

Modulated Filter Banks with Arbitrary System Delay: Efficient Implementations and the Time-Varying Case

Gerald D. T. Schuller, *Member, IEEE*, and Tanja Karp, *Member, IEEE*

Abstract—In this paper, we present a new method for the design and implementation of modulated filter banks with perfect reconstruction. It is based on the decomposition of the analysis and synthesis polyphase matrices into a product of two different types of simple matrices, replacing the polyphase filtering part in a modulated filter bank. Special consideration is given to cosine-modulated as well as time-varying filter banks. The new structure provides several advantages. First of all, it allows an easy control of the input-output system delay, which can be chosen in single steps of the input sampling rate, independent of the filter length. This property can be used in audio coding applications to reduce pre-echoes. Second, it results in a structure that is nearly twice as efficient as performing the polyphase filtering directly. Perfect reconstruction is a structurally inherent feature of the new formulation, even for nonlinear operations or time-varying coefficients. Hence, the structure is especially suited for the design of time-varying filter banks where both the number of bands as well as the prototype filters can be changed while maintaining perfect reconstruction and critical sampling. Further, a proof of effective completeness is given, and the design of equal magnitude-response analysis and synthesis filter banks is described. Filter design can be performed by non-constrained optimization of the matrix coefficients according to a given cost function. Design and audio-coding application examples are given to show the performance of the new filter bank.

Index Terms—Audio coding, low system delay, modulated filter banks, polyphase formulation, pre-echoes, time-varying filter banks.

I. INTRODUCTION

MODULATED filter banks are popular because they provide computationally efficient implementation and great design ease since only the FIR baseband analysis and synthesis prototypes need to be designed and evaluated [1]–[4]. Historically, the first modulated filter banks with perfect reconstruction were paraunitary and used cosine modulation [5]–[9]; see also [10] and [11]. However, in paraunitary filter banks, the input-output delay of the filter bank is fixed as the length of the filter minus 1. Thus, FIR filter banks where the filters have a high stopband attenuation and/or narrow transition bandwidth yield a large system delay since long filter impulse responses are needed for the filter design.

In applications like speech and audio coding, it is important to have a low system delay in order to have not only a low round-trip delay and avoid audible distortion (pre-echoes) but

also have narrowband filters with high stopband attenuation. Both features can be obtained when using low-delay filter banks [12] or biorthogonal modulated filter banks [13]–[19]. Low-delay filter banks offer a low system delay independent of the filter length. This enables a higher stopband attenuation and/or a narrower transition bandwidth than an orthogonal filter bank with the same overall system delay. Two different approaches for cosine-modulated filter banks with arbitrary delay were presented in [18] and [19]. Both approaches use different phases for the modulation function. The method presented in [19] explicitly derives the constraints on the prototypes' polyphase components for perfect reconstruction, and a quadratic-constrained optimization is proposed for the filter design. On the other hand, [18] proposes an efficient implementation that automatically guarantees perfect reconstruction and a chosen system delay of the filter bank such that the prototype filters can be designed using unconstrained optimization. This latter approach will be used in this paper.

Most of the real-world signals being treated with filter banks cannot be considered stationary. In order to improve the coding efficiency of the filter bank, it is useful to adapt the filter characteristic and the number of bands to the signal statistics, e.g., for sinusoidal-like signals, it is best to have many narrow bands, yielding long filters, whereas for clicks or attacks in audio signals (or edges in images), it is best to have a few short filters. In [20], it was shown how to design a time-varying cosine-modulated TDAC filter bank [5]. However, its filters are orthogonal with length restricted to $2M$ and a system delay of $2M - 1$ samples. That algorithm is used, e.g., in ISO MPEG audio coders [21]. Later approaches were mainly for nonmodulated filter banks [22]–[28] and for the orthogonal case [29].

To address the above issues, this paper treats the following main points: It describes how to design filter banks with a truly arbitrary delay, which can be specified in single sample steps, so that fine-tuned compromises between filter length and system delay can be obtained. A proof of effective completeness for the new design method by extracting factors is provided. The proof shows that all contiguous impulse responses for prototypes of biorthogonal cosine-modulated filter banks that yield perfect reconstruction can be implemented using the new structure. We show how to design equal-magnitude response analysis and synthesis filter banks, which is important for coding applications. Finally, we describe how to design modulated time-varying filter banks with arbitrary system delay based on a new polyphase description for time-varying filter banks [26], [30]. The time variation includes changing the number of bands and/or the prototypes while maintaining perfect reconstruction

Manuscript received February 1, 1997; revised August 18, 1999. The associate editor coordinating the review of this paper and approving it for publication was Dr. Henrique Malvar.

G. D. T. Schuller is with the Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974, USA.

T. Karp is with the University of Mannheim, Mannheim, Germany.

Publisher Item Identifier S 1053-587X(00)01544-0.

and critical sampling. The proposed structure is very general and holds true for odd and even numbers of subbands.

The outline of the paper is as follows. After providing some definitions and notations, we recall the polyphase formulation of a modulated filter bank with perfect reconstruction and a given system delay in Section III. In Section IV, the new filter structure is described. It is based on a factorization of the polyphase matrix into a cascade of two types of simple matrices. Section V explains how to use the new structure for filter design. A proof for the effective completeness of the new factorization is given in the Appendix. Useful symmetries for obtaining equal-magnitude responses of the analysis and synthesis filters, even for low-delay filter banks, are then discussed. Design examples are given. In Section V-B, we show that the factorization is not only useful for the mathematical description, but it also yields an efficient ladder-like filter structure. In Section VI, we demonstrate that the same filter structure can be used for time-varying modulated filter banks. Design examples are shown at the end of the section, including a practical example of a proposed ISO MPEG standard audio coder whose filter bank was replaced by the described time-varying low-delay filter bank to improve its performance.

II. NOTATION AND DEFINITIONS

Boldface letters denote matrices or vectors, capital letters z -transforms or polynomials, and “:=” means “defined as.” A polynomial matrix $\mathbf{F}(z)$ is causal if it contains no positive powers of z . Symbols \mathbf{I} and \mathbf{J} denote the $M \times M$ identity and counter-identity matrix, respectively. $\mathbf{S}(z)$ is a shift matrix that advances a block or vector by one sample and \mathbf{diag} is an $M \times M$ diagonal matrix

$$\mathbf{S}(z) := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & \vdots \\ \vdots & & 0 & \ddots & 0 \\ 0 & \vdots & \vdots & & 1 \\ z & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$\mathbf{diag}(x_0, \cdots, x_{M-1}) := \begin{bmatrix} x_0 & 0 & \cdots & 0 \\ 0 & x_1 & \cdots & 0 \\ & & \ddots & \vdots \\ 0 & 0 & \cdots & x_{M-1} \end{bmatrix}.$$

The symbol $[\mathbf{F}]_{n,k}$ denotes the element at the n th row and k th column of the matrix \mathbf{F} . The degree of $\mathbf{F}(z)$ is defined as the difference between the highest and the lowest power of z . $\prod_{j=1}^{\nu} \mathbf{L}_j := \mathbf{L}_1 \cdots \mathbf{L}_\nu$; here, the ordering is important. The notation \mathbf{x}^T denotes the transpose of a vector \mathbf{x} .

The filter bank structure is as follows. The analysis filter bank consists of M parallel analysis filters of length N with impulse responses $h_k(n)$, $k = 0, \cdots, M-1$, $n = 0, \cdots, N-1$, and subsequent downsampling by M . The input signal is $x(n)$, and the subband signals are $y_k(m)$, $k = 0, \cdots, M-1$, where m is the time index at the reduced sampling rate. The synthesis filter bank consists of upsamplers by M followed by M synthesis filters with impulse responses $g_k(n)$, $k = 0, \cdots, M-1$, $n = 0, \cdots, N-1$. The filter outputs are summed to form the

reconstructed signal $\hat{x}(n)$. The filter bank provides perfect reconstruction if the output signal is a delayed version of the input signal $\hat{x}(n) = x(n - n_d)$, where n_d is the system delay, assuming that the subband signals are directly passed from the analysis to the synthesis bank.

III. THE POLYPHASE FORMULATION

The new factorization proposed in this paper is based on the well-known polyphase formulation [1]. The effect of downsampling and upsampling in the analysis and synthesis filter bank, respectively, can be viewed as processing the signal in blocks of length M . The input signal is represented by an M -dimensional vector $\mathbf{x}(m)$ composed of sequences of the downsampled $x(n)$

$$\mathbf{x}(m) := [x_0(m), \cdots, x_{M-1}(m)]^T$$

with

$$x_i(m) := x(mM + i). \quad (1)$$

The z -transform of $\mathbf{x}(m)$ is given by

$$\mathbf{X}(z) = [X_0(z), \cdots, X_{M-1}(z)]^T.$$

The subband signals are represented by the z -domain column vector $\mathbf{Y}(z) = [Y_0(z), \cdots, Y_{M-1}(z)]^T$ and the reconstructed signal by $\hat{\mathbf{X}}(z)$ being defined in the same way as $\mathbf{X}(z)$. The analysis polyphase matrix $\mathbf{E}(z)$ for causal filters contains the n th type-2 polyphase component [1, pp. 121, 122] of the k th analysis filter at position $[\mathbf{E}(z)]_{k,n}$

$$[\mathbf{E}(z)]_{k,n} := \sum_{m=0}^{\infty} h_k(mM + M - 1 - n)z^{-m}$$

$$n, k = 0, \cdots, M-1.$$

This formulation can be seen as a generalization of nonoverlapping block transforms, like a DCT, to (multiple) overlapped blocks. The synthesis polyphase matrix consists of type-1 polyphase components according to

$$[\mathbf{R}(z)]_{n,k} := \sum_{m=0}^{\infty} g_k(mM + n)z^{-m}.$$

The operations in the analysis and synthesis filter bank can then be written as

$$\mathbf{Y}(z) = \mathbf{E}(z) \cdot \mathbf{X}(z), \quad \hat{\mathbf{X}}(z) = \mathbf{R}(z) \cdot \mathbf{Y}(z).$$

Fig. 1 shows the polyphase filter bank structure.

The method in [16] only considered system delays in steps of integer multiples of M . However, there can be a tradeoff between the filter quality and the system delay. Therefore, it is important to have a finer choice of the system delay. From [1], a general formulation for perfect reconstruction is known. It can effectively be written with the help of the shift matrix \mathbf{S}

$$\mathbf{R}(z) \cdot \mathbf{E}(z) = z^{-2\mu} \cdot \mathbf{S}^{n_0} \quad (2)$$

with μ and n_0 being non-negative integers. The right side represents a delay of $2\mu M - n_0$ samples. For causal systems, the exponent n_0 has to be small enough so that the above product

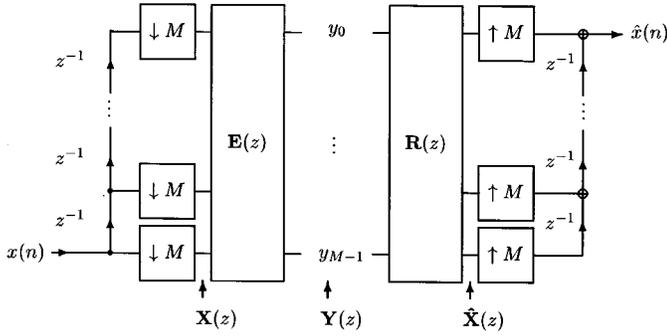


Fig. 1. Polyphase representation of an M -channel filter bank with critical downsampling.

has no positive powers of z , i.e., $n_0 \leq 2\mu M$. The so called block delay of $M - 1$ samples has to be added to this delay to obtain the system delay n_d [1, p. 237]. This results from assembling the signal into blocks of length M (see Fig. 1). Therefore, the overall system delay n_d of the filter bank is

$$n_d = 2\mu M + M - 1 - n_0. \quad (3)$$

It can be seen that n_0 allows the specification of the system delay in single steps of the input sampling rate.

Observe that orthogonal filter banks with filter length N have a system delay of $n_d = N - 1$. This delay will also be called a standard delay.

IV. NEW FACTORIZATION

As a modulation scheme, we consider impulse responses of cosine-modulated filter banks of the form

$$h_k(n) = h(n) \cdot \cos\left(\frac{\pi}{M} (k + 0.5)(n + 0.5 + n_a)\right) \quad (4)$$

$$g_k(n) = h'(n) \cdot \cos\left(\frac{\pi}{M} (k + 0.5)(n + 0.5 - M + n_s)\right) \quad (5)$$

$k = 0, \dots, M - 1$, $n = 0, \dots, N - 1$, where $h(n)$ and $h'(n)$ are the analysis and synthesis baseband prototype filters, respectively. The additional shift of M in (5) is introduced to simplify the following notation. Other modulation schemes are also possible, e.g., different cosine and sine modulations (see also [16] or [10], [11], and [19]), but for clarity, we concentrate on the above form.

It is well known that modulated filter banks provide an efficient implementation based on the polyphase components of the prototypes and a fast transform. The first step in our formulation is to split or decompose the polyphase matrix $\mathbf{E}(z)$ in (2) into the product of a sparse “filter matrix” $\mathbf{F}(z)$ containing the prototype polyphase components, a coefficient or transform matrix \mathbf{T} , and the shift matrices \mathbf{S}^{n_a} and \mathbf{S}^{n_s} , adjusting the system delay

$$\mathbf{E}(z) = \mathbf{T} \cdot \mathbf{F}(z) \mathbf{S}^{n_a}(z) \quad (6)$$

$$\mathbf{R}(z) = \mathbf{S}^{n_s}(z) z^{-2\mu} \mathbf{F}^{-1}(z) \cdot \mathbf{T}^{-1}. \quad (7)$$

Here again, in the product containing $\mathbf{S}^{n_a}(z)$, all coefficients that would lead to positive exponents of z have to be zero in

order to obtain causal filters. The values n_a and n_s can be chosen as

$$n_a = n_s = \frac{n_0}{2} \quad (8)$$

with n_0 from (2). For odd n_0 , the exponent n_a can be chosen as $n_s + 1$.

For the analysis and synthesis filters according to (4) and (5), respectively, a suitable transform matrix \mathbf{T} is

$$[\mathbf{T}]_{k,n} := \cos\left(\frac{\pi}{M} (k + 0.5)(n + 0.5)\right), \quad 0 \leq n, k < M$$

which is the well-known DCT type IV. The filter matrix $\mathbf{F}(z)$ then has a sparse, “bi-diagonal” form

$$\begin{bmatrix} \bullet & & 0 & & \bullet \\ & \ddots & & \ddots & \\ & & \bullet & \bullet & 0 \\ 0 & & \bullet & \bullet & \\ & \ddots & & \ddots & \\ \bullet & & 0 & & \bullet \end{bmatrix}$$

or, more specifically

$$\mathbf{F}(z) = \text{diag}[P_0(-z^2), \dots, P_{M-1}(-z^2)] \cdot \mathbf{J} + z^{-1} \cdot \text{diag}[P_{2M-1}(-z^2), \dots, P_M(-z^2)] \quad (9)$$

$$z^{-2\mu} \mathbf{F}^{-1}(z) = \text{diag}[P'_0(-z^2), \dots, P'_{M-1}(-z^2)] \cdot \mathbf{J} - z^{-1} \text{diag}[P'_M(-z^2), \dots, P'_{2M-1}(-z^2)] \quad (10)$$

with the time-shifted polyphase representation of the prototypes

$$P_k(z) = \sum_{m=0}^{\infty} h(m2M + k - n_a) z^{-m} \quad (11)$$

$$k = 0, \dots, 2M - 1$$

$$P'_k(z) = \sum_{m=0}^{\infty} h'(m2M + k - n_s) z^{-m}. \quad (12)$$

That the filter matrix has this bi-diagonal form also means that this modulated filter bank can be viewed as a set of nested two-band filter banks followed by the cosine transform matrix \mathbf{T} , as can also be seen in [1] and [16].

In the following, we directly use the sparse filter matrix $\mathbf{F}(z)$ for the filter design. For perfect reconstruction and the desired system delay, we have to solve $\mathbf{F}^{-1}(z)$. In general, this approach leads to IIR synthesis filters for a given set of FIR analysis filters. We here restrict ourselves to FIR analysis and synthesis filters in order to guarantee stability. Existing approaches effectively use a factorization of the matrix $\mathbf{F}(z)$ into a product of paraunitary matrices, which leads to orthogonal filter banks [5], [6]. In the following, a different factorization is used to obtain a more general formulation, which also covers nonorthogonal or biorthogonal filter banks.

With the new factorization, we construct the filter matrices as a product or cascade of simple matrices. These simple matrices

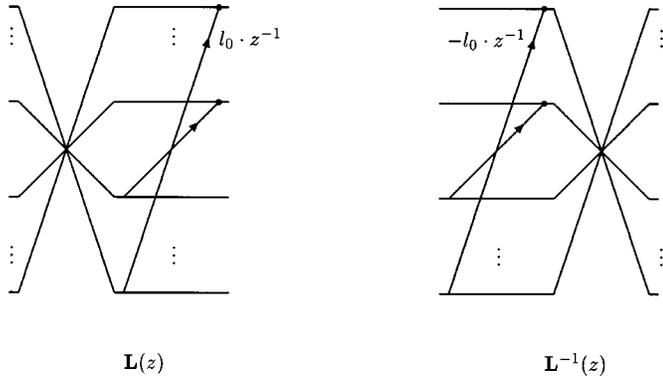


Fig. 2. Structure of the zero-delay matrices.

have an FIR inverse yielding FIR synthesis filters and are sparse with only a few elements not equal to 1 or 0. Since there are two independent variables in the design process (the system delay and the filter length), two types of matrices are needed. These simple matrices, which may also be called filter matrices, are described in the following.

Zero-Delay Matrices: They increase the filter length but not the system delay (see also [16]).

$$\mathbf{L}_i(z) := \mathbf{J} + \text{diag}(l_0^i, \dots, l_{M/2-1}^i, 0, \dots, 0) \cdot z^{-1}$$

where l_j^i are coefficients, and i denotes different sets of coefficients ($i \geq 0$). Another type, where the coefficients are on the other half of the diagonal, is also possible. It is further possible to design these matrices for odd numbers of bands M by placing a zero in the center of the matrix. Observe that this matrix type has degree one and that a product with several of these matrices has a degree that corresponds to the number of matrices. The inverse is

$$\mathbf{L}_i^{-1}(z) = \mathbf{J} - \text{diag}(0, \dots, 0, l_{M/2-1}^i, \dots, l_0^i) \cdot z^{-1}.$$

Observe that the inverse is causal and contains the same coefficients. This means that the inverse can be implemented with causal filters. Hence, the matrix cascaded with its inverse introduces no delay in a signal flow, although both are causal and have a degree of one. This property can be used to construct filter banks with multiple overlapping filters but no delay in addition to the block delay. Fig. 2 shows the structure of the zero-delay matrices and their inverse. It is easy to see that the inverse is always perfect. This is true even for time-variable coefficients or nonlinear operations like rounding. This is an important property that can also be used to design time-varying filter banks. This way, it is possible to design filter banks that can change the filters and the number of bands during signal processing, even for overlapping filters during the transition period. This property can also be used to design boundary filters. These are filters that are used at the boundary regions of a signal with finite duration, e.g., as in images. These filters have no overlap beyond the boundaries of the signal. Observe that the basic structure is analogous to the lifting scheme or ladder structure in [31]–[33] or [34]. The zero-delay matrices alone would only allow us to design filter banks with the minimum system delay. To obtain a more general formulation, which also includes

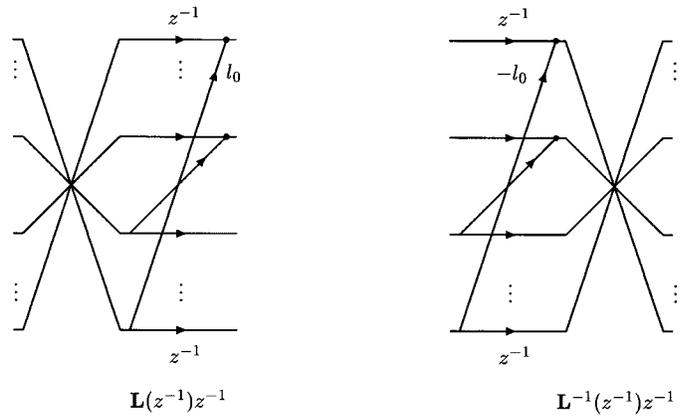


Fig. 3. Structure of the maximum-delay matrices.

orthogonal filter banks, we could additionally use paraunitary matrices, as in [16]. This would allow a range of system delays up to the standard system delay of orthogonal filter banks. However, to obtain a more general formulation, with possible system delays even higher than the standard delay, a different matrix form is proposed. However, most importantly, the following matrix type proves to be very convenient for the design of time-varying filter banks.

Maximum-Delay Matrices: These also increase the filter length but, especially, the system delay. They result by replacing z by z^{-1} in $\mathbf{L}_i(z)$. Therefore, a multiplication with z^{-1} is necessary to obtain a causal system

$$\mathbf{L}_i(z^{-1}) \cdot z^{-1}.$$

The inverse also needs this multiplication for causality

$$\mathbf{L}_i^{-1}(z^{-1}) \cdot z^{-1}$$

which means this matrix type together with its inverse and suitable delays for causality result in a delay of z^{-2}

$$\mathbf{L}_i(z^{-1}) \cdot z^{-1} \cdot \mathbf{L}_i^{-1}(z^{-1}) \cdot z^{-1} = z^{-2}. \quad (13)$$

Their structure can be seen in Fig. 3.

These two types can be used to design a wide range of filter banks. They range from filter banks with the minimum possible delay over orthogonal filter banks to filter banks with a maximum delay. Maximum delay filter banks can be seen as time-reversed minimum delay filter banks.

V. FILTER BANK DESIGN

For the design of a filter bank, several of the zero-delay and maximum-delay matrices are taken together in a product to form the filter matrix $\mathbf{F}(z)$. The total number of zero-delay and maximum-delay matrices determines the degree of the resulting filter matrix and, hence, the filter length. The number of maximum-delay matrices determines the system delay. Let ν be the number of zero-delay matrices and μ the number of maximum-delay matrices. To get the maximum degree of freedom in the design process, a diagonal coefficient matrix \mathbf{D} is needed.

$$\mathbf{D} = \text{diag}(d_0, \dots, d_{M-1}).$$

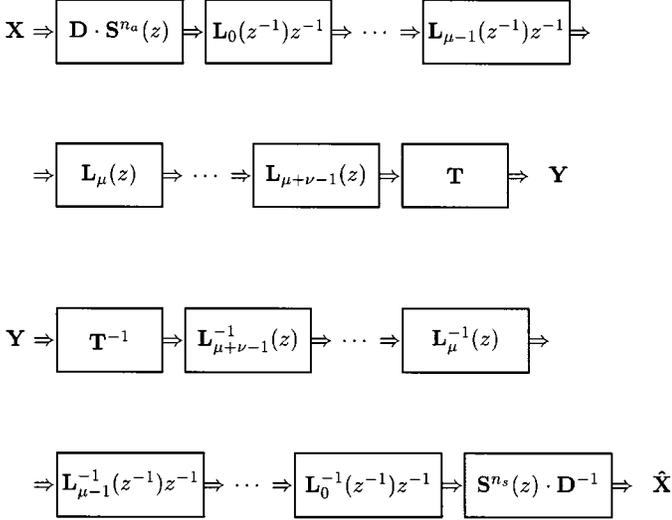


Fig. 4. Block diagram of the filter bank consisting of zero-delay and maximum-delay matrices. The analysis filter bank is above, the synthesis filter bank below.

This can be interpreted as an initialization matrix and leads to the following product for the filter matrix $\mathbf{F}(z)$:

$$\mathbf{F}(z) = \prod_{j=\nu-1}^0 \mathbf{L}_{\mu+j}(z) \cdot \prod_{i=\mu-1}^0 (\mathbf{L}_i(z^{-1}) \cdot z^{-1}) \cdot \mathbf{D}. \quad (14)$$

A block diagram of this structure can be seen in Fig. 4. Since each maximum-delay matrix contributes a factor of z^{-2} to the overall delay (13), the resulting system delay is $n_d = 2\mu M + M - 1 - n_a - n_s$. The exact filter length also depends on the coefficients of the \mathbf{L}_i matrices. Since each matrix \mathbf{L} increases the degree of the matrix \mathbf{F} and, hence, of the polyphase matrix by one, the filter length is approximately $N \approx (\mu + \nu) \cdot M$.

As an example, consider an MDCT type filter bank as described in [5] and [20]. This is an orthogonal filter bank with single overlap between neighboring windows, i.e. the maximum filter length is $2M$, and the system delay is $n_d = 2M - 1$. It is obtained if only one zero-delay matrix and one maximum-delay matrix is used, i.e. $\nu = \mu = 1$ and if $n_a = n_s = M/2$ in (6) and (7) is chosen. Furthermore, to obtain orthogonal filters for this case, it is necessary to restrict the resulting prototype filters to be symmetric or linear phase. The inverse for the synthesis, with a suitable delay for causality, is

$$z^{-2\mu} \mathbf{F}^{-1}(z) = \mathbf{D}^{-1} \cdot \prod_{i=0}^{\mu-1} (\mathbf{L}_i^{-1}(z^{-1}) \cdot z^{-1}) \cdot \prod_{j=0}^{\nu-1} \mathbf{L}_{\mu+j}^{-1}(z). \quad (15)$$

The filter design now consists of the following steps: First, we have to specify the length N of the analysis and synthesis prototype filters and the desired system delay n_d . From the latter, we obtain the values of μ and n_a, n_s according to (3) and (8). The necessary number of zero-delay matrices is then determined by the filter length.

Furthermore, in order to obtain causal filters, i.e., we have to make sure that the first blocks of the cascade $\mathbf{L}_0(z^{-1})z^{-1} \cdot \mathbf{D}$

$\mathbf{S}^{n_a}(z)$, which are implemented together, are also causal. This means that some coefficients l of $\mathbf{L}_0(z)$ have to be set to zero if $n_a > M/2$.

The coefficients of the simple matrices determine the frequency responses of the filter bank. They can be obtained, e.g., with the optimization described in [16]. Note that the minimum possible delay can be obtained for $\mu = 0$ and $n_0 = n_a = n_s = 0$, resulting in a system delay equal to the block-delay $n_d = M - 1$. The maximum possible delay is obtained for $\nu = 0$. It is higher than for the paraunitary case. The system delay of orthogonal or paraunitary filter banks is obtained with $\nu = \mu$ and $n_a = n_s = M/2$.

That the cascade of maximum- and zero-delay matrices is a complete representation of effectively all FIR cosine modulated filter banks, and that the ordering of the maximum-delay-matrices and the zero-delay-matrices in the product is not important, can be seen in the proof of effective completeness in the Appendix.

A. Symmetries for Equal Magnitude Responses

In many applications, it is desirable to have identical magnitude responses for the analysis and synthesis filters, e.g., in audio coders where it is important to have narrow analysis filters for efficient redundancy reduction and narrow synthesis filters for effective application of psycho-acoustic models for the irrelevance reduction. This symmetry is inherent in orthogonal filter banks, where analysis and synthesis filters are time reversed versions of each other. This is, in general, not the case for biorthogonal filters.

We here show that the presented filter bank, for the shown type of modulating function (DCT-IV), can be designed such that it has this symmetry property even in the case of a low system delay. Identical magnitude responses are obtained if the baseband impulse responses for analysis and synthesis are identical, except for the sign $h'(n) = \pm h(n)$ (see also [17] and [19]). We first derive the general relationship between analysis and synthesis polyphase components and then reduce the number of free variables for optimization in order to obtain identical magnitude responses for the analysis and synthesis filters.

Using the fact that the bi-diagonal structure of $\mathbf{F}(z)$ allows us to invert the matrix by inversion of 2×2 submatrices, we obtain

$$\begin{aligned} & \begin{bmatrix} [\mathbf{F}]_{k,k} & [\mathbf{F}]_{k,M-1-k} \\ [\mathbf{F}]_{M-1-k,k} & [\mathbf{F}]_{M-1-k,M-1-k} \end{bmatrix}^{-1} \\ &= \frac{\begin{bmatrix} [\mathbf{F}]_{M-1-k,M-1-k} & -[\mathbf{F}]_{k,M-1-k} \\ -[\mathbf{F}]_{M-1-k,k} & [\mathbf{F}]_{k,k} \end{bmatrix}}{[\mathbf{F}]_{k,k} \cdot [\mathbf{F}]_{M-1-k,M-1-k} - [\mathbf{F}]_{M-1-k,k} \cdot [\mathbf{F}]_{k,M-1-k}}. \end{aligned} \quad (16)$$

Since the synthesis consists of FIR filters, the determinant in the denominator is a constant delay and a factor. Comparing (16) with (9) and (10) shows that the polyphase representations of the prototypes $P'_k(-z^2)$ and $P_k(-z^2)$, and, hence, the prototypes, are equal up to a factor and a delay if that determinant is the same for all the submatrices. Since the determinant of the submatrices of $\mathbf{L}(z)$ and $\mathbf{L}(z^{-1})$ is constant (-1), this condition is fulfilled if the determinant of the submatrices of \mathbf{D} are constant, i.e., $d_{M-1-k} = \pm 1/d_k$. That means, if we chose \mathbf{D}

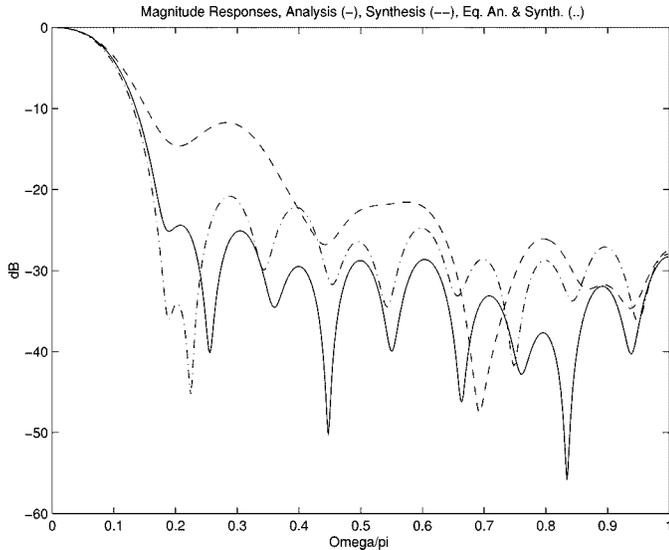


Fig. 5. Magnitude responses of an analysis prototype (solid line) and a synthesis prototype (dashed line) for eight bands, a system delay of seven samples (minimum delay), and a filter length of 20 taps. For comparison, the magnitude response for a filterbank with the same specifications but with equal magnitude responses for analysis and synthesis is also shown (dash-dotted line).

with this property, the the resulting prototype filters will automatically be identical for analysis and synthesis, up to a factor, and, hence, will have the same magnitude response.

B. Efficient Implementation

An efficient implementation of the filter bank can be obtained by building the cascade with the simple matrices \mathbf{L} and the shift matrix \mathbf{S} and to take an efficient algorithm for the $M \times M$ DCT-IV. The number of multiplications and additions for a fast DCT are of order $M \log(M)$ [2]. The number of multiplications necessary for the filter matrices and \mathbf{D} is less than or equal to

$$\text{number of mult's: } M + (\mu + \nu) \cdot M/2 = M + N/2.$$

This number is equal to the number of unconstrained variables, which shows that the number of multiplications is minimal. This is slightly more than half the number of multiplications necessary when implementing the polyphase filters directly. The approximate number of additions for the filter matrices is

$$\text{number of add's: } (\mu + \nu) \cdot M/2.$$

Note that the coefficients for the synthesis matrices result from sign flipping of the coefficients of the \mathbf{L} matrices, and that the input for the multipliers is the same as for the analysis (except for the matrix \mathbf{D}), which means that they provide perfect reconstruction even if they are implemented with low precision arithmetic, as long as the sign flipping is exact.

C. Design Examples

Fig. 5 shows an example of a minimum-delay filter bank with unequal analysis and synthesis prototypes. The coefficients of the resulting cascades or structures were obtained with the optimization algorithm described in [16]. Using higher weights for the analysis magnitude response than for the synthesis, the anal-

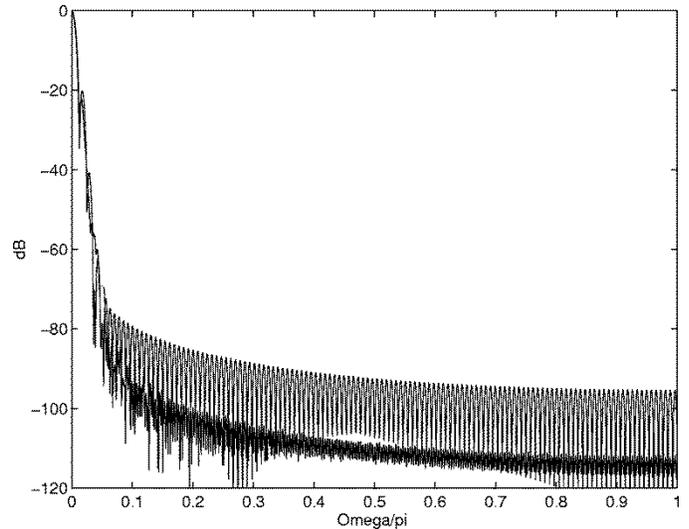


Fig. 6. Magnitude responses of the baseband low-delay prototype (the lower curve) with length 512 taps, identical for the analysis and synthesis filter bank, compared with an orthogonal filter bank (upper curve) with length 256 taps. Both have 128 bands and a system delay of 255 samples.

ysis has a higher stopband attenuation. This analysis magnitude response also has a higher stopband attenuation than for the case of equal magnitude responses for the analysis and synthesis, as can be seen in the figure.

Fig. 6 shows an example of a filter bank with a low system delay, where the symmetry condition for identical magnitude responses for analysis and synthesis was imposed. It is compared with an orthogonal filter bank with a standard system delay. Both were designed with the presented design method. The parameters of the low-delay filter bank are $n_a = n_s = M/2$, $\nu = 3$, $\mu = 1$, and for the standard delay filter bank $n_a = n_s = M/2$, $\nu = 1$, $\mu = 1$. The latter is an MDCT-type filter bank, which is widely used in audio coding. Both filter banks have 128 bands and a system delay of 255 samples, but the orthogonal filter bank is restricted to a filter length of 256 taps due to the given system delay. The low delay filter bank has a filter length of 512 taps and, as a result, has an about 20 dB higher stopband attenuation, as can be seen from Fig. 6.

Figs. 7 and 8 show a similar design for a low-delay filter bank ($n_a = n_s = M/2$, $\nu = 3$, $\mu = 1$, with symmetry condition) but with 1024 bands, length 4096, and a system delay of 2047 samples. This filter will also be used in an example for a time-varying filter bank, which is described in Section VI.

VI. TIME VARIANCE

The decomposition of the filter matrices into zero-delay and maximum-delay matrices now provides a convenient framework for the design of time-varying filter banks. To extend the cosine-modulated filter bank considered in the last sections to time variation, its polyphase matrices must have time-varying entries. In order to express this time dependency, the parameter m , denoting the time instance at the lower sampling rate, is introduced. Thus, $\mathbf{E}(z)$ becomes $\mathbf{E}(z, m)$, and $\mathbf{L}_i(z)$ becomes $\mathbf{L}_i(z, m)$, with

$$\mathbf{L}_i(z, m) := \mathbf{J} + \text{diag}(l_0^i(m), \dots, l_{M-1}^i(m), 0, \dots, 0) \cdot z^{-1}.$$

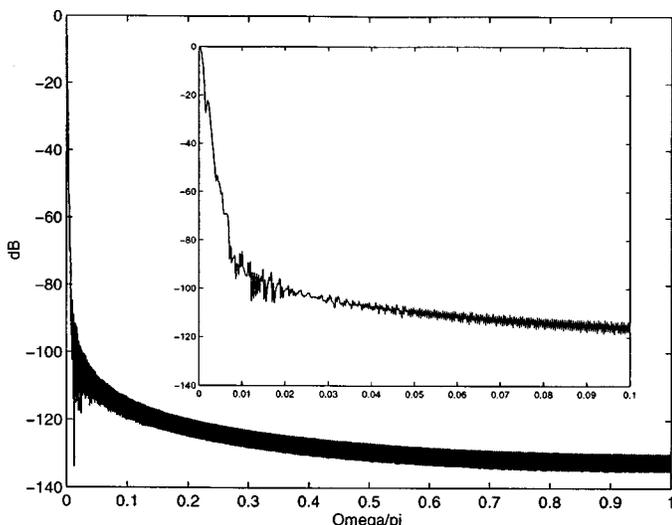


Fig. 7. Magnitude response of the baseband low-delay prototype of a filter bank with 1024 bands length 4096 taps, and system delay of 2047 samples, identical for analysis and synthesis.

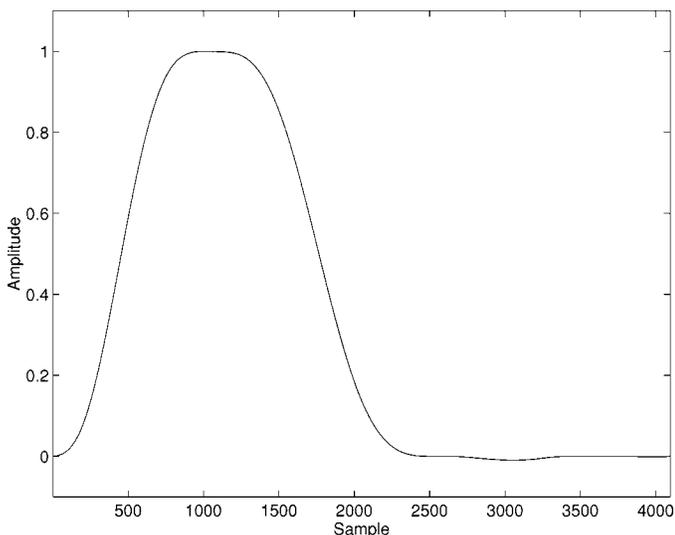


Fig. 8. Impulse response of the baseband low-delay prototype of Fig. 7.

The additional parameter requires a computation that is different from the time-invariant case. Observe that if a signal first passes a time-varying system or matrix $\mathbf{F}(z, m)$ and then a delay z^{-1} , the output is the same as if the signal is first delayed and then passes the system or matrix at the state of the previous time step. This is an important observation for the treatment of time-varying systems in the z -domain and can be written as (see also [26])

$$z^{-1} \cdot \mathbf{F}(z, m) = \mathbf{F}(z, m-1) \cdot z^{-1}. \quad (17)$$

For the computation of the synthesis polyphase matrix for perfect reconstruction, the following observation is useful. A delay between a time-varying matrix and its inverse results in a time shift of the inverse because the input to the inverse matrix is now a delayed version of the original matrix. This can be seen using (17)

$$\begin{aligned} z^{-d} \cdot \mathbf{I} &= z^{-d} \cdot \mathbf{F}^{-1}(z, m) \cdot \mathbf{F}(z, m) \\ &= \mathbf{F}^{-1}(z, m-d) \cdot z^{-d} \cdot \mathbf{F}(z, m). \end{aligned} \quad (18)$$

Because of their special structure, the inverses of the filter matrices are very simple, even in the time-varying case. Thus, we can write

$$\mathbf{L}_i^{-1}(z, m) := \mathbf{J} - \text{diag}(0, \dots, 0, l_{M-1}^i(m), \dots, l_0^i(m)) \cdot z^{-1}$$

and because of (18)

$$\mathbf{L}^{-1}(z^{-1}, m-1) \cdot z^{-1} \cdot \mathbf{L}(z^{-1}, m) \cdot z^{-1} = z^{-2}.$$

which shows that the time index m has to be lowered by 1 for the inverse of the maximum delay matrices. The time-varying analysis filter bank can now be expressed as

$$\begin{aligned} \mathbf{E}(z, m) &= \mathbf{T}(m) \prod_{j=\nu-1}^0 \mathbf{L}_{\mu+j}(z, m) \\ &\cdot \prod_{i=\mu-1}^0 (\mathbf{L}_i(z^{-1}, m) \cdot z^{-1}) \cdot \mathbf{D}(m) \cdot \mathbf{S}^{n_a}. \end{aligned} \quad (19)$$

The inverse for the synthesis, with a suitable delay for causality, is

$$\begin{aligned} \mathbf{R}(z, m) &= \mathbf{S}^{n_s} \cdot \mathbf{D}^{-1}(m) \cdot \prod_{i=\mu-1}^0 \\ &\cdot (\mathbf{L}_{\mu-1-i}^{-1}(z^{-1}, m-1-2i) \cdot z^{-1}) \\ &\cdot \prod_{j=0}^{\nu-1} \mathbf{L}_j^{-1}(z, m) \cdot \mathbf{T}^{-1}(m). \end{aligned} \quad (20)$$

Observe the index $\mu-1-i$ of the maximum-delay matrices and the time indices in (20). The signal can be viewed as passing the matrices from right to left. Since the zero-delay matrices do not introduce any additional system delay, their time index remains m . The maximum-delay matrices are associated with an additional delay of z^{-2} [see (13)]. For this reason, the time index after each of these delays has to be lowered by two, according to the relationship in (18), in order to yield perfect reconstruction.

Keeping the total delay of the filter bank constant, n_a , n_s , and μ are time invariant. In addition, considering ν as time invariant, the filter length N remains the same for all time steps m . When switching between filter banks with different system delays, the impulse responses of the filter bank with the lower delay will be zero padded at the beginning of the impulse response to yield the same overall delay. This zero padding is done implicitly by keeping n_0 constant and always using the same number μ of maximum-delay matrices (with zero-valued coefficients if needed). However, different system delays are still useful for a variety of applications such as the reduction of “pre-echo” in audio coding, as will be seen in an example. Note that here, we provide a very general approach that accommodates many different ways to switch the analysis and synthesis filters. The resulting direct form filters at time m are now determined as follows. The filter coefficients of the time-varying analysis filters at time m are the weights for the input samples that are used to obtain the subband samples at time m . They are

obtained by rewriting $\mathbf{P}_a(z, m)$ with the help of (17) such that it has the form

$$[\mathbf{E}(z, m)]_{k, n} := \sum_{m=0}^{\infty} h_k(mM + n, m)z^{-m}$$

i.e., all delays are on the right side of the coefficients. In this way, all coefficients of $\mathbf{P}_a(z, m)$ contribute to the subband samples at time m . The synthesis filter coefficients at time m are the weights for the output samples of the synthesis using input samples at time m . These coefficients are obtained by rewriting $\mathbf{P}_s(z, m)$ such that it has the form

$$[\mathbf{R}(z, m)]_{n, k} := \sum_{m=0}^{\infty} z^{-m} g_k(mM + M - 1 - n, m)$$

i.e. all delays are on the left side of the coefficients so that the coefficients correspond to one synthesis input vector at time m .

A. Changing the Number of Bands

If we want to use the above formulation for switching between different numbers of bands, we must provide a formalism to treat input signal vectors of different lengths. Assume any two different numbers of bands M_1 and M_2 , such that $M_1 > M_2$. First, consider the time-invariant case, and define

$$\mathbf{X}'(z) := \mathbf{S}^{n_a}(z) \cdot \mathbf{X}(z) \quad (21)$$

which is the input vector for the filter matrices. \mathbf{X}' and \mathbf{X} are of length M_1 , and $\mathbf{S}^{n_a}(z)$ is of size $M_1 \times M_1$. For the case of M_2 subbands, $M_1 - M_2$ zeros are placed in $\mathbf{X}(z)$ such that $\mathbf{X}'(z)$ has length M_1 and zeros around the center

$$\mathbf{X}'(z) = [X'_0(z), \dots, X'_{M_2/2-1}(z), \underbrace{0, \dots, 0}_{M_1-M_2}, X'_{M_2/2}(z), \dots, X'_{M_2-1}(z)]$$

e.g., by computing $\mathbf{X}(z) = \mathbf{S}^{-n_a}(z) \cdot \mathbf{X}'(z)$ so that the number of input samples in one block is M_2 .

The filter matrices in the mode for M_2 bands are now size $M_1 \times M_1$ matrices of the form

$$\mathbf{L}_i(z, m) = z^{-1} \cdot \mathbf{J} + \mathbf{diag}(l^i(m), \dots, l^i_{M_2/2-1}(m), \underbrace{0, \dots, 0}_{M_1-M_2/2})$$

The matrix \mathbf{D} is filled with ones in the center, for invertibility:

$$\mathbf{D} = \mathbf{diag}(d_0, \dots, d_{M_2/2-1}, 1, \dots, 1, d_{M_2/2}, \dots, d_{M_2-1}).$$

Because the filter matrices have nonzero coefficients only on the diagonal and antidiagonal, the introduced zeros also appear at the analysis transform matrix, which is now named $\mathbf{T}_a(m)$ instead of \mathbf{T} . Since the positions of the zeros are known, they do not need to be processed further and can be omitted for the computation of the transform, so that an analysis transform matrix of size $M_2 \times M_1$ can be used which results from splitting the $M_2 \times M_2$ transform matrix into a left and right half (\mathbf{T}_l and \mathbf{T}_r , respectively) and inserting $M_1 - M_2$ columns of zeros in the middle. The synthesis transform $\mathbf{T}_s(m)$ is then an $M_1 \times M_2$

matrix that results from splitting the $M_2 \times M_2$ inverse transform matrix into an upper and lower half (\mathbf{T}_u and \mathbf{T}_d , respectively) and inserting $M_1 - M_2$ rows of zeros in the middle.

$$\mathbf{T}_a = [\mathbf{T}_l \mid \mathbf{0} \mid \mathbf{T}_r] M_2, \quad \mathbf{T}_s = \left[\begin{array}{c} \mathbf{T}_u \\ \mathbf{0} \\ \mathbf{T}_d \end{array} \right] M_1. \quad (22)$$

Their product is the $M_1 \times M_1$ matrix

$$\mathbf{T}_s \cdot \mathbf{T}_a = \left[\begin{array}{c|c|c} \mathbf{I} & & \mathbf{0} \\ \hline & \mathbf{0} & \\ \hline \mathbf{0} & & \mathbf{I} \end{array} \right].$$

The so-defined length M_1 signal vectors $\mathbf{X}(z)$ and size $M_1 \times M_1$ filter matrices now represent a filter bank with M_2 bands. Note that the computational complexity is the same as for signal vectors and matrices of size M_2 since only operations with nonzero coefficients need to be computed.

This formulation can now be used for switching between different numbers of bands. The transition between M_1 and M_2 bands can be done by inserting or removing zeros in $\mathbf{x}(m)$ such that they appear or disappear together after the shift matrix $\mathbf{S}(z)$ in $\mathbf{x}'(m)$. Therefore, if $n_a > M_2/2$, an intermediate number of input samples has to appear in $\mathbf{x}(m)$ at the beginning of the transition. For example, with $n_a = M_1/2$ and the beginning of the transition at time m_0 , the input vectors have the form

$$\begin{aligned} \mathbf{x}(m_0) &= [x_0(m_0), \dots, x_{M_1/2+M_2/2-1}(m_0), 0, \dots, 0]^T \\ \mathbf{x}(m_0+1) &= [0, \dots, 0, x_0(m_0+1) \\ &\quad \dots, x_{M_2-1}(m_0+1), 0, \dots, 0]^T \end{aligned}$$

and so on. Observe that at time $m_0 - 1$, the signal vector contains M_1 samples; at time m_0 , it contains $M_1/2 + M_2/2$ samples; and at time $m_0 + 1$, it contains M_2 samples. They are consecutive pieces of the signal, i.e., (1) is not valid for the case of changing numbers of bands.

Now that the $M_1 - M_2$ zeros appear together after the shift matrix, consider how the signal with the zeros pass the filter matrices. Here, another advantage of the cascade of zero-delay and maximum-delay matrices becomes apparent. Their coefficients can be chosen such that they keep the set of zeros together throughout the cascade. The maximum-delay matrices delay them all by one block, and the zero-delay matrices do not delay them at all. This means they arrive together at the transform matrix \mathbf{T} so that no intermediate number of bands is needed and a critical sampling is guaranteed also during the transition. If the cascade contained paraunitary matrices, e.g., as in [16] or [25], they would delay the zeros differently, depending on their position, and hence, they would arrive at the transform matrix \mathbf{T} at different times. That would lead to intermediate numbers of bands or a noncritical sampling during the transition.

When the zeros appear at the transform matrix \mathbf{T} , it is switched to the form (22) with a suitable number of zeros in it. The reverse process is used for switching to a higher number of bands.

If $M_2 = 0$ is chosen, the switching can be used to obtain filters for the boundary regions of a signal in order to process

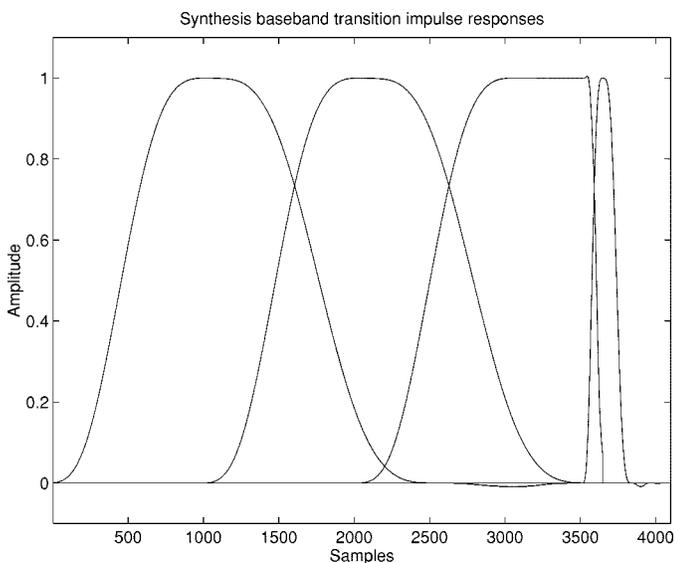


Fig. 9. Synthesis transition baseband impulse responses of a time-varying low-delay filter bank for a switch from 1024 bands and length 4096 filters to 128 bands with length 512 filters.

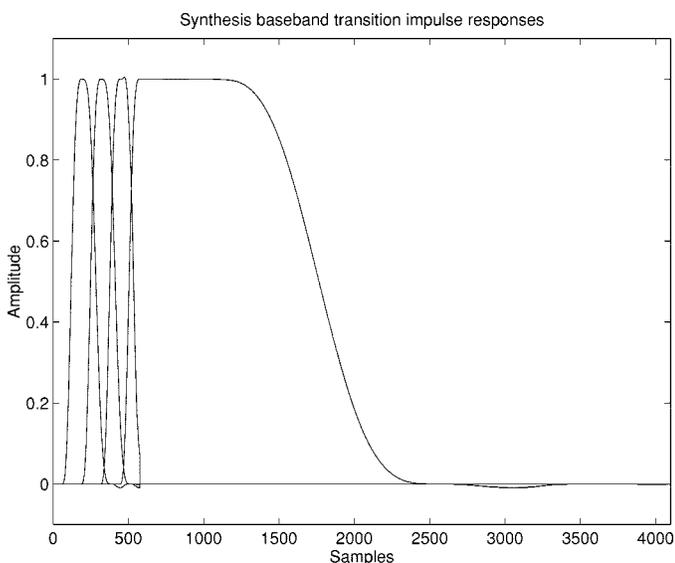


Fig. 10. Synthesis transition baseband impulse responses of the time-varying low-delay filter bank for a switch from 128 bands, length 512, to 1024 bands, length 4096.

signals with finite support. The beginning of a signal can then be treated as a switch from zero bands, and vice versa, for the end of the signal. In this way, the analysis filter bank produces the same number of samples as the finite input signal contains. The signal can then still be completely reconstructed, including the boundary regions.

B. Audio Coding Examples

Figs. 9 and 10 show an example of the prototype filter impulse responses of a filter bank with a low system delay, which is switched from 1024 to 128 bands and from 128 to 1024 bands, respectively. This results in an intermediate input block size of 576 samples during transition. The synthesis prototype impulse responses for different times are shown in their actual

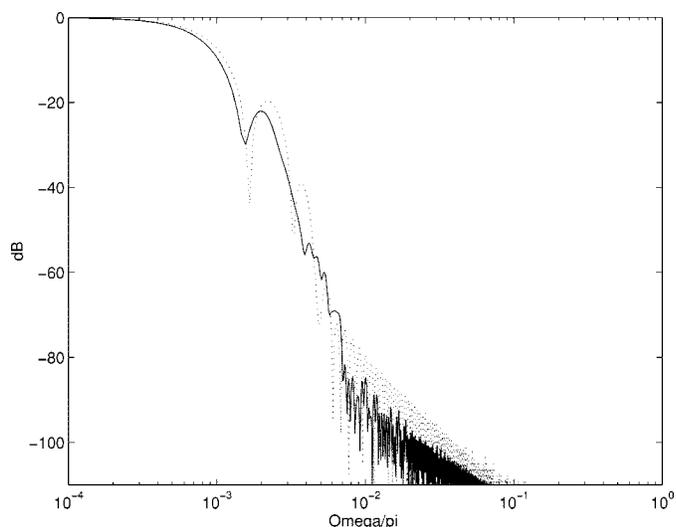


Fig. 11. Comparison of the magnitude responses of the baseband prototype of the 1024-band orthogonal MPEG filter bank of length 2048 (dotted line) to the low-delay filter bank of the same system delay and length 4096.

relative position. The filters for the steady-state case have a length of 4096 and 512 taps, respectively, and a system delay of 2047 and 255 samples, respectively, with $\mu = 1$, $\nu = 3$, $n_a = n_s = 1024/2$. The symmetry condition of Section V-A was used for their design, i.e., the analysis prototype is identical to the synthesis prototype filter. The 1024 band filter can also be seen in Fig. 8. The transition filters were designed such that they are also modulated filters by limiting the overlap between filters of different modes to one block. Switching the analysis from 1024 to 128 bands and from 128 to 1024 bands is obtained by time reversing Figs. 9 and 10, respectively.

The next example is for the “MPEG-2 advanced audio coding” audio coder [21], which is a proposed standard, targeted at delivering CD quality sound at a bit rate of 64 kb/s mono. It has a filter bank with two modes and uses the switching algorithm described in [20]. Both modes have orthogonal filters with a standard system delay: One has 1024 bands and a filter length of 2048 taps (i.e. delay 2047), and the other has 128 bands and a length of 256 taps (delay 255). The mode with 128 bands is used for signal transients to reduce pre-echo, which, however, is still audible. A time-varying low-delay filter bank was then implemented in this coder by replacing its built-in filter bank. The low-delay filter bank was designed such that the 1024 band mode has the same system delay but a higher attenuation in the transition and stopband region to improve the coding efficiency. It has 1024 bands, a filter length of 4096 taps, and a system delay of 2047 samples (see Figs. 7 and 8). The 128-band mode was designed such that it has about the same magnitude response as the original “sine” filter but a lower system delay to reduce pre-echo. It has 480 taps and a system delay of 191 samples. The filters of both modes were designed with the symmetry condition of Section V-A to obtain identical prototypes for analysis and synthesis, as the coders original filters. Fig. 11 shows the improved magnitude response of the 1024-band baseband prototype for analysis and synthesis in comparison with the original 1024-band mode “Dolby” filter on a logarithmic frequency scale. Fig. 12 shows reduced

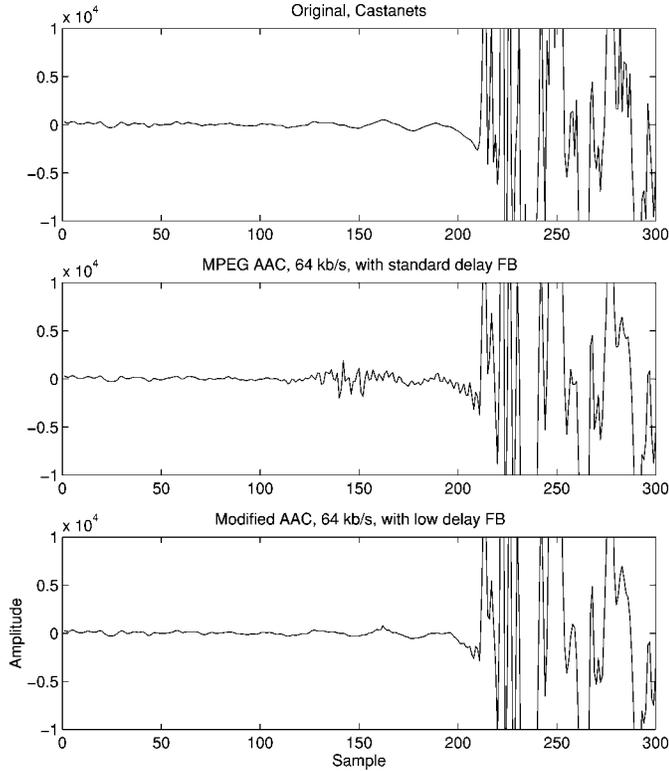


Fig. 12. MPEG 2 advanced audio coding (AAC) audio coder with the original filter bank with a standard system delay of 255 samples in the 128-band mode leads to audible pre-echo (center window), whereas the same coder with a filter bank with a low system delay of 191 samples has no audible pre-echo (lower window). The original is in the top window.

pre-echo, which results from using the low-delay filter bank. The upper part shows a segment of the original “castanets” signal, the middle part is the same segment, coded and decoded with the original filter bank, and the lower part is with the low-delay filter bank. The pre-echo of the original coder is still audible, whereas it is not audible after coding and decoding with the new filter bank. Listening tests with this modified MPEG-AAC coder confirmed that the castanets’ signal was rated better with the low-delay filter bank.

VII. CONCLUSION

In this paper, we have presented a new formulation for critically downsampled time-varying cosine-modulated filter banks with perfect reconstruction. As was seen, the presented formulation has the following advantages.

- The system delay can be specified in terms of single samples of the input sampling rate. This allows for fine-tuned compromises between filter quality, filter length, and system delay.
- A low system delay is possible, down to the block-delay of $M - 1$ samples, independent of the filter length.
- Equal-magnitude response filters for analysis and synthesis can be easily designed.
- A proof of effective completeness is provided.
- Perfect reconstruction results even for time-varying coefficients or nonlinear operations like rounding.

- The number of bands can be changed during signal processing while maintaining PR, critical sampling, and temporal overlap between filters.
- Finite-length signals can be processed without an overhead in the number of subband samples.
- The implementation has low complexity and is simple to design.

These advantages make the proposed technique very useful in practical applications like audio coding, where it can be used in existing audio coders to reduce pre-echo, as shown in the examples.

APPENDIX EFFECTIVE COMPLETENESS

In this section, we prove that all FIR cosine-modulated filter banks with perfect reconstruction that lead to bi-diagonal filter matrices $\mathbf{F}(z)$ and $\mathbf{F}^{-1}(z)$, for analysis and synthesis, respectively, and whose prototype impulse responses $h(n)$ and $h'(n)$ are contiguous, can be represented by the factorization given in (14) and (15). The proof is constructive, presenting an iterative algorithm for the extraction of the zero-delay and maximum-delay matrices from the filter matrices. The filter matrices can be written as polynomials of matrices, with a filter length of $N \leq KM$

$$\mathbf{F}(z) = \sum_{m=0}^{K-1} \mathbf{f}_a(m) \cdot z^{-m}$$

$$z^{-2\mu} \mathbf{F}^{-1}(z) = \sum_{m=0}^{K-1} \mathbf{f}_s(m) \cdot z^{-m}$$

where $K - 1$ is the degree of the polynomials. Perfect reconstruction results in

$$z^{-2\mu} \cdot \mathbf{I} = \mathbf{F}(z) \cdot z^{-2\mu} \mathbf{F}^{-1}(z)$$

$$= \sum_{m=0}^{2K-2} z^{-m} \sum_{i+j=m} \mathbf{f}_a(i) \cdot \mathbf{f}_s(j). \quad (23)$$

Now, consider the matrices for certain exponents m . If $2\mu < 2K - 3$, then for $m = 2K - 2$, it follows that

$$\mathbf{0} = \sum_{i+j=m} \mathbf{f}_a(i) \cdot \mathbf{f}_s(j) = \mathbf{f}_a(K-1) \cdot \mathbf{f}_s(K-1) \quad (24)$$

and for $m = 2K - 3$

$$\mathbf{f}_a(K-1) \cdot \mathbf{f}_s(K-2) + \mathbf{f}_a(K-2) \cdot \mathbf{f}_s(K-1) = \mathbf{0}. \quad (25)$$

Since $h(n)$ and $h'(n)$ are contiguous nonzero filters and $\mathbf{f}_a(K-1)$ and $\mathbf{f}_s(K-1)$ contain the end of the nonzero part of the baseband impulse responses, $\mathbf{f}_a(K-2)$ and $\mathbf{f}_s(K-2)$ are diagonal or anti-diagonal matrices with full rank (also compare with (11)). From (24), we obtain

$$\text{rank}(\mathbf{f}_a(K-1)) + \text{rank}(\mathbf{f}_s(K-1)) \leq M$$

and from (25)

$$\text{rank}(\mathbf{f}_a(K-1)) = \text{rank}(\mathbf{f}_s(K-1)).$$

it follows that

$$\text{rank}(\mathbf{f}_a(K-1)) = \text{rank}(\mathbf{f}_s(K-1)) \leq M/2.$$

Thus, since $\mathbf{f}_a(K-1)$ and $\mathbf{f}_s(K-1)$ are diagonal or anti-diagonal matrices, the number of their nonzero elements is less than or equal to $M/2$. Since the prototype impulse responses are contiguous, these nonzero elements must also be contiguous on the diagonal or anti-diagonal, bordering on the right or left side of the matrix. The relationships (24) and (25) still hold true if they are multiplied by $(\mathbf{f}_s(K-2))^{-1} \cdot \mathbf{J}$ from the right and $\mathbf{J} \cdot (\mathbf{f}_a(K-2))^{-1}$ from the left, respectively. This leads to the construction or extraction of a zero-delay matrix

$$\begin{aligned} \mathbf{L}_i(z) &= \mathbf{J} \cdot (\mathbf{I} + z^{-1} \cdot \mathbf{f}_a^{-1}(K-2) \cdot \mathbf{f}_a(K-1)) \\ \mathbf{L}_i^{-1}(z) &= (\mathbf{f}_s(K-1)\mathbf{f}_s^{-1}(K-2) + z^{-1} \cdot \mathbf{I}) \cdot \mathbf{J}. \end{aligned}$$

Applying the above equations to the filter matrices, we obtain new filter matrices $\mathbf{F}'(z)$ and $\mathbf{F}'^{-1}(z)$

$$\mathbf{F}'(z) := \mathbf{F}(z) \cdot \mathbf{L}_i^{-1}(z), \quad \mathbf{F}'^{-1}(z) := \mathbf{L}_i(z) \cdot \mathbf{F}^{-1}(z)$$

that have a degree reduced by 1 while the system delay remains unchanged. The index i of the zero-delay matrix starts with $i = \nu - 1$ and is reduced by 1 for each step of the iteration. This corresponds to extracting the last matrices of the cascade first. Matrices $\mathbf{F}'_a(z)$ and $\mathbf{F}'_s(z)$ are, again, filter matrices of a cosine-modulated filter bank since \mathbf{L}_i has the suitable bi-diagonal form. They again result in an FIR filter bank with FIR inverse since \mathbf{L}_i has an FIR inverse. The condition for a further reduction of the polyphase filter length K is that $\mathbf{F}'_a(z)$ and $\mathbf{F}'_s(z)$ lead, again, to contiguous prototype impulse responses in order to obtain the next full-rank $\mathbf{f}_a(K-2)$ and $\mathbf{f}_s(K-2)$. This is usually the case. If not, the objectionable zeros in it can be replaced by some small number ε . In this way, an arbitrary close approximation is possible. The process of reducing the degree can be continued, whereas $2K-3 > 2\mu$, and thus, all the zero-delay matrices are obtained.

Here, it can also be seen that the effect of the zero-delay matrices is the extension of the impulse response to later times m . Similarly, the effect of the maximum-delay matrices is the extension of the impulse response to earlier times m and its shifting to later times. This is why the same extraction and reduction can now be done for the other side, i.e., the beginning of the impulse response. If $2\mu > 1$, maximum-delay matrices can be extracted. For $m = 0$ in (23), we obtain

$$\mathbf{f}_a(0) \cdot \mathbf{f}_s(0) = \mathbf{0} \quad (26)$$

and for $m = 1$

$$\mathbf{f}_a(0) \cdot \mathbf{f}_s(1) + \mathbf{f}_a(1) \cdot \mathbf{f}_s(0) = \mathbf{0}. \quad (27)$$

At this point, we can conclude that

$$\text{rank}(\mathbf{f}_a(0)) = \text{rank}(\mathbf{f}_s(0)) \leq M/2$$

and that $\mathbf{f}_a(1)$ and $\mathbf{f}_s(1)$ have full rank. Defining

$$\begin{aligned} \mathbf{L}_i(z^{-1}) \cdot z^{-1} &= \mathbf{J} \cdot (\mathbf{f}_a^{-1}(1) \cdot \mathbf{f}_a(0) + z^{-1} \cdot \mathbf{I}) \\ \mathbf{L}_i^{-1}(z^{-1}) \cdot z^{-1} &= (\mathbf{f}_s(0) \cdot \mathbf{f}_s^{-1}(1) + z^{-1}\mathbf{I}) \cdot \mathbf{J} \end{aligned}$$

then

$$\begin{aligned} \mathbf{F}'(z) &= \mathbf{F}(z) \cdot \mathbf{L}_i^{-1}(z^{-1}) \quad \text{and} \\ \mathbf{F}'^{-1}(z) &= \mathbf{L}_i(z^{-1}) \cdot \mathbf{F}^{-1}(z) \end{aligned}$$

are causal. The iteration starts with $i = \mu - 1$. For the new filter matrices, K and μ are both reduced by 1. As above, the resulting matrices \mathbf{F}' and \mathbf{F}'^{-1} are again filter matrices of an FIR cosine-modulated filter bank. The process of reducing the degree of $\mathbf{F}(z)$ is continued until $K = 1$. In this way, the maximum-delay matrices are obtained. The matrix which is left at the end of the iteration is \mathbf{D} .

REFERENCES

- [1] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [2] H. Malvar, *Signal Processing with Lapped Transforms*. Norwell, MA: Artech House, 1992.
- [3] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [4] A. Akansu and M. J. T. Smith, *Subband and Wavelet Transforms, Design and Applications*. Boston, MA: Kluwer, 1996.
- [5] J. P. Princen and A. B. Bradley, "Analysis/synthesis filter bank design based on time domain alias cancellation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1153-1161, Oct. 1986.
- [6] H. S. Malvar, "Extended lapped transforms: Properties, applications, and fast algorithms," *IEEE Trans. Signal Processing*, vol. 40, pp. 2703-2714, Nov. 1992.
- [7] T. A. Ramstad and J. P. Tanem, "Cosine modulated analysis synthesis filter bank with critical sampling and perfect reconstruction," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Toronto, Ont., Canada, 1991, pp. 1789-1792.
- [8] R. D. Koilpillai and P. P. Vaidyanathan, "Cosine modulated fir filter banks satisfying perfect reconstruction," *IEEE Trans. Signal Processing*, vol. 40, pp. 770-783, Apr. 1992.
- [9] T. Q. Nguyen and R. D. Koilpillai, "The theory and design of arbitrary-length cosine-modulated filter banks and wavelets, satisfying perfect reconstruction," *IEEE Trans. Signal Processing*, vol. 44, pp. 473-483, Mar. 1996.
- [10] R. A. Gopinath and C. S. Burrus, "Some results in the theory of modulated filter banks and modulated tight frames," *Appl. Comput. Harmon. Anal.*, vol. 2, pp. 303-326, 1995.
- [11] R. A. Gopinath, "Modulated filter banks and wavelets—A general unified theory," in *Proc. IEEE ICASSP*, Atlanta, GA, May 1996, pp. 1585-1588.
- [12] K. Nayebi, T. P. Barnwell III, and M. J. T. Smith, "Low delay FIR filter banks: Design and evaluation," *IEEE Trans. Signal Processing*, vol. 42, no. 11, pp. 24-31, Jan. 1994.
- [13] T. Q. Nguyen, "A class of generalized cosine-modulated filter bank," in *Proc. Int. Symp. Circuits Syst.*, San Diego, CA, 1992, pp. 943-946.
- [14] G. Schuller and M. J. T. Smith, "A general information for modulated perfect reconstruction filter banks with variable system delay," in *Proc. NJIT Symp. Appl. Subbands Wavelets*, Mar. 1994.
- [15] K. Nayebi, T. P. Barnwell, III, and M. J. T. Smith, "On the design of FIR analysis-synthesis filter banks with high computational efficiency," *IEEE Trans. Signal Processing*, vol. 42, pp. 825-834, Apr. 1994.
- [16] G. D. T. Schuller and M. J. T. Smith, "New framework for modulated perfect reconstruction filter banks," *IEEE Trans. Signal Processing*, vol. 44, pp. 1941-1954, Aug. 1996.
- [17] T. Q. Nguyen and P. N. Heller, "Biorthogonal cosine-modulated filter bank," in *Proc. IEEE ICASSP*, vol. 3, Atlanta, GA, May 1996, pp. 1471-1474.
- [18] G. Schuller, "A new factorization and structure for cosine modulated filter banks with variable system delay," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 2, Pacific Grove, CA, Nov. 6, 1996, pp. 1310-1314.
- [19] P. N. Heller, T. Karp, and T. Q. Nguyen, "A general formulation of modulated filter banks," *IEEE Trans. Signal Processing*, vol. 47, pp. 986-1002, Apr. 1999.
- [20] B. Edler, "Coding of audio signals with overlapping block transform and adaptive window functions (in German)," *Frequenz*, pp. 252-256, Sept. 1989.

- [21] M. Bosi *et al.*, "ISO/IEC MPEG-2 advanced audio coding," in *Proc. 101 AES Conv.*, Los Angeles, CA, Nov. 1996.
- [22] K. Nayebi, T. P. Barnwell, III, and M. J. T. Smith, "Analysis-synthesis systems based on time-varying filter bank structures," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 4, Mar. 1992, pp. 947–950.
- [23] C. Herley, J. Kovacević, and M. Vetterli, "Tilings of the time-frequency plane: Construction of arbitrary orthogonal bases and fast tiling algorithms," *IEEE Trans. Signal Processing*, vol. 41, pp. 3341–3359, Dec. 1993.
- [24] C. Herley and M. Vetterli, "Orthogonal time-varying filter banks and wavelet packets," *IEEE Trans. Signal Processing*, vol. 42, pp. 2650–2663, Oct. 1994.
- [25] J. L. Arrowood and M. J. T. Smith, "Exact resolution analysis/synthesis filter banks with time-varying filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 3, Apr. 1993, pp. 233–236.
- [26] S. M. Phoong and P. P. Vaidyanathan, "A polyphase approach to time-varying filter banks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 3, Atlanta, GA, May 1996, pp. 1554–1557.
- [27] I. Sodagar, K. Nayebi, T. P. Barnwell, and M. J. T. Smith, "Time-varying analysis-synthesis systems based on filter banks and post filtering," *IEEE Trans. Signal Processing*, vol. 43, pp. 2512–2524, Nov. 1995.
- [28] R. A. Gopinath and C. S. Burrus, "Factorization approach to unitary time-varying filter bank trees and wavelets," *IEEE Trans. Signal Processing*, vol. 43, pp. 666–680, Mar. 1995.
- [29] R. L. de Queiroz and K. R. Rao, "Variable-block-size lapped transforms," *IEEE Trans. Signal Processing*, vol. 44, pp. 3139–3142, Dec. 1996.
- [30] G. Schuller, "Time-varying filter banks with variable system delay," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 3, Munich, Germany, Apr. 1997, pp. 2469–2472.
- [31] A. A. C. Kalker and I. A. Shah, "Ladder structures for multidimensional linear phase perfect reconstruction filter banks and wavelets," in *Proc. Visual Commun. Image Process.*, 1992, pp. 12–20.
- [32] W. Sweldens, "The lifting scheme: A new philosophy in biorthogonal wavelet constructions," *Proc. SPIE, Wavelet Appl. Signal Image Process. III*, pp. 68–79, 1995.
- [33] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," Preprint, Bell Labs., Lucent Technol., 1996.
- [34] S. Borac and R. Seiler, "Loop group factorization of biorthogonal wavelet bases," Preprint 281, SFB 288, Fachbereich Mathematik, Tech. Univ. Berlin, 1997.



Gerald D. T. Schuller (M'98) received the "Vordiplom" (B.S.) degree in mathematics from the Technical University of Clausthal, Clausthal, Germany, in 1984, the "Vordiplom" and "Diplom" (M.S.) degree in electrical engineering from the Technical University of Berlin, Berlin, Germany, in 1986 and 1989, respectively, and the Ph.D. degree from the University of Hanover, Hanover, Germany, in 1997.

He received a fellowship to study at the Massachusetts Institute of Technology, Cambridge, in 1989 and 1990; was a Research Assistant at the Technical University of Berlin from 1990 to 1992, where he worked on speech coding; a Teaching Assistant at the Georgia Institute of Technology, Atlanta, in 1993, where he worked on low-delay perfect reconstruction filter banks; a Research Assistant at the University of Bonn, Bonn, Germany, in 1994, where he worked on filter banks for vision and their optimization; and a Research Assistant and Teaching Assistant at the University of Hanover. He is currently with Bell Laboratories, Lucent Technologies, Murray Hill, NJ, where he works in the Multimedia Communications Research Laboratory. His research interests include multirate signal processing; filter banks; speech, audio, and image coding; and electronics.



Tanja Karp (M'98) was born in Germany in 1969. She received the Dipl.-Ing. degree in electrical engineering and the Dr.-Ing. degree from Hamburg University of Technology, Hamburg, Germany, in 1993 and 1997, respectively.

In 1995 and 1996, she spent two months as a Visiting Researcher at the Signal Processing Department, ENST, Paris, France, and at the Multirate Signal Processing Group, University of Wisconsin, Madison, respectively, working on modulated filter banks. Since 1997, she has been with Mannheim University, Mannheim, Germany, as a Research and Teaching Associate. Since 1998, she has also taught as a Guest Lecturer at Freiburg University, Freiburg, Germany. Her research interests include multirate signal processing, filter banks, audio coding, images coding, multicarrier modulation, and signal processing for communications.