use a parameterization of all factorable orthogonal wavelets, which is a much more general subset. The subset of factorable wavelets was shown in [12] to have the same number of free parameters as the set of all orthogonal wavelets and, contrary to the complete set of orthogonal wavelet, has a simple parameterization. This subset includes the previously suggested 2-D orthogonal wavelets, such as separable wavelets [20] and the special (actually 1-D) quincunx wavelets [24].

In order to obtain wavelets with several vanishing moments, one has, in general, to solve numerically a set of nonlinear equations. Even when focusing our attention only on wavelets with maximal number of vanishing moments for given McMillan degrees, considerable freedom is still left since the set of nonlinear equations have many solutions. Furthermore, usually, satisfying the conditions of having a large number of

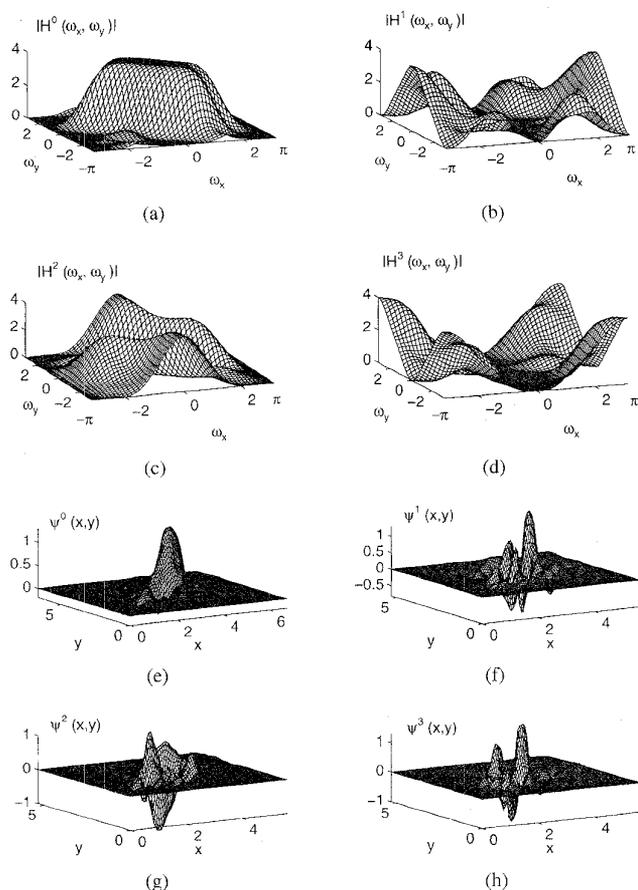


Fig. 4. Four-band, separable sampling wavelet with four vanishing moments. (a)–(d) show the magnitude response of the four wavelet filters, and (e)–(h) show the wavelet functions.

vanishing moments is not enough, as is illustrated in Figs. 2 and 3. Extra care should be taken in choosing the best solution for a given 2-D application. Theorem 2 and our examples show that a sufficient number of vanishing moments is a necessary, but by no means sufficient, condition for the regularity of the wavelet function. The regularity of the original Daubechies orthogonal wavelets (minimal phase wavelets) was found to grow at least linearly with the number of vanishing moments, yet in 2-D, we were not able to select any special subset of solutions and to give a lower bound for their regularity. For a high degree of regularity, one should probably *sacrifice* some of the vanishing moments and use the extra degrees of freedom to optimize the regularity. For a given order of moments, one can relax only part of the partial derivative constraints (9) and only at some of the aliasing points. Our treatment of the regularity issue is quite preliminary and mainly serves as a motivation for investigating 2-D wavelets with vanishing moments. Further work is required in order to establish the exact relation between the number of vanishing moments and regularity and its dependency on the choice of sampling matrix and coset vectors.

As in many other areas of mathematics, most of the difficulties encountered when trying to generalize 1-D theory to higher dimensions arise already at the transition from 1-D to

2-D. Once these difficulties have been overcome, the design of N -dimensional wavelets will be a simple generalization of the 2-D method.

APPENDIX A

In this appendix, we will develop the constraint equations for vanishing moments in the case of factorable quincunx wavelets. Each factor of degree 1 is defined by one parameter θ_i

$$W_i = (I + (z_i^{-1} - 1)V_i V_i^T)$$

where $z_i = z_1$ or z_2 and V_i is a unit vector of length 2. If we use $V_i = (\cos(\theta_i), \sin(\theta_i))^T$, we get

$$W_i(\theta_i) = \frac{1}{2}(z_i^{-1} + 1)I + \frac{1}{2}(z_i^{-1} - 1) \begin{pmatrix} \cos(2\theta_i) & \sin(2\theta_i) \\ \sin(2\theta_i) & -\cos(2\theta_i) \end{pmatrix}. \quad (\text{A.1})$$

Next, we define two matrices: the 2-D rotation matrix $R(\theta)$ and the rotation-plus-reflection matrix $T(\theta)$

$$R(\theta) = \begin{pmatrix} \cos(2\theta) & -\sin(2\theta) \\ \sin(2\theta) & \cos(2\theta) \end{pmatrix}$$

$$T(\theta) = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}$$

and recall the relations

$$T(\alpha)T(\beta) = R(\alpha - \beta); \quad R(\alpha)R(\beta) = R(\alpha + \beta);$$

$$R(\alpha)T(\beta) = T(\alpha + \beta); \quad T(\alpha)R(\beta) = T(\alpha - \beta). \quad (\text{A.2})$$

From this, we see that the 2-D polyphase matrix of McMillan degrees (m, n) can be written as

$$H_I(z_1, z_2) = \prod_{i=1}^{m+n} W_i(\theta_i)$$

$$= \sum_{k=0}^m \sum_{q=0}^n \gamma_{k,q}^{m,n}(z_1, z_2) T_{k,q}(\{\theta\}) \quad (\text{A.3})$$

where we define

$$\gamma_{k,q}^{m,n}(z_1, z_2) \triangleq \left(\frac{1}{2}\right)^{m+n} (z_1^{-1} + 1)^{m-k} (z_1^{-1} - 1)^k \cdot (z_2^{-1} + 1)^{n-q} (z_2^{-1} - 1)^q$$

and $T_{k,q}(\{\theta\})$ is the sum of all possible combinations of an ordered product of k T -matrices of z_1 factors and q T -matrices of z_2 factors. With use of the relations (A.2), one can see that for a given k and q , $T_{k,q}(\{\theta\})$ will include terms of the following type:

$$T_{k,q}(\{\theta\}) \triangleq \begin{cases} \begin{pmatrix} C_{k,q} & -S_{k,q} \\ S_{k,q} & C_{k,q} \end{pmatrix} \\ \begin{pmatrix} C_{k,q} & S_{k,q} \\ S_{k,q} & -C_{k,q} \end{pmatrix} \end{cases}$$

$$\triangleq \begin{cases} \sum R(\theta_{j_1} - \theta_{j_2} + \theta_{j_3} \cdots - \theta_{j_{k+q}}), & (k+q) \text{ even} \\ \sum T(\theta_{j_1} - \theta_{j_2} + \theta_{j_3} \cdots + \theta_{j_{k+q}}), & (k+q) \text{ odd.} \end{cases}$$

For example, with degrees $(m, n) = (2, 2)$, where the first two factors are in z_1 and the last in z_2 , we have

$$\begin{aligned}
16H(z_1, z_2) = & (z_1^{-1} + 1)^2(z_2^{-1} + 1)^2 I \\
& + (z_1^{-2} - 1)(z_2^{-1} + 1)^2 [T(\theta_1) + T(\theta_2)] \\
& + (z_1^{-1} + 1)^2(z_2^{-2} - 1) [T(\theta_3) + T(\theta_4)] \\
& + (z_1^{-2} - 1)(z_2^{-2} - 1) \\
& \times [R(\theta_1 - \theta_3) + R(\theta_1 - \theta_4) + R(\theta_2 - \theta_3) \\
& + R(\theta_2 - \theta_4)] \\
& + (z_1^{-1} - 1)^2(z_2^{-1} + 1)^2 [R(\theta_1 - \theta_2)] \\
& + (z_1^{-1} + 1)^2(z_2^{-1} - 1)^2 [R(\theta_3 - \theta_4)] \\
& + (z_1^{-1} - 1)^2(z_2^{-2} - 1) \\
& \times [T(\theta_1 - \theta_2 + \theta_3) + T(\theta_1 - \theta_2 + \theta_4)] \\
& + (z_1^{-2} - 1)(z_2^{-1} - 1)^2 \\
& \times [T(\theta_1 - \theta_3 + \theta_4) + T(\theta_2 - \theta_3 + \theta_4)] \\
& + (z_1^{-1} - 1)^2(z_2^{-1} - 1)^2 [R(\theta_1 - \theta_2 \\
& + \theta_3 - \theta_4)].
\end{aligned}$$

With an explicit expression for $H(z_1, z_2)$, we can obtain the vanishing moments constraints as constraints on the parameters θ_i . We construct the scaling filter from the polyphase matrix

$$H^0(z_1, z_2) = (1, z_1) \mathbf{H}(z_1 z_2, z_1 z_2^{-1}) \frac{1}{\sqrt{2}} (1, 1)^T \quad (\text{A.4})$$

and then we write the partial derivatives of the scaling function

$$\begin{aligned}
\sqrt{2} \frac{\partial^{p+l} H^0(z_1, z_2)}{\partial z_1^p \partial z_2^l} \Big|_{z_1=z_2=-1} \\
= \sum_{k=0}^m \sum_{q=0}^n [\Gamma_{k,q}^{m,n}(p, l) (C_{k,q}(1 \mp 1) - S_{k,q}(1 \pm 1)) \\
+ p \Gamma_{k,q}^{m,n}(p-1, l) (S_{k,q} \pm C_{k,q})]
\end{aligned}$$

where the upper sign should be used when $(k+q)$ is even and the lower sign otherwise, and we used the notation

$$\Gamma_{k,q}^{m,n}(p, l) \triangleq \frac{\partial^{p+l} \gamma_{k,q}^{m,n}(z_1 z_2, z_1 z_2^{-1})}{\partial z_1^p \partial z_2^l} \Big|_{z_1=z_2=-1}$$

Finally, we arrive at the constraints for L vanishing moment. By equating all partial derivatives $(p+l) = 1, 2, \dots, (L-1)$ to zero, we get $\binom{L+1}{2} - 1$ constraints

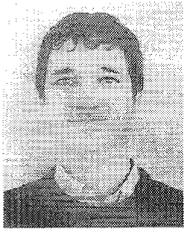
$$\begin{aligned}
0 = & \sum_{(k+q)\text{even}} [-2\Gamma_{k,q}^{m,n}(p, l) S_{k,q} \\
& + p \Gamma_{k,q}^{m,n}(p-1, l) (S_{k,q} + C_{k,q})] \\
& + \sum_{(k+q)\text{odd}} [2\Gamma_{k,q}^{m,n}(p, l) C_{k,q} \\
& + p \Gamma_{k,q}^{m,n}(p-1, l) (S_{k,q} - C_{k,q})]. \quad (\text{A.5})
\end{aligned}$$

ACKNOWLEDGMENT

The authors would like to thank N. Cohen and M. Zibulski of the Department of Electrical Engineering and Y. Binyamini of the Department of Mathematics at the Technion for their help. They also wish to thank L. Villemoes for his help in applying his methods to our examples.

REFERENCES

- [1] S. Mallat, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, pp. 674-693, 1989.
- [2] T. Chen and P. P. Vaidyanathan, "Multidimensional multirate filters and filter banks derived from one-dimensional filters," *IEEE Trans. Signal Processing*, vol. 41, May 1993.
- [3] J. Kovačević, "Filter banks and wavelets: Extensions and applications," PhD thesis, Dept. of Elect. Eng., Columbia Univ., New York, NY, 1991.
- [4] C. Lien and C. Huang, "The design of 2-D nonseparable directional perfect reconstruction filter banks," *Multidimensional Syst. Signal Processing*, vol. 5, p. 289, 1994.
- [5] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, vol. 41, pp. 909-996, 1988.
- [6] O. Rioul, "On the choice of wavelet filters for still image compression," in *Proc. IEEE ICASSP 93*, Minneapolis, MN, vol. 5, 1993, p. 550.
- [7] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.
- [8] I. Daubechies and J. C. Lagarias, "Two-scale difference equations II. Local regularity, infinite products of matrices and fractals," *SIAM J. Math. Anal.*, vol. 23, no. 4, pp. 1031-1079, 1992.
- [9] O. Rioul, "Simple regularity criteria for subdivision schemes," *SIAM J. Math. Anal.*, vol. 23, no. 6, pp. 1544-1576, 1992.
- [10] L. F. Villemoes, "Continuity of nonseparable quincunx wavelets," *Appl. Comput. Harmon. Anal.*, vol. 1, pp. 180-187, 1994.
- [11] G. Karlsson and M. Vetterli, "Theory of two dimensional multirate filter banks," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, June 1990.
- [12] S. Venkataraman and B. C. Levy, "State space representations of two dimensional FIR lossless transfer matrices," *IEEE Trans. Circuits Syst.*, vol. 41, p. 117, Feb. 1994.
- [13] S. Venkataraman and B. C. Levy, "A comparison of design methods for 2-D FIR orthogonal perfect reconstruction filter banks," submitted to *IEEE Trans. Circuits Syst.*, 1994.
- [14] J. Kovačević and M. Vetterli, "Nonseparable two and three-dimensional wavelets," *IEEE Trans. Signal Processing*, vol. 43, May 1995.
- [15] E. Viscito and J. P. Allebach, "The analysis and design of multidimensional FIR perfect reconstruction filter banks for arbitrary sampling lattices," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 29-41, Jan. 1991.
- [16] W. M. Lawton and H. L. Resnikoff, "Multidimensional wavelet bases," Tech. Rep., Aware, Aware Inc., Cambridge MA, 1991.
- [17] P. N. Heller, H. L. Resnikoff, and R. O. Wells, "Wavelet matrices and the representation of discrete functions," in *Wavelets: A Tutorial in Theory and Applications* (C. K. Chui, Ed.). New York: Academic, 1992.
- [18] P. P. Vaidyanathan, T. Q. Nguyen, Z. Doganata, and T. Saramaki, "Improved technique for design of perfect reconstruction FIR QMF banks with lossless polyphase matrices," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, July 1989.
- [19] V. C. Liu and P. P. Vaidyanathan, "On factorization of a subclass of 2-D digital FIR lossless matrices for 2-D QMF bank applications," *IEEE Trans. Circuits Syst.*, vol. 37, June 1990.
- [20] D. Stanhill and Y. Y. Zeevi, "Separable 2-D lossless filter-banks as a special case of factorable filter-banks," submitted to *IEEE Trans. Circuits Syst.*, 1996.
- [21] P. Steffen *et al.*, "Theory of regular m-band wavelet bases," Tech. Rep. TR-93-02, Computational Math. Lab., Rice Univ., Houston, TX, 1993.
- [22] A. Cavaretta, W. Dahmen, and C. Micchelli, "Subdivision for Computer Aided Geometric Design," *Memoirs Amer. Math. Soc.*, vol. 93, 1991.
- [23] K. Gröchenig and W. R. Madych, "Multiresolution analysis, Haar bases, and self-similar tilings of R^n ," *IEEE Trans. Inform. Theory*, vol. 38, Mar. 1992.
- [24] A. Cohen and I. Daubechies, "Non-separable bidimensional wavelet bases," *Rev. Mat. Iberoamericana*, vol. 9, no. 1, pp. 51-137, 1993.
- [25] G. G. Walter, *Wavelets and Other Orthogonal Systems with Applications*. Boca Raton, FL: CRC, 1994.
- [26] W. H. Press *et al.*, *Numerical Recipes in C*. Cambridge, UK: Cambridge University Press, 1992.



David Stanhill was born on August 20, 1963 in Israel. He received the B.Sc degree in physics and computer science and M.Sc in physics from the Hebrew University, Jerusalem, Israel, in 1988 and 1991 respectively. He is currently pursuing the Ph.D. degree in electrical engineering at the Technion, Haifa, Israel. His research interests include wavelet theory and image processing.

Yehoshua Y. Zeevi is a graduate of the Technion—Israel Institute of Technology, Haifa. He received the M.Sc. degree from the University of Rochester, Rochester, NY, and the Ph.D. degree from the University of California at Berkeley.

Currently, he is the Barbara and Norman Seiden Professor of Computer Sciences and the Dean of the Faculty of Electrical Engineering at the Technion. He was a Vinton Hayes Fellow at Harvard University, Cambridge, MA, and has been a regular visitor there. He was also a Visiting Professor at the Massachusetts Institute of Technology, Cambridge, MA, and at the CAIP Center at Rutgers University, Piscataway, NJ. His major research is devoted to biological and computer vision, visual communication, and image representation. He is also one of the founders of i Sight, Inc., which is a company devoted to digital video cameras that mimic visual systems and real-time image processing.

Dr. Zeevi is Editor-in-Chief of the *Journal of Visual Communication and Image Representation* and a member of the editorial boards of several journals.