Optimal Filter Banks for Signal Reconstruction from Noisy Subband Components

Anastasios N. Delopoulos, *Member, IEEE,* and Stefanos D. Kollias, *Member, IEEE*

Abstract— **Conventional design techniques for analysis and synthesis filters in subband processing applications guarantee perfect reconstruction of the original signal from its subband components. The resulting filters, however, lose their optimality when additive noise due, for example, to signal quantization, disturbs the subband sequences. In this paper, we propose filter design techniques that minimize the reconstruction mean squared error (MSE) taking into account the second order statistics of signals and noise in the case of either stochastic or deterministic signals. A novel recursive, pseudo-adaptive algorithm is proposed for efficient design of these filters. Analysis and derivations are extended to 2-D signals and filters using powerful Kronecker product notation. A prototype application of the proposed ideas in subband coding is presented. Simulations illustrate the superior performance of the proposed filter banks versus conventional perfect reconstruction filters in the presence of additive subband noise.**

I. Introduction

MULTIRATE signal processing has recently gained great popularity in applications such as digital audio/speech coding, progressive image coding, spectrum analysis, and time-varying system identification. In conjunction with its analog counterpart, namely time-scale transformations, subband processing has become a topic of extensive research. The core idea is to divide the signal domain into complementary frequency bands and separately process each band.

The use of decimating filters to produce the subband signals is accompanied by appropriate interpolators that resynthesize the original signal. Extensive research has been carried out on the design of decimators/interpolators that allow perfect reconstruction [6], [22]. Efficient implementation schemes that use the so-called polyphase decomposition structures are employed to reduce the complexity of decimators/interpolators to levels as low as FFT.

When perfect reconstruction is the issue, the design of decimation/interpolation filters encounters removal of both phase and amplitude distortions. The associated analysis is based on the assumption that all subband signals are available to the interpolation bank with infinite precision. Thus, perfect reconstruction filter banks can be thought of as a combination of filters with all-pass and linear phase overall transfer function. In most practical applications, though, infinite accuracy is not possible. Subband signals are quantized and/or corrupted by external disturbances. Typical examples are subband coding applications, where quantization of subband signals is necessary due to limitations in the bandwidth of transmission channels or in the space of storage media; consequently, reconstruction is performed on the basis of subband signals that include quantization error (see, e.g., [24] for a description of the influence of subband quantization error to the reconstructed signals.) In addition, external disturbances of the transmission channels can be modeled as an extra subband noise term; the compensation of noise in problems of data fusion available in various resolutions can also be modeled in the same framework. It should also be mentioned that quantization noise may have time-varying spectral characteristics, especially in variable bit rate situations that require on-line switching between modes with different number of quantization levels. The goal of the present work is to propose methods for designing interpolation filters that suppress the effects of additive subband noise. This is possible by adjusting the impulse responses of the interpolating filters to the particular autocovariance structures of the original signal and of the additive noise. Recent publications (see, e.g., [9], [11], and [18]) examine the design of appropriate compensators that, preceding the synthesis stage, suppress the effects of quantization noise; the noise is assumed to obey the gain plus noise model that is suitable for optimal Lloyd–Max quantizers.

The method that we propose here compensates quantization—or any other type—of noise, by appropriately adjusting the synthesis filter bank, i.e., the transfer functions of the interpolators. In general, no assumption regarding the nature of the noise is necessary; colored noise even correlated to the input signal can be easily handled. The latter is of particular importance in coarse quantization scenarios where the resulting noise is highly correlated to the subband signals (see, e.g., [1], [10], and [24]).

Our approach is developed purely in the time domain and essentially uses the mean squared error (MSE) criterion as a measure of similarity between the original signal and the reconstructed version of it. Both stochastic and deterministic inputs are considered. The proposed design algorithm is time adaptive, being able to track signals with slowly time-varying spectral characteristics.

In Section II, we cover some preliminary definitions and tools that are used in multirate signal processing. We quote the standard notation for decimation and interpolation, and briefly refer to the necessary and sufficient conditions that establish perfect reconstruction. The notation covers the general case,

Manuscript received March 9, 1994; revised July 28, 1995. The associate editor coordinating the review of this paper and approving it for publication was Prof. Roberto H. Bamberger.

The authors are with Computer Science Division, National Technical University of Athens, Athens, Greece (email: stefanos@cs.ntua.gre).

Publisher Item Identifier S 1053-587X(96)01634-9.

1053–587X/96$05.00 © 1996 IEEE

for arbitrary $P \geq 2$ subbands; for convenience, however, in the sections that follow we adopt the simplest paradigm of $P = 2$. In order to derive the optimal filter design analysis of the next section, we present the relation between decimators/interpolators with: i) matrix operations (refer to [19] for a similar approach); and ii) multichannel filters (see, e.g., [6], [21], and [22]).

In Section III the matrix and multichannel approaches introduced in Section II are used to derive optimal reconstruction in the presence of noise. The matrix approach that is considered first yields the overall optimal synthesis filters at the cost of solving a set of linear equations of order equal to the length of the available data record. The resulting optimal interpolating filters are, in general, time-varying and not of prespecified length; they appear as rows of a square matrix.

On the other hand, multichannel formulation yields optimal filters within the class of linear time-invariant FIR filters of prespecified length. The set of linear equations to be solved in this case is of order equal to the size of the filters rather than the size of the data record. In addition, we show that the resulting system can be solved recursively in time by means of a recursive least squares (RLS) type of algorithm. Comparison between the two approaches is included in the simulation Section VI.

By the end of Section III, the optimal filter bank expressions are revisited for the special cases as follows: i) additive noise uncorrelated to the input signals and ii) quantization noise produced by optimal Lloyd–Max quantizers.

Section IV extends the methodology proposed in Section III to 2-D signal processing. Separable 2-D impulse responses are encountered, which lead to optimal filters of the same structure as in the 1-D case. Effective Kronecker product notation and properties are used in order to handle the 2-D problem via 1-D optimization tools.

Section V refers to real-world applications of the proposed design techniques. We mainly consider the variable bit rate scenario in subband coding of 1-D and 2-D data, making the assumption that quantization error can be modeled as second-order stationary additive.

Section VI contains simulation experiments that illustrate the advantages of the proposed filters versus the fixed perfect reconstruction filters that are currently used. Simulations are performed on synthetic as well as on real 1-D and 2-D image data.

II. Decimation and Interpolation: Definition and Representations

Subband processing and related applications assume a digitally implementable mechanism that decomposes a given signal, $x(n)$, into, for example, P sequences, where $P \geq 2$, each containing different frequency information of the original signal. In addition, an inverse mechanism that perfectly resynthesizes $x(t)$ from its subband components is required. Forward signal decomposition, called decimation, is implemented using digital filters, e.g., $h_0, h_1, \cdots, h_{P-1}$ with transfer functions centered at different frequencies. The ith subband component is formed by passing $x(n)$ through $h_i(n)$ and subsampling by P. The decimation procedure is therefore described as

$$y_i(n) = \sum_k h_i(Pn - k)x(k). \tag{2.1}$$

Subsampling in (2.1) is accomplished by multiplying n by P in the RHS of the expression. The operation of (2.1) is linear but not time invariant. A slight modification of the above expression turns it into a sum of linear and time-invariant filtering operations, which involves appropriately defined subsequences of $h_i(n)$ and $x(n)$, namely their polyphase components as follows:

$$y_i(n) = \sum_{l=0}^{P-1} \sum_k h_{il}(n - k)x_l(k) \tag{2.2}$$

where $h_{il}(n)$ and $x_l(n)$ are the polyphase components of $h_i(n)$ and $x(n)$, respectively, defined as

$$h_{il}(n) \triangleq h_i(Pn + l) \quad l = 0, \cdots, P - 1 \tag{2.3a}$$

and

$$x_l(n) \triangleq x(Pn - l) \quad l = 0, \cdots, P - 1. \tag{2.3b}$$

Decimation itself decomposes the original signal in subband sequences that can be used in various signal processing tasks such as pattern recognition [15], [20], modeling [4], [5], identification [2], [5], and image analysis [15]. In most of these applications and especially in digital communications applications [22], reconstruction of the original signal should be possible through the inverse procedure, called interpolation.

Interpolation is performed by first inserting $P - 1$ zeros between successive samples of the subband signals and then passing the resulting sequences through P filters $f_i(n)$, $i = 0, \cdots, P - 1$; finally, the outputs of all filters are summed up to form an approximation $\hat{x}(n)$ of $x(n)$

$$\hat{x}(n) = \sum_{i=0}^{P-1} \sum_k f_i(n - Pk)y_i(k). \tag{2.4}$$

It turns out (see, e.g., [14], [22]) that the use of appropriate combinations of decimation and interpolation filter banks guarantees perfect reconstruction, i.e., $\hat{x}(n) = x(n)$.

For convenience in the following analysis we shall consider only the $P = 2$ case; all derivations and assertions can be easily generalized to any number of bands at the cost of notational complexity. By the end of this section we shall translate (2.1) and (2.4) first into a pair of matrix operations and next into a couple of multichannel filters, the latter based on the polyphase notation.

A. Matrix Representation of Decimators/Interpolators

The computations performed in (2.1) and (2.4) are of convolution type; thus, they can be written in matrix form. For a finite impulse response (FIR) digital filter $h(n)$ and for $P = 2$, let us define the matrix

$$H_i(k, l) \triangleq h_i(2k - l), \quad k = 0, \cdots, \frac{N}{2} - 1, \quad l = 0, \cdots, N - 1 \tag{2.5}$$

where N is the length of the input signal $x(n)$. Using (2.5), decimation can be written as

$$\mathbf{y}_i = H_i \mathbf{x} \tag{2.6}$$

where the elements of matrix H_i correspond to the impulse response coefficients $h_i(n)$ according to (2.5), i.e., $\mathbf{x} \triangleq [x(0), \cdots, x(N-1)]^T$ and $\mathbf{y}_i \triangleq \{y_i(0), \cdots, y_i[(N/2)-1]\}^T$. Equation (2.6) is an exact duplicate of (2.1) except on the boundaries of vector $\mathbf{y}_i$; these boundary effects are, however, of decreasing importance for large values of N.

Let matrix F_i be formed similarly to (2.5), based on the ith interpolation filter $f_i(n)$; then the output of the corresponding interpolator is given by $F_i^T \mathbf{y}_i$ and the reconstructed signal vector is

$$\begin{aligned}\hat{\mathbf{x}} &= \sum_{i=0}^{1} F_i^T \mathbf{y}_i \\ &= [F_0^T, F_1^T] \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \end{bmatrix}. \end{aligned} \tag{2.7}$$

Let us define vector $\mathbf{y}$ as the concatenation of the 2 subband vectors $\mathbf{y}_i$, i.e., $\mathbf{y} \triangleq [\mathbf{y}_0^T, \mathbf{y}_1^T]^T$ and form the augmented matrices

$$H = \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \tag{2.8a}$$

and

$$F = \begin{bmatrix} F_0 \\ F_1 \end{bmatrix}. \tag{2.8b}$$

Then, decimation is described as follows:

$$\mathbf{y} = H\mathbf{x} \tag{2.9a}$$

and interpolation by

$$\hat{\mathbf{x}} = F^T \mathbf{y}. \tag{2.9b}$$

Equations (2.9a) and (2.9b) represent the decimation and interpolation procedures in the form of vector space projections. A similar approach along with properties of matrices H and F can be found in [19], where infinite size matrices are considered in order to overcome boundary effects. In the following we prefer, however, to restrict our approach to the finite matrix formulation since, as $N \to \infty$, boundary effects introduce a finite error that does not affect the validity of the optimal design analysis presented in this paper.

It should be mentioned that the perfect reconstruction property of the filter banks described by (2.9a) and (2.9b) is very conveniently expressed as

$$F^T H = I_N,$$

using matrix formulation, where I_N stands for the $N \times N$ identity matrix.

When additive noise disturbs the subband signals the input to the interpolation filter bank can be written in vector form as $\mathbf{z} = \mathbf{y} + \mathbf{v}$, where the $N \times 1$ vector $\mathbf{v}$ is a concatenation of P noise components each of length N/P. The overall effect of decimation/interpolation is then described as

$$\hat{\mathbf{x}} = F^T \mathbf{z} = F^T H \mathbf{x} + F^T \mathbf{v}.$$

B. Multichannel Representation of Decimators/Interpolators

In the previous paragraph, interpolation and decimation were formulated in a matrix/vector multiplication representation. Alternatively, decimating and interpolating filter banks can be shown (see, e.g., [1], [16]) to correspond to multichannel filters. Namely, (2.1) and (2.4) are equivalent to

$$\mathbf{y}(n) = \sum_k H(k)\mathbf{x}(n-k) \tag{2.10a}$$

and

$$\hat{\mathbf{x}}(n) = \sum_k F(k)^T \mathbf{y}(n-k) \tag{2.10b}$$

respectively, where,

$$\begin{aligned} \mathbf{x}(n) &\triangleq \begin{bmatrix} x(2n) \\ x(2n-1) \end{bmatrix} \\ \hat{\mathbf{x}}(n) &\triangleq \begin{bmatrix} \hat{x}(2n) \\ \hat{x}(2n-1) \end{bmatrix} \\ \mathbf{y}(n) &\triangleq \begin{bmatrix} y_0(n) \\ y_1(n) \end{bmatrix} \end{aligned} \tag{2.11}$$

and $H(k)$, $F(k)$ are multichannel impulse response matrices of the form

$$\begin{aligned} H(k) &\triangleq \begin{bmatrix} h_{00}(k) & h_{01}(k) \\ h_{10}(k) & h_{11}(k) \end{bmatrix} \\ h_{i0}(k) &\triangleq h_i(2k) \\ h_{i1}(k) &\triangleq h_i(2k+1), \qquad i = 0, 1 \end{aligned} \tag{2.12a}$$

and

$$\begin{aligned} F(k) &\triangleq \begin{bmatrix} f_{00}(k) & f_{01}(k) \\ f_{10}(k) & f_{11}(k) \end{bmatrix} \\ f_{i0}(k) &\triangleq f_i(2k) \\ f_{i1}(k) &\triangleq f_i(2k-1), \qquad i = 0, 1. \end{aligned} \tag{2.12b}$$

The advantage of the above multichannel formulation is that it can lead to the derivation of optimal FIR interpolators of prespecified length. Let $h_i(n)$ be FIR filters. Then (2.10a) yields

$$\begin{aligned} \mathbf{y}(n) &= \sum_{k=0}^{p} H(k)\mathbf{x}(n-k) \\ &= H_a \mathbf{x}_a(n) \end{aligned} \tag{2.13}$$

where H_a and $\mathbf{x}_a(n)$ are augmented versions of $H(k)$, $\mathbf{x}(k)$.

$$H_a \triangleq [H(0) \ \cdots \ H(p)] \tag{2.14a}$$

is of size $2 \times 2(p+1)$ and

$$\begin{aligned} \mathbf{x}_a(n) &\triangleq \begin{bmatrix} \mathbf{x}(n) \\ \vdots \\ \mathbf{x}(n-p) \end{bmatrix} \\ &\equiv \begin{bmatrix} x(2n) \\ x(2n-1) \\ \vdots \\ x(2n-2p) \\ x(2n-2p-1) \end{bmatrix} \end{aligned} \tag{2.14b}$$

is of length $2(p+1)$. On the other hand, assuming that $f_i(k)$ are also FIR filters, (2.10b) yields

$$\begin{aligned}\hat{\mathbf{x}}(n) &= \sum_{k=0}^{q} F(k)^T \mathbf{y}(n-k) \\ &= F_a^T \mathbf{y}_a(n),\end{aligned} \tag{2.15}$$

where

$$F_a \triangleq \begin{bmatrix} F(0) \\ \vdots \\ F(q) \end{bmatrix} \tag{2.16a}$$

is a $2(q+1) \times 2$ matrix and

$$\begin{aligned}\mathbf{y}_a(n) &\triangleq \begin{bmatrix} \mathbf{y}(n) \\ \vdots \\ \mathbf{y}(n-q) \end{bmatrix} \\ &\equiv \begin{bmatrix} y_0(n) \\ y_1(n) \\ \vdots \\ y_0(n-q) \\ y_1(n-q) \end{bmatrix}\end{aligned} \tag{2.16b}$$

is a vector of length $2(q+1)$. Combining (2.13) and (2.15), we obtain

$$\begin{aligned}\mathbf{y}_a(n) &= \begin{bmatrix} H_a \mathbf{x}_a(n) \\ \vdots \\ H_a \mathbf{x}_a(n-q) \end{bmatrix} \\ &= \mathcal{H}\mathcal{X}(n)\end{aligned} \tag{2.17a}$$

where the $2(q+1) \times 2(p+q+1)$ matrix $\mathcal{H}$ is formed by augmenting shifts of H_a, i.e.

$$\mathcal{H} \triangleq \begin{bmatrix} H_a & & & \cdot & \cdot & \cdot \\ \mathbf{0}_2 & H_a & & \cdot & \cdot & \cdot \\ \mathbf{0}_2 & \mathbf{0}_2 & H_a & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & H_a \end{bmatrix} \tag{2.17b}$$

and the $2(p+q+1)$-dimensional vector $\mathcal{X}(n)$ is defined as

$$\begin{aligned}\mathcal{X}(n) &\triangleq \begin{bmatrix} x(2n) \\ x(2n-1) \\ \vdots \\ x(2n-2q-2p) \\ x(2n-2q-2p-1) \end{bmatrix} \\ &\equiv \begin{bmatrix} \mathbf{x}(n) \\ \vdots \\ \mathbf{x}(n-p-q) \end{bmatrix}\end{aligned}$$

in the definition of $\mathcal{H}$, $\mathbf{0}_2$ is the 2×2 null matrix. Inserting (2.17a) into (2.15), we get an expression of the total effect of decimation/interpolation

$$\hat{\mathbf{x}}(n) = F_a^T \mathcal{H}\mathcal{X}(n). \tag{2.18}$$

Equation (2.18) holds provided that all subband signals are noise free [c.f., (2.15)]. If additive noise components are

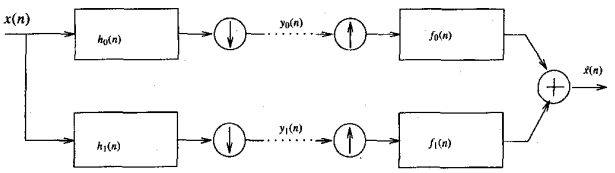


Fig. 1. Noise-free decimation interpolation with $P = 2$.

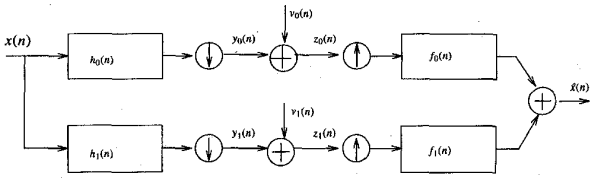


Fig. 2. Decimation interpolation with additive subband noise.

included in the subband signals, the interpolator of (2.15) accepts as input some

$$\mathbf{z}(n) = \mathbf{y}(n) + \mathbf{v}(n) \tag{2.19}$$

where $\mathbf{v}(n) = [v_0(n) \quad v1(n)]^T$ includes the subband noise disturbances. In the form of (2.16b)

$$\mathbf{z}_a(n) = \mathbf{y}_a(n) + \mathbf{v}_a(n). \tag{2.20}$$

In this case, the RHS of (2.18) should be modified to include a term that is due to the noise, as follows:

$$\begin{aligned}\hat{\mathbf{x}}(n) &= F_a^T \mathbf{z}_a(n) \\ &= F_a^T \mathcal{H}\mathcal{X}(n) + F_a^T \mathbf{v}_a(n).\end{aligned} \tag{2.21}$$

III. Optimal Interpolators

Perfect reconstruction filters, as presented in Section II, guarantee that $\hat{x}(n) = x(n)$ provided that all subband signals are exactly computed and used in the synthesis filter bank. In practice, however, there are a variety of reasons for deviations from this assumption. In [22], for example, the effect of finite precision arithmetic was examined as an error source, and optimal bit assignment for the filter taps was proposed as a means to reduce distortions of this origin. Still, though, subband noise is always added to $y_i(n)$ due to necessary quantization of the transmitted signals. On top of that, external additive noise can model transmission imperfections. In the following discussions all kinds of noise are assumed to be stationary. No assumptions regarding the spectrum of the noise, nor intersubband independence, are due. The general model that we consider is depicted in Fig. 2, where $\{v_i(n)\}_{i=0,1}$ are the additive noise components and

$$z_i(n) = y_i(n) + v_i(n)$$

are the noisy versions of the subband signals that are fed to the interpolation bank.

If $\{H_i(z), F_i(z)\}$ consist of an analysis/synthesis structure and due to linearity, the reconstruction error will be

$$\begin{aligned}\boldsymbol{\varepsilon} &\triangleq \mathbf{x} - \hat{\mathbf{x}} \\ &= \mathbf{x} - F^T \mathbf{z} \\ &= (I - F^T H)\mathbf{x} - F^T \mathbf{v}\end{aligned} \tag{3.1}$$

in matrix representation and

$$\begin{aligned}\varepsilon(n) &\triangleq \mathbf{x}(n) - \hat{\mathbf{x}}(n) \\ &= \mathbf{x}(n) - F_a^T \mathbf{z}_a(n) \\ &= \mathbf{x}(n) - F_a^T \mathcal{H}\mathcal{X}(n) - F_a^T \mathbf{v}_a(n) \end{aligned} \tag{3.2}$$

in multichannel formulation.

In the following, for a given analysis bank $\{H_i(z)\}_{i=0,1}$, we seek the best synthesis bank $\{F_i(z)\}_{i=0,1}$ that minimizes the power of the error signal $\varepsilon(n)$. Both formulations can be used in order to establish optimality. It is shown that in the matrix formulation, optimal design reduces to a vector space projection, while in the multichannel formulation optimality is achieved using standard Wiener filtering methods. In both cases, knowledge of second-order statistics of the signal and the noise are necessary. Our approach is general enough in order to include stationary and nonstationary input signals, the latter including deterministic sequences as well. In any case, the noise processes are stochastic.

A. Matrix Formulation

In this formulation, we adopt the Euclidean norm of the $N \times 1$ vector ε as reconstruction error measure, namely

$$\begin{aligned} J &\triangleq E\{\varepsilon^T \varepsilon\} \\ &\equiv \operatorname{tr} E\{\varepsilon\varepsilon^T\}. \end{aligned} \tag{3.3}$$

Substituting (3.1) in (3.3), we get a least-squares minimization problem that accepts an optimal solution w.r.t. F given by (see, e.g., [13])

$$F_{opt} = R_{zz}^{-1} R_{zx}. \tag{3.4}$$

where hereafter $Rabi$ denotes the (cross-) covariance matrix $E\{ab^T\}$ of the column vectors a and b.

The error covariance is then given by

$$\begin{aligned} C &= E\{\varepsilon\varepsilon^T\} \\ &= R_{xx} - R_{xz} R_{zz}^{-1} R_{zx} \end{aligned} \tag{3.5a}$$

and the corresponding minimum reconstruction MSE is

$$J_{\min} = \operatorname{tr} C. \tag{3.5b}$$

It can be proved that expression (3.4) reduces to the conventional synthesis filter bank $(H^T)^{-1}$ in the absence of subband noise.

Expression (3.4) for F_{opt} was derived without imposing any constraints on the structure of F. As a consequence, it does not have the structure of (2.8b) in which F_0, F_1 are of the form described by (2.5). This implies that the overall optimal (in the mean-square sense) reconstruction, achieved by linear projections of $\mathbf{z}$, is not in general possible using time invariant interpolators.

Summarizing, optimal filter selection based on the preceding matrix formulation has the advantage of yielding the overall optimal interpolating filter banks among the class of linear projection operators from the subband to the original signal domain. Unfortunately, this major property is accompanied by two discouraging disadvantages:

1) The size of the linear set of equations that yield these optimal filters is of order N, namely, the size of the signal. This implies that, unless appropriate segmentation of the original signal is used, the problem is intractable.
2) The length of the resulting FIR interpolators is not possible to be specified.
3) The interpolating filters are in general time varying.

In the sequel, we turn to the multichannel approach that yields optimal filters in the class of time-invariant ones with prespecified length. The derivations that follow show that these filters can be computed by solving a set of linear equations of order equal to the size of the FIR filters used for decimation/interpolation.

B. Multichannel Formulation

In this formulation, optimality of the interpolators is equivalent to the minimization of the distance between the reconstructed$\hat{\mathbf{x}}(n)$ and the original $\mathbf{x}(n)$ vector sequences. As a measure of the interpolation accuracy we pick the norm, as follows:

$$\begin{aligned} J_F &= \overline{E}\{[\mathbf{x}(n) - \hat{\mathbf{x}}(n)]^T [\mathbf{x}(n) - \hat{\mathbf{x}}(n)]\} \\ &= \operatorname{tr} \overline{E}\{[\mathbf{x}(n) - \hat{\mathbf{x}}(n)][\mathbf{x}(n) - \hat{\mathbf{x}}(n)]^T\} \\ &= \operatorname{tr} (\overline{E}\{\mathbf{x}(n)\mathbf{x}(n)^T\} \\ &\quad + F_a^T \overline{E}\{\mathbf{z}_a(n)\mathbf{z}_a(n)^T\} F_a \\ &\quad + F_a^T \overline{E}\{\mathbf{z}_a(n)\mathbf{x}(n)^T\} - \overline{E}\{\mathbf{x}(n)\mathbf{z}_a(n)^T\} F_a). \end{aligned} \tag{3.6}$$

The $\overline{E}\{.\}$ operator is defined (see, e.g., [12]) as

$$\overline{E}\{x(n)\} \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} E\{x(n)\}$$

in order to take into account nonstationary input signals. In all cases we make the assumption that for all signals under consideration the above limit exists and is finite; this is true under common mixing conditions regarding the moments up to order four of the involved signals. It should be mentioned that when we consider signal sequences of finite length, the limit in the definition of $\overline{E}\{.\}$ should be replaced by summation over all available data samples.

For fixed decimators, minimization of J_F w.r.t. F_a is accomplished setting

$$\frac{\partial J_F}{\partial F_a} = \mathbf{0}$$

which yields the following system of normal equations:

$$R_{z_a z_a} F_a = R_{z_a x} \tag{3.7}$$

where

$$R_{z_a z_a} \triangleq \overline{E}\{\mathbf{z}_a(n)\mathbf{z}_a(n)^T\} \tag{3.8a}$$

is the autocovariance matrix of the noisy subband signals $\mathbf{z}_a(n)$ of order $2(q+1)$ and

$$R_{z_a x} \triangleq \overline{E}\{\mathbf{z}_a(n)\mathbf{x}(n)^T\} \tag{3.8b}$$

is the cross-covariance matrix of the noisy subband vectors $\mathbf{z}_a(n)$ and the input vector $\mathbf{x}(n)$ of order $2(q+1)$ by 2.

Using simple algebraic manipulations one can prove that in the absence of additive subband noise the conventional perfect reconstruction synthesis filters obey (3.7).

Both expressions (3.4) and (3.7) for the computation of optimal reconstruction filters include the (cross-) covariance matrices of the input signal and the noisy subbands. Hence, estimation of the necessary covariance lags is required in some stage of the filter design procedure. In Section V we describe a generic application of the above ideas to subband coding that solves the problem of estimating R_{zz} and R_{zx} matrices at the transmitter's side.

C. Recursive Implementation

The normal equations that yield optimal interpolators involve the theoretical values of the covariance structures of the signals. In practice, we replace the expectations by summations over the available data record; consistency of the resulting parameter estimators can be established along the lines of [12]. Given a data record of length N, estimates of $R_{z_a z_a}$ and $R_{z_a x}$ can be obtained in a batch mode and then (3.4) or (3.7) are solved by replacing $R_{z_a z_a}$ and $R_{z_a x}$ by their estimates.

Alternatively, time-recursive algorithms can be employed for solving these equations in a computationally efficient manner; moreover, by inserting an appropriate forgetting factor in the recursive definition of the auto- (cross-) covariance estimators, we make the resulting parameter estimators able to track slow variations of the signals' spectral characteristics. In the sequel, we concentrate to the multichannel formulation and correspondingly to (3.7) in order to illustrate one of the possible recursive schemes that can be used.

Let

$$\Phi_N = \frac{1}{N}\sum_{n=0}^{N-1} \mathbf{z}_a(n)\mathbf{z}_a(n)^T,$$
$$\phi_N = \frac{1}{N}\sum_{n=0}^{N-1} \mathbf{z}_a(n)\mathbf{x}(n)^T \tag{3.9a}$$

be the estimators of the covariance matrices $R_{z_a z_a}$ and $R_{z_a x}$ or, in a recursive form

$$\Phi_N = \lambda\Phi_{N-1} + \mathbf{z}_a(N)\mathbf{z}_a(N)^T,$$
$$\phi_N = \lambda\phi_{N-1} + \mathbf{z}_a(N)\mathbf{x}(N)^T \tag{3.9b}$$

where the forgetting factor $\lambda < 1$ is a value, usually close to one, that tunes the memory of the estimation procedure; subscript N denotes the time at which Φ_N and ϕ_N were estimated. Using the *matrix inversion lemma* and following the derivation of RLS algorithm (see, e.g., [12], pp. 306–307), we obtain the RLS type of algorithm of Table I, where $P_N \triangleq \Phi_N^{-1}$, $K(N)$ is the so-called gain matrix and F_N is the optimal filter at time point N. The above RLS algorithm yields, for $\lambda = 1$, exactly the same solution with the batch algorithm, while for $\lambda < 1$ it behaves as a pseudo-adaptive filter parameter estimation scheme.

TABLE I
RECURSIVE IMPLEMENTATION OF OPTIMAL SYNTHESIS FILTER DESIGN

Proposed Recursive Algorithm	
Update Procedure:	
I	$K(N) = \dfrac{P_{N-1}\,\mathbf{z}_a(N)}{\lambda + \mathbf{z}_a(N)^T\,P_{N-1}\,\mathbf{z}_a(N)}$
II	$F_N = F_{N-1} + K(N)\,[\,\mathbf{x}(N)^T - \mathbf{z}_a(N)^T\,F_{N-1}\,]$
III	$P_N = \dfrac{1}{\lambda}\left[\,P_{N-1} - \dfrac{P_{N-1}\,\mathbf{z}_a(N)\,\mathbf{z}_a(N)^T\,P_{N-1}}{\lambda + \mathbf{z}_a(N)^T\,P_{N-1}\,\mathbf{z}_a(N)}\right]$
Initialization: Batch estimation of $P_0 = \Phi_0^{-1}$	

Special Cases of Additive Noise: The general optimal solutions of (3.4) or (3.7) can be straightforwardly rewritten for two special scenarios appearing in the literature, as follows:

1) *Noise Uncorrelated to the Input Signal:*
In the *matrix approach* since $E\{\mathbf{x}\mathbf{v}^T\} \equiv \mathbf{0}$ and using (2.9b) we get $R_{zz} = HR_{xx}H^T + R_{vv}$ and $R_{zx} = HR_{xx}$. Hence, (3.4) can be rewritten as

$$F_{opt} = (HR_{xx}H^T + R_{vv})^{-1}HR_{xx}. \tag{3.10}$$

Similarly, in the *multichannel approach* using (2.17a), after some manipulation we conclude to

$$F_{opt} = (\mathcal{H}R_{\mathcal{X}\mathcal{X}}\mathcal{H}^T + R_{v_a v_a})^{-1}\mathcal{H}R_{\mathcal{X}\mathbf{x}}. \tag{3.11}$$

2) *Lloyd–Max Quantizers—The Gain-Plus-Noise Model:*
When quantization of the subband signals is performed following the Lloyd–Max algorithm (see, e.g., [8], Section 6.2), the distance between $y_i(n)$ and $z_i(n) = y_i(n) + v_i(n)$ is minimized. It turns out that under this condition the quantizers are, in general, nonuniform, and the quantization noise $v_i(n)$ is correlated to $y_i(n)$ but orthogonal to $z_i(n)$. It can be also proved that a gain-plus-noise equivalent model,

$$z_i(n) = \alpha_i y_i(n) + r_i(n), \quad i = 1, 2, \tag{3.12}$$

exists where

$$\alpha_i \equiv 1 - \frac{E\{v_i(n)^2\}}{E\{y_i(n)^2\}}, \quad E\{r_i(n)\} = 0$$
$$\text{and} \quad E\{y_i(n)r_i(n)\} = 0.$$

Optimization in subband reconstruction for such types of quantization noise have been treated in the literature by inserting appropriate scalar compensators before the synthesis band (see, e.g., [9], [11], and [18]). In the present work, we incorporate the compensator in the filters of the synthesis bank by adjusting their impulse responses to the spectral characteristics of signals and noise.

In matrix notation, let $A \triangleq \left[\begin{smallmatrix}\alpha_0 & 0\\ 0 & \alpha_1\end{smallmatrix}\right]$ and $\mathbf{A} \triangleq A \otimes I_{N/2}$. Then, the gain-plus-noise model along with (2.9a) yield

$$\mathbf{z} = \mathbf{A}\mathbf{y} + \mathbf{r} = \mathbf{A}H\mathbf{x} + \mathbf{r} \tag{3.13}$$

where $\mathbf{r}$ is defined similarly to $\mathbf{z}$, $\mathbf{y}$, and $\mathbf{v}$. Consequently, we have $R_{zz} = \mathbf{A}HR_{xx}H^T\mathbf{A}^T + R_{rr}$ and $R_{zx} = \mathbf{A}HR_{xx}$. Hence, (3.4) rewrites as

$$F_{opt} = (\mathbf{A}HR_{xx}H^T\mathbf{A}^T + R_{rr})^{-1}\mathbf{A}HR_{xx}. \quad (3.14)$$

In multichannel notation

$$\mathbf{z}_a(n) = \mathcal{A}\mathbf{y}_a(n) + \mathbf{r}_a(n) \quad (3.15)$$

where $\mathbf{r}_a(n)$ is defined from $r_i(n)$ in a manner similar to $\mathbf{v}_a(n)$ and $\mathcal{A} \triangleq I_q \otimes A$. Plugging the above expression to the definitions of $R_{z_az_a}$ and R_{z_ax} and using (2.17a), we get

$$F_{opt} = (\mathcal{A}\mathcal{H}R_{\mathcal{X}\mathcal{X}}\mathcal{H}^T\mathcal{A}^T + R_{r_ar_a})^{-1}\mathcal{A}\mathcal{H}R_{\mathcal{X}\mathbf{x}}. \quad (3.16)$$

IV. Extension to 2-D Signals

In the area of still and moving image compression, subband coding plays an important role. In fact, subband processing competes with transform coding, mainly discrete cosine transform (DCT)-based compression schemes, in the race to establish a standard for lossy image coding.

Pioneering work on image compression algorithms that employ separate coding of spatial frequency bands can be found in [23] and [25]. 2-D perfect reconstruction filter banks are also considered in [14]. The dominant approach is to implement 2-D decimators/interpolators using 1-D perfect reconstruction filter banks; this approach is equivalent to using 2-D filter banks with separable impulse responses. In this setup the original 2-D signal $x(m, n)$ is decimated into four subband components $y_{ij}(m, n)$ $i, j = 0, 1$ according to the following equation:

$$y_{ij}(m, n) = \sum_k \sum_l h_i^c(l)h_j^r(k)x(2m - l, 2n - k). \quad (4.1)$$

Superscripts c and r denote column- and row-wise operations, respectively. As in the previous section, we have adopted decimation by $P = 2$, which is commonly used in practice.

The reconstruction counterpart of (4.1) is given by

$$\hat{x}(m, n) = \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_k\sum_l f_i^c(m - 2l) \cdot f_j^r(n - 2k)y_{ij}(l, k) \quad (4.2)$$

which yields a reconstructed version $\hat{x}(m, n)$ of the original image $x(m, n)$. Perfect reconstruction, i.e., $\hat{x}(m, n) \equiv x(m, n)$ is guaranteed, provided that the 1-D filter banks applied successively on the columns and rows of the image separately satisfy (2.5a) and (2.5b). In particular, QMF filters (2.6a)–(2.6e) should be satisfied.

In the sequel, we assume that the analysis filter banks are of perfect reconstruction type and seek optimal synthesis filter banks in the sense of minimizing the effect of additive subband noise. This means that (4.1) is still holding true, while in (4.2) the subband components $y_{ij}(l, k)$ should be replaced by

$$z_{ij}(l, k) = y_{ij}(l, k) + v_{ij}(l, k) \quad (4.3)$$

where $v_{ij}(l, k)$ are 2-D, zero-mean, additive stationary noises. Both matrix and multichannel representations used in the previous sections are extended to the 2-D case, in order to reformulate (4.1) and (4.2) and to express the reconstruction error in a way that allows easy derivation of optimal (in the mean-squared-error sense) reconstruction filters.

A. Matrix Representation

Let

$$\mathbf{x} = [x(m, n)]_{m=0,\cdots,N-1,\ n=0,\cdots,N-1}$$

be an $N \times N$ matrix corresponding to the original image and, ignoring boundary effects, let

$$\mathbf{y}_{ij} = [y_{ij}(m, n)]_{m=0,\cdots,(N/2)-1, n=0,\cdots,(N/2)-1}, \quad i, j = 0, 1$$

be the four $N/2 \times N/2$ subband components. If the decimating filters have FIR impulse responses of order much smaller than N, equation (4.1) can be well approximated, apart from boundary inaccuracies, as

$$\mathbf{y}_{ij} = H_i^c\mathbf{x}H_j^{rT} \quad (4.4a)$$

where the $N/2 \times N$ matrices H_i^c and H_j^r are of the form of (2.5). Then, the $N \times N$ matrix that contains all four subband components is defined by,

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_{00} & \mathbf{y}_{01} \\ \mathbf{y}_{10} & \mathbf{y}_{11} \end{bmatrix} = H^c\mathbf{x}H^{rT} \quad (4.4b)$$

where $H^c \triangleq \begin{bmatrix} H_0^c \\ H_1^c \end{bmatrix}$, $H^r \triangleq \begin{bmatrix} H_0^r \\ H_1^r \end{bmatrix}$, in accordance with (2.8a).

Equation (4.4b) is the 2-D counterpart of matrix formulation describing 1-D decimation by (2.9a). Following the same notation, interpolation described by (4.2) can be written in matrix form as

$$\hat{\mathbf{x}} = F^{cT}\mathbf{y}F^r \quad (4.5)$$

where the $N \times N$ matrices F^c and F^r are generated from f_j^c and f_i^r, respectively. In this setup, perfect reconstruction requirement is equivalent to

$$F^{cT}H^c = F^{rT}H^r = I_N. \quad (4.6)$$

Expressions of (4.4b) and (4.5) can be both rewritten in a more convenient form using Kronecker product notation, namely,

$$\tilde{\mathbf{y}} = (H^r \otimes H^c)\tilde{\mathbf{x}}, \quad (4.7a)$$

and

$$\hat{\tilde{\mathbf{x}}} = (F^r \otimes F^c)^T\tilde{\mathbf{y}} \quad (4.7b)$$

respectively, where $\tilde{\mathbf{x}} \triangleq \text{vec}\{\mathbf{x}\}$, $\hat{\tilde{\mathbf{x}}} \triangleq \text{vec}\{\hat{\mathbf{x}}\}$ and $\tilde{\mathbf{y}} \triangleq \text{vec}\{\mathbf{y}\}$ are $N^2 \times 1$ column vectors. Notation $\alpha = \text{vec}\{\alpha\}$, which is frequently used in the following, means that the column vector α is formed by stacking up all columns of α, i.e., $\alpha(iN + j) = A(i, j)$. In (4.7a) and (4.7b) we used the property $\text{vec}\{ABC\} = (C^T \otimes A)\text{vec}\{B\}$ of Kronecker products (for definition and properties of Kronecker products refer, for example, to [17, pp. 237–241]).

When additive noise disturbs $y_{ij}(m, n)$, (4.7b) should be replaced by

$$\hat{\tilde{\mathbf{x}}} = (F^r \otimes F^c)^T \tilde{\mathbf{z}} \tag{4.7c}$$

where $\tilde{\mathbf{z}} = \tilde{\mathbf{y}} + \tilde{\mathbf{v}}$, using obvious notation. Formulation of (4.7a)–(4.7c) resembles the structure of (2.9a) and (2.9b). Hence, for fixed decimators H_0^c, H_0^r the optimal in the mean-squared-error sense interpolator is given by

$$\begin{aligned} F_{opt} &= F_{opt}^r \otimes F_{opt}^c \\ &= R_{\tilde{z}\tilde{z}} R_{\tilde{z}\tilde{x}} \end{aligned} \tag{4.8}$$

where $R_{\tilde{z}\tilde{z}} \triangleq \overline{E}\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T\}$ and $R_{\tilde{z}\tilde{x}} \triangleq \overline{E}\{\tilde{\mathbf{z}}\tilde{\mathbf{x}}^T\}$. Equation (4.8) yields the optimal $N^2 \times N^2$ matrix F_{opt} rather than F_{opt}^c and F_{opt}^r separately. The latter can be next computed within an unimportant scalar ambiguity from their Kronecker product exploiting the following lemma, which provides a somewhat general way for approximating 2-D nonseparable problems by separable ones:

Lemma 4.1: Let A, B be $M \times N$ matrices and $C = A \otimes B$. Define the $MN \times MN$ matrix $\tilde{C}$ via the following rearrangement of the entries of C:

$$\begin{aligned} \tilde{C}(lN + n, kN + m) \triangleq\ & C(mN + n, kN + l) \\ & k, l = 0, \cdots, N-1 \\ & m, n = 0, \cdots, M-1. \end{aligned}$$

If $\tilde{A} = \text{vec}\{A\}$ and $\tilde{B} = \text{vec}\{B\}$, then we have the following:

1) $\tilde{C} = \tilde{A}\tilde{B}^T$.
2) $\text{rank}\,\tilde{C} = 1$.
3) If s is the single nonzero singular value of $\tilde{C}$ and u, v the corresponding singular vectors, then $\tilde{A} = s_1 u$, $\tilde{B} = s_2 v$, and $s_1 s_2 = s$.

Proof: From the definition of Kronecker products it can be easily checked that the particular rearrangement $\tilde{C}$ of C satisfies 1); assertions 2) and 3) are immediate consequences of part 1).

Due to this lemma, F_{opt}^c and F_{opt}^r can be obtained from F_{opt} by first forming $\tilde{F}_{opt}$ and by then using singular value decomposition as described by the aforementioned lemma. It should be mentioned that in practice more than one singular values may be nonzero; in this situation, the maximum singular value and the corresponding singular vectors should be used. Due to SVD properties [3] the resulting F_{opt}^c and F_{opt}^r matrices will have the closest to F_{opt} Kronecker product among all matrices of the same dimension.

B. Multichannel Representation

Decimation and interpolation for 2-D signals can be transformed into a multichannel linear filtering using polyphase analysis in a manner similar to that used in the 1-D case. Decimation can be rewritten using (4.1) as follows:

$$\begin{aligned} & y_{ij}(m, n) = \\ & \sum_k \sum_l [h_i^c(2l) \quad h_i^c(2l+1)] \\ & \times \begin{bmatrix} x(2m-2l, 2n-2k) & x(2m-2l, 2n-2k-1) \\ x(2m-2l-1, 2n-2k) & x(2m-2l-1, 2n-2k-1) \end{bmatrix} \\ & \times \begin{bmatrix} h_j^r(2k) \\ h_j^r(2k+1) \end{bmatrix} \end{aligned} \tag{4.9a}$$

or

$$\begin{aligned} \mathbf{y}(m, n) &\triangleq \begin{bmatrix} y_{00}(m, n) & y_{01}(m, n) \\ y_{10}(m, n) & y_{11}(m, n) \end{bmatrix} \\ &= \sum_k \sum_l H^c(l)\mathbf{x}(m-l, n-k)H^r(k)^T \end{aligned} \tag{4.9b}$$

where $H^c(l)$ and $H^r(k)$ are defined similarly to $H(k)$ in (2.12a) and

$$\mathbf{x}(m, n) \triangleq \begin{bmatrix} x(2m, 2n) & x(2m, 2n-1) \\ x(2m-1, 2n) & x(2m-1, 2n-1) \end{bmatrix}.$$

In particular, for FIR decimators of order p_c and p_r for columns and rows, respectively, (4.9b) yields

$$\mathbf{y}(m, n) = H_a^c \mathbf{x}_a(m, n) H_a^{rT} \tag{4.10}$$

where

$$\begin{aligned} H_a^c &\triangleq [H^c(0) \quad H^c(1) \quad \cdots \quad H^c(p_c)], \\ H_a^r &\triangleq [H^r(0) \quad H^r(1) \quad \cdots \quad H^r(p_r)] \end{aligned} \tag{4.11a}$$

and

$$\begin{aligned} & \mathbf{x}_a(m, n) \triangleq \\ & \begin{bmatrix} \mathbf{x}(m, n) & \cdots & \mathbf{x}(m, n-p_r) \\ \mathbf{x}(m-1, n) & \cdots & \mathbf{x}(m-1, n-p_r) \\ \vdots & & \vdots \\ \mathbf{x}(m-p_c, n) & \cdots & \mathbf{x}(m-p_c, n-p_r) \end{bmatrix} \end{aligned} \tag{4.11b}$$

is a $2(p_c+1) \times 2(p_r+1)$ sliding window of the input image $x(m, n)$.

Interpolation by separable 2-D filters can be expressed using (4.2) and grouping together appropriate output values, as follows:

$$\begin{aligned} \hat{\mathbf{x}}(m, n) &\triangleq \begin{bmatrix} \hat{x}(2m, 2n) & \hat{x}(2m, 2n-1) \\ \hat{x}(2m-1, 2n) & \hat{x}(2m-1, 2n-1) \end{bmatrix} \\ &= \sum_k \sum_l F^c(l)^T \mathbf{y}(m-l, n-k) F^r(k) \end{aligned} \tag{4.12}$$

where

$$F^c(n) \triangleq \begin{bmatrix} f_0^c(2n) & f_0^c(2n-1) \\ f_1^c(2n) & f_1^c(2n-1) \end{bmatrix}$$

and

$$F^r(n) \triangleq \begin{bmatrix} f_0^r(2n) & f_0^r(2n-1) \\ f_1^r(2n) & f_1^r(2n-1) \end{bmatrix}.$$

Equation (4.12) can be written in a more compact form for FIR interpolators of order (q_c, q_r) in a manner similar to (4.10), namely

$$\hat{\mathbf{x}}(m, n) = F_a^{cT} \mathbf{y}_a(m, n) F_a^r \tag{4.13}$$

where the augmented matrices F_a^c, F_a^r are defined from $F^c(n)$ $F^r(n)$ as in (2.16a), and

$$\begin{aligned} & \mathbf{y}_a(m, n) \triangleq \\ & \begin{bmatrix} \mathbf{y}(m, n) & \cdots & \mathbf{y}(m, n-q_r) \\ \vdots & & \vdots \\ \mathbf{y}(m-q_c, n) & \cdots & \mathbf{y}(m-q_c, n-q_r) \end{bmatrix}. \end{aligned} \tag{4.14}$$

Using Kronecker product notation, we can rewrite (4.13) as

$$\hat{\tilde{\mathbf{x}}}(m, n) \triangleq \text{vec}\{\hat{\mathbf{x}}(m, n)\} = (F_a^{rT} \otimes F_a^{cT})\tilde{\mathbf{y}}_a(m, n) \quad (4.15)$$

where $\tilde{\mathbf{y}}_a(m, n) \triangleq \text{vec}\{\mathbf{y}_a(m, n)\}$; in (4.15), we used the property of Kronecker products

$$\text{vec}\{ABC\} = (C^T \otimes A)\,\text{vec}\,B.$$

Expressions (4.12) and consequently (4.13) are the interpolation formulas when separable 2-D interpolators are used. For the general 2-D case, however, i.e., when *nonseparable* interpolators are used, (4.2) should be modified to

$$\hat{x}(m, n) = \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k}\sum_{l} f_{ij}(m - 2l, n - 2k)y_{ij}(l, k).$$

Hence

$$\begin{aligned}\hat{\tilde{\mathbf{x}}}(m, n) &\triangleq \text{vec}\{\hat{\mathbf{x}}(m, n)\} \\ &= \sum_k\sum_l F(m - l, n - k)\tilde{\mathbf{y}}(l, k) \\ &= \sum_k\sum_l F(l, k)\tilde{\mathbf{y}}(m - l, n - k) \quad (4.16)\end{aligned}$$

where we have the matrix $F(m,n)$ at the bottom of the page, and $\tilde{\mathbf{y}}(m, n) \triangleq \text{vec}\{\mathbf{y}(m, n)\}$. For FIR interpolators $f_{ij}(m, n)$ of order (q_c, q_r), (4.16) can be written as

$$\begin{aligned}&\hat{\tilde{\mathbf{x}}}(m, n) = \\ &\quad [F(0, 0) \quad F(1, 0) \quad \cdots \quad F(q_c, 0) \quad \cdots \\ &\quad\ F(0, q_r) \quad F(1, q_r) \quad \cdots \quad F(q_c, q_r)] \\ &\quad \cdot \begin{bmatrix} \tilde{\mathbf{y}}(m, n) \\ \tilde{\mathbf{y}}(m - 1, n) \\ \vdots \\ \tilde{\mathbf{y}}(m - q_c, n) \\ \cdots \\ \tilde{\mathbf{y}}(m, n - q_r) \\ \tilde{\mathbf{y}}(m - 1, n - q_r) \\ \tilde{\mathbf{y}}(m - q_c, n - q_r) \end{bmatrix}.\end{aligned}$$

The data vector of the RHS of this expression contains the same samples with $\mathbf{y}_a(m, n)$; thus, $\mathbf{y}$ rearranging its entries, we can write

$$\hat{\tilde{\mathbf{x}}}(m, n) = \mathcal{F}\,\text{vec}\{\mathbf{y}_a(m, n)\} \quad (4.17)$$

where

$$\begin{aligned}&\mathcal{F} = \\ &[F(0, 0) \quad F(1, 0) \quad \cdots \quad F(q_c, 0) \quad \cdots \\ &\quad F(0, q_r) \quad F(1, q_r) \quad \cdots \quad F(q_c, q_r)]\mathbf{E}\end{aligned} \quad (4.18)$$

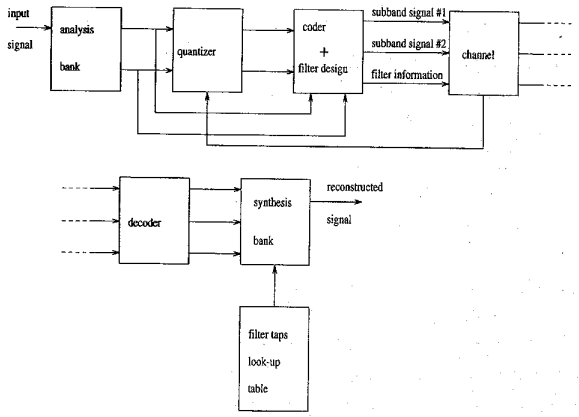


Fig. 3. Subband coder-decoder implementing optimal filter design.

$\mathbf{y}_a(m, n)$ is given by (4.14) and $\mathbf{E}$ is a unitary permutation matrix. Note that (4.15) is a special case of (4.17) with $\mathcal{F} = F_a^{rT} \otimes F_a^{cT}$.

When additive subband noise has been added to $y_{ij}(m, n)$, $\mathbf{y}_a(m, n)$ in (4.13) and (4.17) should be replaced by $\mathbf{z}_a(m, n) = \mathbf{y}_a(mn) + \mathbf{v}_a(m, n)$ where $\mathbf{v}_a(m, n)$ is defined similarly to $\mathbf{y}_a(m, n)$. Then, reconstruction (4.18) transforms to

$$\hat{\tilde{\mathbf{x}}}(m, n) = \mathcal{F}\tilde{\mathbf{z}}_a(m, n) \quad (4.19)$$

with $\tilde{\mathbf{z}}_a(m, n) \triangleq \text{vec}\{\mathbf{z}_a(m, n)\}$. Equation (4.19) is of the same form as (2.24), hence. following the same steps we conclude that for fixed decimators $\mathcal{H}^r$, $\mathcal{H}^c$ the optimal in the mean-squared-error sense interpolator is given by

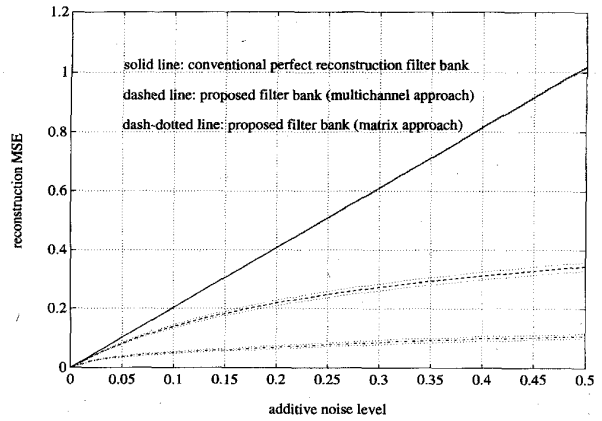
$$\mathcal{F}_{opt} = R_{\tilde{z}_a\tilde{z}_a}^{-1} R_{\tilde{z}_a\tilde{x}} \quad (4.20)$$

where $R_{\tilde{z}_a\tilde{z}_a} \triangleq \overline{E}\{\tilde{\mathbf{z}}_a(m, n)\tilde{\mathbf{z}}_a(m, n)^T\}$ and $R_{\tilde{z}_a\tilde{x}} \triangleq \overline{E}\{\tilde{\mathbf{z}}_a(m, n)\tilde{\mathbf{x}}(m, n)^T\}$. In general, the RHS of (4.20) yields some $\mathcal{F}_{opt}$, which is not a "perfect" Kronecker product. Thus, in general, optimal 2-D interpolators have not separable transfer functions, i.e., $\mathcal{F}_{opt}$ cannot be split into some $F_{a,\,opt}^{rT}$, $F_{a,\,opt}^{cT}$ such that $\mathcal{F}_{opt} = F_{a,\,opt}^{rT} \otimes F_{a,\,opt}^{cT}$. In practice this can be checked out by estimating $\mathcal{F}_{opt}$ and by using Lemma 4.1. As in the matrix approach, suboptimal separable filters can be obtained by computing $F_{a,\,opt}^{rT}$, $F_{a,\,opt}^{cT}$ via the singular vectors of Lemma 4.1 that correspond to the largest singular value.

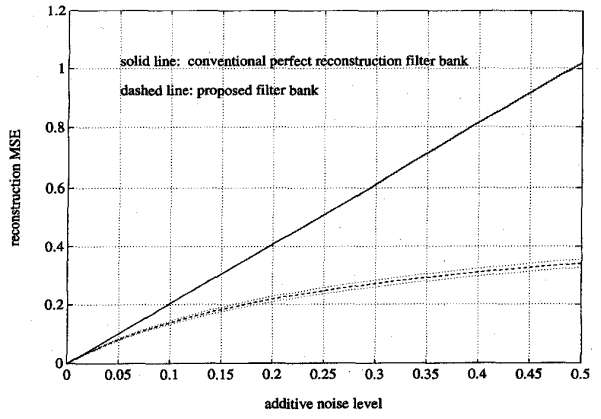
V. An Application

The improvement in the quality of the reconstructed signal, obtained through the proposed optimal synthesis filter design, is of great interest in coding applications for transmission and storage of 1-D signals and 2-D images. The signal-to-noise ratio (SNR) increase in the output, without a corresponding

$$F(m, n) = \begin{bmatrix} f_{00}(2m, 2n) & f_{10}(2m, 2n) & f_{01}(2m, 2n) & f_{11}(2m, 2n) \\ f_{00}(2m - 1, 2n) & f_{10}(2m - 1, 2n) & f_{01}(2m - 1, 2n) & f_{11}(2m - 1, 2n) \\ f_{00}(2m, 2n - 1) & f_{10}(2m, 2n - 1) & f_{01}(2m, 2n - 1) & f_{11}(2m, 2n - 1) \\ f_{00}(2m - 1, 2n - 1) & f_{10}(2m - 1, 2n - 1) & f_{01}(2m - 1, 2n - 1) & f_{11}(2m - 1, 2n - 1) \end{bmatrix}$$



(a)



(b)

Fig. 4. (a) Reconstruction mean squared error versus additive noise level—stationary 1-D signal, stationary additive noise, optimal filters designed based on theoretically computed statistics; (b) reconstruction mean squared error versus additive noise level—stationary 1-D signal, stationary additive noise, optimal filters designed based on estimated statistics.

bit rate overload, is perhaps the most important contribution of the proposed filter design algorithms. On-line determination of the optimal synthesis filter banks in variable bit rate coding situations is also of particular interest.

The structure and the associated flow of data/information in a subband coding/transmission system that exploits the above tools is depicted in Fig. 3. The original input signal is analyzed in, say, $P = 2$ subband sequences. Next, the two subband signals are quantized. The quantizer proceeds to the determination of the number of quantization levels based on channel load information, i.e., it performs coarser or finer quantization according to the available channel capacity at each time point. Both the quantized and the quantization-free subband signals are fed to the next stage ("coder + filter design" in Fig. 3), which implements the design task. Since this stage has access to both signals, it can estimate the covariance matrices that are necessary for the optimal filter design. As an alternative, the recursive algorithm of Table I can be used to adaptively compute the optimal filter taps.

In the sequel, the quantized subband signals along with information regarding the structure of the optimal synthesis

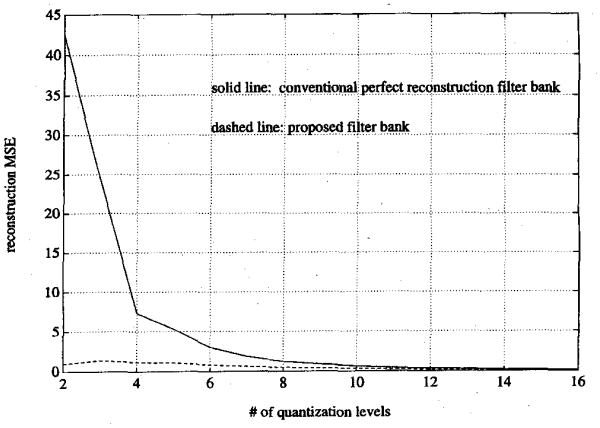


Fig. 5. Reconstruction mean squared error versus number of quantization levels: 1-D nonstationary signal, quantization noise, optimal filter design based on estimated statistics.

filter bank are coded and transmitted through the channel. Coder operation is errorless. Consequently, assuming a lossless channel, the decoder's output will consist of exact duplicates of the quantized subband signals and optimal filter information. Synthesis bank uses the available optimal filter information in order to adjust the impulse responses of the reconstruction filters. This information could be either an explicit description of the optimal impulse responses or a pointer to a lookup table that is accessible from the synthesis bank and contains a variety of possible impulse responses for the synthesis filters. The latter choice consists of a suboptimal solution but is more economical, since it does not require transmission of the entire impulse response.

It is important to note that the aforementioned overhead can be eliminated using a quite different approach, the design of optimal analysis filters for fixed synthesis banks; the corresponding procedure is more complicated, since the spectral characteristics of the quantization noise may rely heavily on the choice of analysis filters. In this case, a joint design of optimal analysis filters and associated quantizers seems natural.

The above scheme describes the simplest possible subband transmission structure employing the proposed design methods. In practice, for example, in moving picture coding applications, more complex structures should be incorporated (e.g., including motion compensators). The main structure regarding the filter design procedure, however, remains unaffected.

In the case of compression methods based on subband coding, we can employ almost the same flow of information. In this scenario the compression program computes the optimal reconstruction filter banks and stores their impulse responses in the header of the file that contains the compressed signal in order to be accessible to the decompression program.

VI. Simulations

In this section we present a series of simulated examples in order to illustrate the advantages of using the proposed optimal reconstruction filter over the conventional "perfect

reconstruction" ones. The design procedure is tested in noisy environments both on stochastic and deterministic signals of one and two dimensions. We compare the performance of the proposed filters versus that of optimal reconstruction filters when both are used to reconstruct signals based on their noisy subband components at various SNR levels.

In experiments 1 and 3 the noise was generated by a pseudorandom external source and the SNR was common to both subbands. Experiments 2 and 4 refer to quantization noise produced by a simple uniform quantizer; the number of quantization levels varied between successive experiments, but was the same for both channels.

In all examples, we used as a starting point the quadrature miror filters (QMF's) corresponding to the lowpass FIR filter of length 8

$$h(n) = [0.0094 \quad -0.0707 \quad 0.0694 \quad 0.4900 \quad 0.4900 \quad 0.0694 \quad -0.0707 \quad 0.0094]$$

which was proposed in [6, Appendix 7.1, pp. 401–404]. The impulse responses of the analysis band were given by $h_0(n) = h(n)$, $h_1(n) = (-1)^n h(n)$ while the corresponding. *conventional* synthesis filters were $f_0(n) = 2h(n)$ and $f_1(n) = -2(-1)^n h(n)$.

Experiment 1—Stationary Stochastic Signal Stationary External Noise: The input signal $x(n)$ was an ARMA process created by exciting the system

$$\frac{1 + 0.3z^{-1} + 0.4z^{-2} - 0.5z^{-3} + 0.6z^{-4} - 0.7z^{-5}}{1 - 0.25z^{-1} + 0.3z^{-2} - 0.5z^{-3}}$$

by zero-mean white driving noise. Realizations of length $N = 1000$ of the signal were decimated via $h_0(n)$, $h_1(n)$ generating the signals $y_0(n)$, $y_1(n)$ in accordance to (2.2) with $P = 2$. Noise signals $v_0(n)$ and $v_1(n)$ were also zero-mean ARMA processes generated by the systems

$$\frac{1 + 0.7z^{-1} - 0.7z^{-2} + 0.4z^{-3} - 0.5z^{-4} + 0.6z^{-5} - 0.7z^{-6}}{1 + 0.5z^{-1} - 0.3z^{-2}}$$

and

$$\frac{1 + 0.3z^{-1} - 0.4z^{-2} + 0.5z^{-3} + 0.8z^{-4} - 0.8z^{-5} - 0.6z^{-6}}{1 - 0.25z^{-1} + 0.5z^{-2}}$$

respectively. For optimal filter design, we used both the matrix and the multichannel approaches.

Case 1a: The theoretical expressions of the matrices $R_{z_a z_a}$, R_{zz}, and $R_{z_a x}$, R_{zx} were initially employed in order to obtain solutions for (3.4) and (3.7). The designed optimal synthesis filters were subsequently used for the reconstruction of the original signals based on their noisy subband components $z_0(n) = y_0(n) + v_0(n)$ and $z_1(n) = y_1(n) + v_1(n)$. Fig. 4(a) depicts the mean-squared reconstruction error ((± standard deviation) as approximated using 100 Monte Carlo iterations versus the ratio of signal's power over noise's power. The level of noise was the same in both channels. The solid line corresponds to the use of conventional perfect reconstruction filter banks, the dashed line corresponds to the filters designed using (3.7), and the dashed-dotted line represents the reconstruction error resulting when (3.4) was used for filter design.

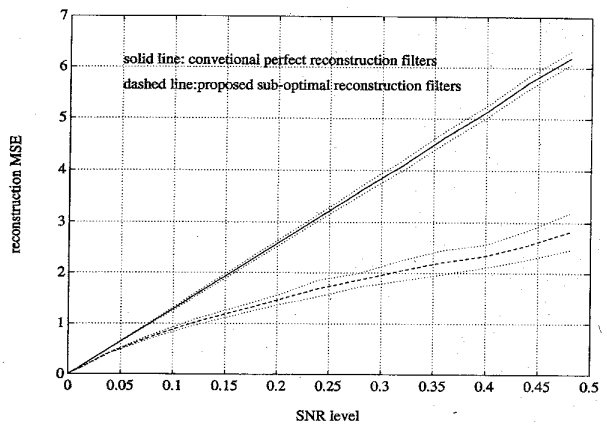


Fig. 6. Reconstruction mean squared error versus additive noise level: stationary 2-D signal, stationary additive noise, filter design based on estimated statistics.

Case 1b: Instead of the true values of $R_{z_a z_a}$ and $R_{z_a x}$, we used their estimates $R_{z_a z_a} hat$ and $\hat{R}_{z_a x}$. The previous experiment (100 Monte Carlo iterations) was repeated for filters designed on the basis of the multichannel approach [i.e., (3.7)]. In the diagram of Fig. 4(b) we show the performance of the conventional and proposed filter banks (obtained using the multichannel approach) for various noise levels. Again, the solid line corresponds to the conventional filters and the dashed line refers to the reconstruction capability of the proposed filters.

Experiment 2—Deterministic Input Signal Quantization Noise: We generated a single deterministic input signal $x(n)$ of length $N = 1000$, scanning row-by-row a gray-scale image created by a scanner device from a printed real image. We used the analysis filters $h_0(n)$, $h_1(n)$ of the previous experiment in order to obtain two subband components of $x(n)$. Both subband signals were quantized via a uniform quantizer. Their quantized versions were then fed to the interpolation filter banks corresponding to the conventional perfect reconstruction filters and to the proposed (multichannel approach) optimal reconstruction filters of (3.7). The autocovariance structures $R_{z_a z_a}$ and $R_{z_a x}$ in (3.7) were estimated using standard autocovariance estimators. The resulting reconstruction MSE's for both filter banks is plotted in Fig. 5 versus the varying number of adopted quantization levels. It should be noted that the impulse responses of the optimal synthesis banks were recalculated for all different numbers of quantization levels, since any change of this number would result in a change of the autocovariance matrices $R_{z_a z_a}$ and $R_{z_a x}$.

Experiment 3—Stationary Stochastic 2-D Signal (Texture) Stationary External Noise: The 2-D input signal $x(m, n)$ was an MA process created by convolving an uncorrelated signal $u(m, n)$ with the mask

$$b = \begin{bmatrix} 1 & -0.5 \\ -0.2 & 0.4 \end{bmatrix}.$$

Realizations of size 128×128 of $x(m, n)$ were decimated using separable analysis filters with

$$h_0^c(n) = h_0^r(n) = h_0(n)$$

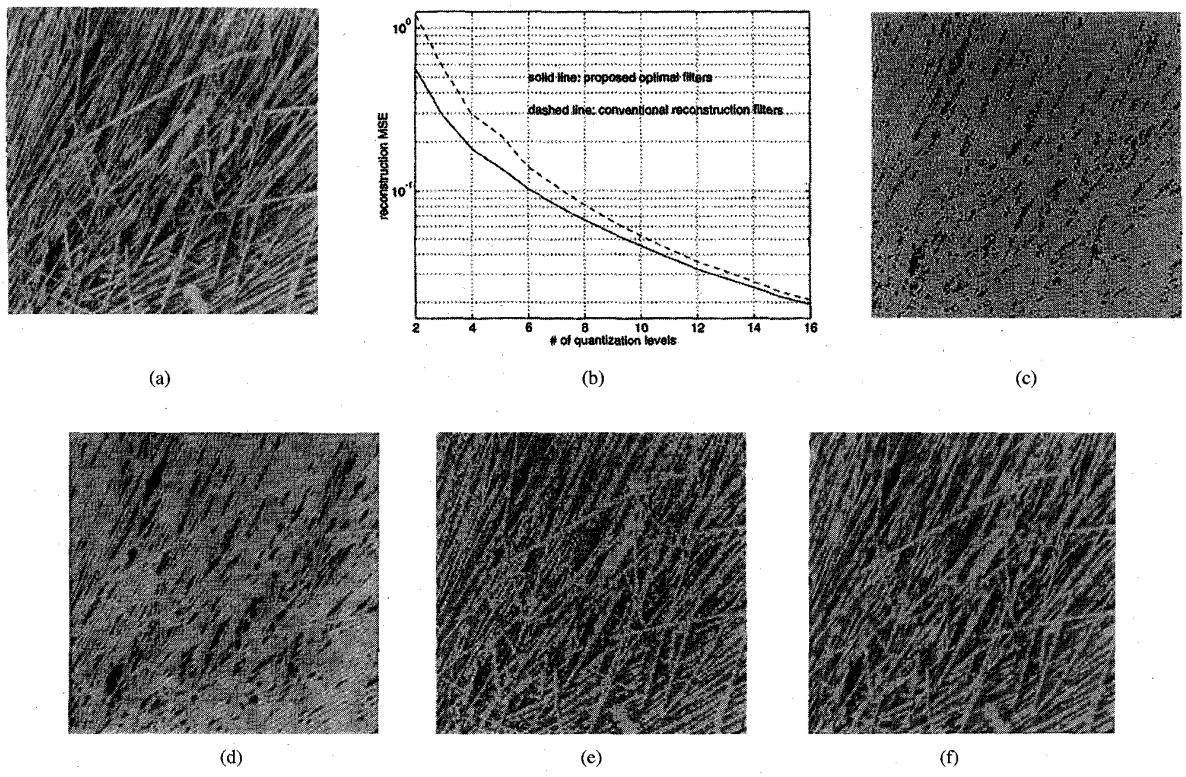


Fig. 7. (a) Original image of Experiment 4; (b) reconstruction SNR versus number of quantization levels: 2-D real-life signal, quantization subband noise, filter design based on estimated statistics; (c) image reconstructed via conventional filters; Subband images uniformly quantized into two levels; (d) image reconstructed via optimally designed filters (subband images uniformly quantized into two levels); (e) image reconstructed via conventional filters (subband images uniformly quantized into four levels; (f) image reconstructed via optimally designed filters (subband images uniformly quantized into four levels.

$$h_1^c(n) = h_1^r(n) = h_1(n).$$

In the case of perfect reconstruction filter banks, we used the corresponding synthesis filters

$$f_0^c(n) = f_0^r(n) = f_0(n)$$
$$f_1^c(n) = f_1^r(n) = f_1(n).$$

White additive subband noise at various SNR levels was added to all four subband images. After estimating the autocovariance statistics of signal and noise, we used the proposed multichannel approach (cf., (4.20)) to design the optimal reconstruction filters. The resulting optimal 2-D interpolators were nonseparable; we next used Lemma 4.1 to approximate them by suboptimal separable ones. Fig. 6 illustrates the performance (mean $\pm$ st. dev. of 100 Monte Carlo runs) of the conventional (solid line) and the proposed separable filters (dashed line). In the diagram, the resulting reconstruction MSE is plotted versus the level of additive noise.

Experiment 4—2-D Signal Quantization Noise: The 256×256 real-life image of Fig. 7(a) was used here in the form of a 2-D input signal. We used the same analysis filters that we used in Experiment 3. All four subband signals were uniformly quantized in different numbers of quantization levels. Signal and noise statistics were estimated and used for the design of optimal reconstruction filters according to (4.20) of the proposed multichannel approach. The optimal reconstruction impulse responses were *not* separable. In Fig. 7(b), we illustrate the achieved reconstruction SNR for the conventional perfect reconstruction synthesis filters (corresponding to the particular analysis bank; see Experiment 3) and the computed optimal synthesis filter bank versus the number of quantization levels. Fig. 7(c) and (d) depict the reconstructed images when using conventional and optimal reconstruction filter banks, respectively. Although the uniform quantization of the subband signals was extremely coarse (two quantization levels, i.e., 1 b/pixel), the optimally reconstructed image, contrary to the conventionally reconstructed one, preserves the main structure and texture of the original picture. Fig. 7(e) and (f) depicts similar results obtained after using a four-level uniform quantizer (2 b/pixel).

VII. Conclusions

In this paper, we examined the effect of additive noise in subband processing applications and propose optimal reconstruction filter design methods. These optimal filters are designed on the basis of input signal and noise statistics and minimize the MSE between the original and the reconstructed signal. We have shown that the proposed design technique can be recursively implemented in a way that allows tracking of slow variations of signal and noise spectral characteristics. Both the analysis and the design are extended to the 2-D

signal (images) case, which is accomplished with the use of Kronecker products.

The proposed techniques can be used in the subband coding of 1-D signals and images, resulting in improved quality of reconstructed signals without corresponding cost in terms of bit rate or storage space. The overhead required for the coding of the optimal filters is negligible when the underlying signals are stationary or piecewise stationary; in this case, the filter coefficients should be coded once or only a few times. In the case of nonstationary signals, an appropriate vector quantization is proposed for the indexing of the filter space; additional research work is currently in progress regarding this topic.

Using synthetic and real data, we performed a set of simulation experiments that illustrated the superiority of the proposed optimal synthesis filters against the conventionally used perfect reconstruction filters.

REFERENCES

[1] A. N. Akansu and R. A. Haddad, *Multiresolution Signal Decomposition: Transforms, Subbands and Wavelets.* New York: Academic, 1992.

[2] M. Basseville and A. Benveniste, "Multiscale statistical signal processing," in *Proc. Int. Conf. Acoust., Speech, Signal Processing,* Glasgow, Scotland, 1989, pp. 2065–2068.

[3] D. R. Brillinger, *Time Series, Data Analysis, and Theory.* New York: McGraw-Hill, 1981.

[4] K. C. Chou, A. S. Willsky, A. Benveniste, and M. Basseville, "Recursive and iterative estimation algorithms for multiresolution stochastic processes," in *Proc. 28th IEEE Conf. Dec. Contr.,* Tampa, FL, USA, Dec. 1989.

[5] K. C. Chou, S. Golden, and A. S. Willsky, "Modeling and estimation of multiscale stochastic processes," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing,* Ontario, Canada, 1991, pp. 1709–1712.

[6] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1983.

[7] I. Daubechies, "The wavelet transform: Time-frequency localization, and signal analysis," *IEEE Trans. Inform. Theory,* vol. 36, no. 5, pp. 961–1005, 1990.

[8] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston: Kluwer, 1993, 2nd printing.

[9] R. A. Haddad and N. Uzun, "Modeling, analysis, and compensation of quantization effects in M-band subband codecs," in *Proc. ICASSP'93,* Minneapolis, MN, vol. III, pp. 173–176.

[10] W. S. Jayant and P. Noll, *Digital Coding of Waveforms.* Englewood Cliffs, NJ: Prentice-Hall, 1984.

[11] J. Kovacevic, "Eliminating correlated errors in subband/wavelet coding systems with quantization," in *Proc. 27th Asilomar Conf. Signals, Syst., Comput.,* Pacific Grove, CA, Nov. 1993, pp. 881–885.

[12] L. Ljung, *System Identification, Theory for the User.* Englewood Cliffs, NJ: Prentice-Hall, 1987.

[13] D. G. Luenbeger, *Optimization by Vector Space Methods.* New York: Wiley, 1969.

[14] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Patt. Anal. Machine Intell.,* vol. 11, no. 7, pp. 674–693, July 1989.

[15] ——, "Zero-crossings of a wavelet transform," *IEEE Trans. Inform. Theory,* vol. 37, pp. 1019–1033, July 1991.

[16] K. Nayebi, T. P. Barnwell, and M. T. J. Smith, "Time-domain filter bank analysis: A new design theory," *IEEE Trans. Signal Processing,* vol. 40, no. 6, June 1992.

[17] J. M. Ortega, *Matrix Theory, A Second Course.* New York and London: Plenum, 1987.

[18] K. Park and R. A. Haddad, "Optimum subband filter design and compensation in presence of quantizers," in *Proc. 27th Asilomar Conf. Signals, Syst., Comput.,* Pacific Grove, CA, Nov. 1993, pp. 80–83.

[19] G. Strang, "Wavelets and dilation equations: A brief introduction," *SIAM Rev.,* vol. 31, no. 4, pp. 614–627, Dec. 1989.

[20] A. Tirakis and S. Kollias, "Efficient image classification using neural networks and multiresolution analysis," in *Proc. Int. Conf. Acoust., Speech, Signal Processing,* Minneapolis, MN, 1993.

[21] M. Tsatsanis, "On wavelets and time-varying system identification," Ph.D. dissertation, Univ. of Virginia, Charlotesville, Sept. 1992.

[22] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial," *Proc. IEEE,* vol. 78, no. 1, Jan. 1990,

[23] M. Veterli, "Multi-dimensional sub-band coding: Some theory and algorithms," *Signal Processing,* vol. 6, pp. 97–112, Apr. 1984.

[24] P. H. Westerink, J. Biemond, and D. E. Boekee, "Scalar quantization error analysis for image subband coding using QMF's," *IEEE Trans. Signal Processing,* vol. 40, no. 2, pp. 421–428, Feb. 1992.

[25] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 34, pp. 1278–1288, Oct. 1986.

Anastasios N. Delopoulos (M'94) was born in Athens, Greece, in 1964. He graduated from the Department of Electrical Engineering, National Technical University of Athens (NTUA), Greece, in 1987, received the M.Sc. degree in electrical engineering from the University of Virginia, Charlotesville, USA, in 1990 and the Ph.D. degree in electrical and computer engineering from NTUA in 1993.

His current research interests lie in the areas of system identification, video coding, multimedia, massively parallel architectures, and biomedical engineering.

Dr. Delopoulos is a member of the Technical Chamber of Greece and a member of the IEEE Signal Processing Society.

Stefanos D. Kollias (M'87) was born in Athens, Greece, in 1956. He received the Diploma degree in electrical engineering from National Technical University of Athens (NTUA), in 1979, the M.Sc degree in communication engineering from the University of Manchester Institute of Science and Technology, Manchester, UK, in 1980, and the Ph.D. degree in signal processing from the Computer Science Division, NTUA, in 1984.

Since 1986, he has been the with Computer Science Division of NTUA, where he is currently an associate professor. From 1987 to 1988, he was a visiting research scientist with the Electrical Engineering Department of Columbia University, New York, NY, USA. His research interests include image processing and analysis, video coding, multimedia systems, artificial neural networks, and multimedia engineering.