Time-Varying Lapped Transforms and Wavelet Packets

Ricardo L. de Queiroz, *Student Member, IEEE*, and K. R. Rao, *Senior Member, IEEE*

***Abstract*—The perfect reconstruction conditions for a time-varying lapped transform (paraunitary filter bank) are developed through the factorization of the transform matrix into sparse factors. A general formulation is presented allowing one to switch between two paraunitary filter banks. However, the extended lapped transform (ELT) is often used as an example. Furthermore, an adaptive wavelet packet is developed employing a time varying tree association of ELT's. In all cases perfect reconstruction is inherently assured.**

I. Introduction

MULTIRATE filter banks are well-known powerful tools in modern digital signal processing allowing easy data processing transform domain, and flexible time-frequency analysis. References such as [1]–[3] cover the main issues on filter banks. The common denomination of "uniform paraunitary filter bank" [2], [4] is analogous to the term "lapped transform" as defined in [3]. Although studied independently in the past, both represent the same (see [3] for the demonstration and definitions). Similar relation is valid for orthogonal discrete wavelet basis, paraunitary filter banks and lapped transforms, mainly for the case of 2-channel banks [5]–[7]. In this paper, we will refer to uniform FIR paraunitary filter banks as lapped transforms. Furthermore, terms like extended lapped transform (ELT) [3], [8], and lapped orthogonal transform (LOT) [9] stand for particular choices of lapped transforms, obeying well defined structures, and are not synonymous for the general term lapped transform.

We are assuming the use of a uniform analysis bank of M FIR filters, each one of length L as shown in Fig. 1. L is related to M as $L = NM = 2KM$, where K is the overlapping factor. The analysis filters are time-reversed versions of the synthesis filters [3]. If the analysis and synthesis filters are represented by $f_m(n)$ and $g_m(n)$, respectively, for $m = 0, 1, \cdots, M-1$ and $n = 0, 1, \cdots, L-1$, we can define a matrix $\boldsymbol{P}$ with elements p_{mn} as

$$p_{mn} = g_m(n) = f_m(L-1-n). \tag{1}$$

Manuscript received July 2, 1992; revised June 10, 1993. The Guest Editor coordinating the review of this paper and approving it for publication was Dr. Ahmed Tewfik. This work was supported in part by the Conselho Nacional de Desenvolvimento Cientifico e Tecnologico (CNPq), Brazil under Grant 200.804-90-1.

The authors are with the Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX 76019-0016.

IEEE Log Number 9212181.

$\boldsymbol{P}$ is a $M \times L$ matrix, which will throughout this paper be called the transform matrix. Define the counter-identity matrix $\boldsymbol{J}_k$ of size $k \times k$ as

$$\boldsymbol{J}_k = \begin{bmatrix} 000 \cdots 01 \\ 000 \cdots 10 \\ \vdots \quad\quad \vdots \\ 010 \cdots 00 \\ 100 \cdots 00 \end{bmatrix}.$$

Similar notation is used for the identity matrix and for the square null matrix ($\boldsymbol{I}_k$ and $\boldsymbol{0}_k$).

Filter banks are generally thought of as stationary forms, and several perfect reconstruction (PR) conditions have been presented under different viewpoints [4], [10]. Recently, Nayebi *et al.* presented a study on the structure of time-varying filter banks [11]. In their work, PR conditions were stated and it was mainly considered the transition between two known PR systems. In this paper, we will present a structure that is inherently orthogonal and is based on lapped transforms with fast algorithms. In Section II, a discussion concerning the factorization of the transform matrix and of the system of transforms that operates over an infinite sequence will be carried. Section III presents the approach for maintaining orthogonality under time variations. Also, time-domain PR conditions are stated as a function of time, applying the concept of an instantaneous filter bank. Section IV studies the ELT in a variational environment, and Section V is concerned with hierarchical systems, employing a time-varying tree-structure using the ELT. In this, the format of the tree is changed along the time axis, forming a time-varying wavelet packet. The paper ends with the conclusions in Section VI.

II. Factorization Into Sparse Matrices

A. General Case

A lapped transform matrix of dimensions $M \times L$ ($L = NM$) can be divided into square $M \times M$ submatrices $\boldsymbol{P}_i$ ($i = 0, 1, \cdots, N-1$) as

$$\boldsymbol{P} = [\boldsymbol{P}_0, \boldsymbol{P}_1, \cdots, \boldsymbol{P}_{N-1}]. \tag{2}$$

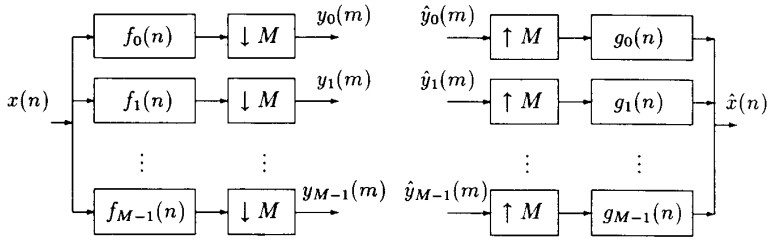

Fig. 1. Critically decimated uniform filter bank. Analysis (left) and synthesis (right) sections are shown.

It can be shown [3] that the PR conditions for a lapped transform, or for any paraunitary filter bank, are given by

$$\sum_{m=0}^{N-1-l} \boldsymbol{P}_m \boldsymbol{P}_{m+l}^T = \sum_{m=0}^{N-1-l} \boldsymbol{P}_{m+l} \boldsymbol{P}_m^T = \delta(l)\boldsymbol{I}_M \quad (3)$$

for $l = 0, 1, \cdots, N-1$, where $\delta(l)$ is the Kronecker delta. Alternately, (3) can be stated as [3]

$$\sum_{m=0}^{N-1-l} \boldsymbol{P}_m^T \boldsymbol{P}_{m+l} = \sum_{m=0}^{N-1-l} \boldsymbol{P}_{m+l}^T \boldsymbol{P}_m = \delta(l)\boldsymbol{I}_M. \quad (4)$$

In [12], it was shown that any paraunitary filter bank (in our notation, any lapped transform) can have its polyphase component matrix $\boldsymbol{E}(z)$ [2], [3], [12] decomposed into a cascade of N_z zero-order lossless (orthogonal) matrices and $N_z - 1$ diagonal matrices containing delays, where $N_z - 1$ is the degree of det $\{\boldsymbol{E}(z)\}$. Then,

$$\boldsymbol{E}(z) = \boldsymbol{Z}\boldsymbol{D}_0(z)\boldsymbol{B}_0\boldsymbol{D}_1(z)\boldsymbol{B}_1 \cdots \boldsymbol{D}_{N_z-2}(z)\boldsymbol{B}_{N_z-2} \quad (5)$$

where $\boldsymbol{D}_k(z)$ ($k = 0, 1, \cdots, N_z - 2$) are diagonal matrices, with ones along the diagonal, except by one delay of the form z^{-1}. $\boldsymbol{Z}$ is a general orthogonal transform, which has $M(M-1)/2$ degrees of freedom, one for each possible plane rotation of the form

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

while the matrices $\boldsymbol{B}_i$ can be implemented with only $M-1$ plane rotations [12].

Let us consider the time domain representation of the analysis–synthesis process. Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be the infinite-length time and frequency domain vectors, respectively. The analysis process of Fig. 1 can be represented using the transform matrix $\tilde{\boldsymbol{P}}$, which is defined as

$$\tilde{\boldsymbol{P}} = \begin{pmatrix} \ddots & & \ddots & & & \\ & \boldsymbol{P}_0 & \boldsymbol{P}_1 & \cdots & \boldsymbol{P}_{N-1} & \boldsymbol{0} & \boldsymbol{0} \\ & \boldsymbol{0} & \boldsymbol{P}_0 & \boldsymbol{P}_1 & \cdots & \boldsymbol{P}_{N-1} & \boldsymbol{0} \\ & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{P}_0 & \boldsymbol{P}_1 & \cdots & \boldsymbol{P}_{N-1} \\ & & & & \ddots & & \ddots \end{pmatrix}. \quad (6)$$

Thus,

$$\boldsymbol{y} = \tilde{\boldsymbol{P}}\boldsymbol{x} \quad (7)$$

and, from (3) and (4), we have the synthesis equation

$$\hat{\boldsymbol{x}} = \tilde{\boldsymbol{P}}^T\boldsymbol{y} = \tilde{\boldsymbol{P}}^T\tilde{\boldsymbol{P}}\boldsymbol{x} = \boldsymbol{x}. \quad (8)$$

Given that the matrices $\boldsymbol{D}_i(z)$ have only pure delays and the matrices $\boldsymbol{B}_i$ and $\boldsymbol{Z}$ are orthogonal matrices, we can combine the above relations with (5) and we can see that the matrix $\tilde{\boldsymbol{P}}$ can be factored into

$$\tilde{\boldsymbol{P}} = \tilde{\boldsymbol{Z}}\tilde{\boldsymbol{D}}_0\tilde{\boldsymbol{B}}_0\tilde{\boldsymbol{D}}_1\tilde{\boldsymbol{B}}_1 \cdots \tilde{\boldsymbol{D}}_{N_z-2}\tilde{\boldsymbol{B}}_{N_z-2} \quad (9)$$

where $\tilde{\boldsymbol{D}}_i$ is a permutation matrix and the matrices $\tilde{\boldsymbol{B}}_i$ and $\boldsymbol{Z}$ are block diagonal matrices, so that

$$\tilde{\boldsymbol{B}}_i = \text{diag}\{\cdots \boldsymbol{B}_i, \boldsymbol{B}_i, \boldsymbol{B}_i \cdots\} \quad (10)$$

$$\tilde{\boldsymbol{Z}} = \text{diag}\{\cdots \boldsymbol{Z}, \boldsymbol{Z}, \boldsymbol{Z} \cdots\}. \quad (11)$$

The inverse operation is, therefore, given by

$$\tilde{\boldsymbol{P}}^{-1} = \tilde{\boldsymbol{B}}_{N_z-2}^T\tilde{\boldsymbol{D}}_{N_z-2}^T \cdots \tilde{\boldsymbol{B}}_1^T\tilde{\boldsymbol{D}}_1^T\tilde{\boldsymbol{B}}_0^T\tilde{\boldsymbol{D}}_0^T\tilde{\boldsymbol{Z}}^T. \quad (12)$$

Fig. 2 shows an example illustrating the relation among z-domain flow-graph (with delays), time-domain non-causal flow-graph and the sparse matrices. In this figure, the order of the input–output M-samples blocks is indicated, with each branch carrying $M/2$ samples. The synthesis is accomplished by following the flow-graphs from the right to the left, since the matrices are all orthogonal. $\tilde{\boldsymbol{P}}$ is, then, a layered implementation of transforms. If we could follow the paths connecting each output block to the input blocks, we would be able to also factorize $\boldsymbol{P}$ in a sense that

$$\boldsymbol{P} = \boldsymbol{Z}\boldsymbol{D}_0\boldsymbol{B}_0'\boldsymbol{D}_1\boldsymbol{B}_1' \cdots \boldsymbol{D}_{N_z-2}\boldsymbol{B}_{N_z-2}'. \quad (13)$$

This would also be a general expression for a factorization of $\boldsymbol{P}$. The matrices $\boldsymbol{B}_i'$ are square and block-diagonal in a size-limited version of $\tilde{\boldsymbol{B}}_i$. The number of blocks in the diagonal of these matrices is dictated by the delays of the $\boldsymbol{D}_i(z)$ matrices, and its a priori calculation for all cases is impractical for our purposes. The matrices $\boldsymbol{D}_i$ are not nec-

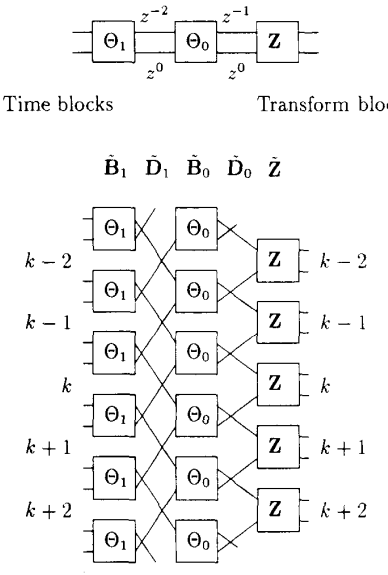

Fig. 2. Example: relation between z-domain factors and its correspondent time-domain noncausal analysis flow-graph. Each branch corresponds to $M/2$ samples, the blocks have M samples, and the boxes are $M \times M$ orthogonal matrices. The sparse factors are also indicated, aligning their symbols with the flow-graph stages. Since all elements are orthogonal, the synthesis is implemented by following the paths of the flow-graph from right to left.

essarily square and are size limited versions of the $\tilde{\boldsymbol{D}}_i$. They should not be confused with $\boldsymbol{D}_i(z)$ which has a different meaning and has complex entries. As a remark, using the layering viewpoint, the matrices $\boldsymbol{D}_i$ have to be lapped transforms, i.e., obey (3), if the whole system is supposed to maintain the PR property.

We will not attempt to design any lapped transform and we will assume that, given the lapped transform, its sparse factor decomposition is known, as in (5).

B. Example: The ELT

We will now address a practical factorization. We consider the extended lapped transform (ELT) as defined by Malvar in [3], [8], among the most efficient factorization methods for a paraunitary filter bank. Although imposing restrictions (that lead to fast algorithms) the resulting filter bank presents very good frequency response [3]. Other choices of modulated filter banks could also be made, such as those in [13], [14]. However, the performance differences, if of any significance, would not affect the general results. In the ELT's, the filters' length is basically an even multiple of the blocksize M, as $L = 2KM$ and $\boldsymbol{P}$ can be split into only $K + 1$ stages.

$$\boldsymbol{E}(z) = \boldsymbol{Z}\boldsymbol{D}_0(z)\Theta_0\boldsymbol{D}_1(z)\Theta_1 \cdots \boldsymbol{D}_{K-1}(z)\Theta_{K-1}, \quad (14)$$

$$\tilde{\boldsymbol{P}} = \tilde{\boldsymbol{Z}}\tilde{\boldsymbol{D}}_0\tilde{\boldsymbol{B}}_0\tilde{\boldsymbol{D}}_1\tilde{\boldsymbol{B}}_1 \cdots \tilde{\boldsymbol{D}}_{K-1}\tilde{\boldsymbol{B}}_{K-1}. \quad (15)$$

$\boldsymbol{Z}$ is an $M \times M$ DCT type IV matrix [15], with inverted inputs and Θ_i are orthogonal matrices composed by only $M/2$ plane rotations [3], [8]. The components of the above factorization are

$$\boldsymbol{Z} = \boldsymbol{DCT}^{IV}\begin{bmatrix} \boldsymbol{0}_{M/2} & \boldsymbol{I}_{M/2} \\ \boldsymbol{I}_{M/2} & \boldsymbol{0}_{M/2} \end{bmatrix}. \quad (16)$$

$$\boldsymbol{D}_0(z) = \begin{bmatrix} z^{-1}\boldsymbol{I}_{M/2} & \boldsymbol{0}_{M/2} \\ \boldsymbol{0}_{M/2} & \boldsymbol{I}_{M/2} \end{bmatrix} \quad \boldsymbol{D}_n(z) = \begin{bmatrix} z^{-2}\boldsymbol{I}_{M/2} & \boldsymbol{0}_{M/2} \\ \boldsymbol{0}_{M/2} & \boldsymbol{I}_{M/2} \end{bmatrix}$$
$$(n = 1, \cdots, K-1), \quad (17)$$

$$\tilde{\boldsymbol{B}}_n = \text{diag}\,\{\cdots \Theta_n, \Theta_n, \Theta_n, \cdots\}, \quad (18)$$

$$\Theta_n = \begin{bmatrix} -\boldsymbol{C}_n & \boldsymbol{S}_n\boldsymbol{J}_{M/2} \\ \boldsymbol{J}_{M/2}\boldsymbol{S}_n & \boldsymbol{J}_{M/2}\boldsymbol{C}_n\boldsymbol{J}_{M/2} \end{bmatrix},$$
$$\boldsymbol{C}_n = \text{diag}\,\{\cos(\theta_{0,n}), \cos(\theta_{1,n}), \cdots, \cdot \cos(\theta_{(M/2)-1,n})\}$$
$$\boldsymbol{S}_n = \text{diag}\,\{\sin(\theta_{0,n}), \sin(\theta_{1,n}), \cdots, \sin(\theta_{(M/2)-1,n})\} \quad (19)$$

$\theta_{i,j}$ are the rotation angles and free parameters in the design of an ELT [3].

Fig. 3 shows the flow graph for the case $K = 2$. From the information above, we can also factorize $\boldsymbol{P}$ as

$$\boldsymbol{P} = \boldsymbol{Z}\boldsymbol{D}_0\boldsymbol{B}_0'\boldsymbol{D}_1\boldsymbol{B}_1' \cdots \boldsymbol{D}_{K-1}\boldsymbol{B}_{K-1}' \quad (20)$$

where $\boldsymbol{B}_n'$ matrices are block diagonal with $2(n+1)$ rows of $2(n+1)$ blocks, each block of size $M \times M$ ($0 \le n \le K-1$). The $\boldsymbol{D}_n$ matrices have dimensions: $M \times 2M$ ($n = 0$), and $2nM \times (2n+2)M$, for $1 \le n \le K-1$. These matrices are more easily described as block matrices generated by a Kronecker product

$$\boldsymbol{D}_n = \boldsymbol{F}_n \otimes \boldsymbol{I}_{M/2} \quad (21)$$

where $\otimes$ denotes the Kronecker product. Then, we can describe the $\boldsymbol{F}_n$ matrices (with elements $f_{ij,n}$) by

$$\boldsymbol{F}_0 = \{f_{ij,0}\} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (22)$$

and $\boldsymbol{F}_n = \{f_{ij,n}\}$, where

$$f_{ij,n} = \begin{cases} 1 & i = j = 2k \\ 1 & i = 4n - 1 - 2k \text{ and } j = 4n + 3 - 2k \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

for $0 \le k \le 2n-1$ and $1 \le n \le K-1$.

III. Perfect Reconstruction and Time-Variation

Suppose the orthogonal building blocks in Fig. 3 were no longer identical, for each column. For example, if the angles belonging to Θ_0 are changed as we go from top to the bottom of the flow-graph, what should be the consequences for the filter bank, and more precisely, can we still maintain the PR property? If each building block is

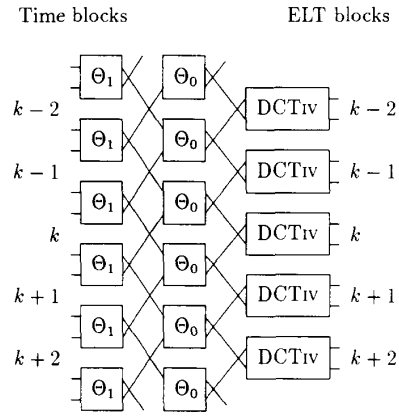

Fig. 3. Flow graph for the ELT with $K = 2$. Time and ELT domain samples are grouped into blocks of M samples and each branch corresponds to $M/2$.

orthogonal, it is clear that the PR conditions are satisfied, with the same flow-graph for analysis and synthesis process. When nonorthogonal blocks are included, their inverse must be applied in the synthesis process leading to biorthogonal filter banks. Although possible, we will disregard this case since we are mainly concerned with lapped transforms (paraunitary filter banks).

Let the matrix $\tilde{\boldsymbol{B}}_n$ contain different matrices along its diagonal. Let us also assume a time-index k to characterize each block. From a decimation–interpolation viewpoint, arising from the classical filter bank analysis [1], this index represents the decimated-rate sampling instants of the subbands. On the lapped transform viewpoint, k is the block number [3], [4]. Note that PR is inherently maintained, but the transform matrix is now a function of k. $\boldsymbol{P}$ is updated from block to block and we may refer to it as $\boldsymbol{P}(k)$ [16]. Then,

$$\tilde{\boldsymbol{P}} = \begin{pmatrix} \ddots & & \ddots & & \\ & \boldsymbol{P}_0(k-1) & \cdots & \boldsymbol{P}_{N-1}(k-1) & \boldsymbol{0} & \boldsymbol{0} \\ & \boldsymbol{0} & \boldsymbol{P}_0(k) & \cdots & \boldsymbol{P}_{N-1}(k) & \boldsymbol{0} \\ & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{P}_0(k+1) & \cdots & \boldsymbol{P}_{N-1}(k+1) \\ & & & & \ddots & \end{pmatrix} \tag{24}$$

with

$$\boldsymbol{P}(k) = [\boldsymbol{P}_0(k), \boldsymbol{P}_1(k), \cdots, \boldsymbol{P}_{N-1}(k)]. \tag{25}$$

This means that $\boldsymbol{P}(k)$ contains the instantaneous filter bank impulse responses. We can also rewrite $\tilde{\boldsymbol{Z}}$ and $\tilde{\boldsymbol{B}}_n$ as

$$\tilde{\boldsymbol{Z}} = \text{diag}\,\{\cdots \boldsymbol{Z}(k-1), \boldsymbol{Z}(k), \boldsymbol{Z}(k+1), \cdots\} \tag{26}$$

$$\tilde{\boldsymbol{B}}_n = \text{diag}\,\{\cdots \boldsymbol{B}_n(k-1), \boldsymbol{B}_n(k), \boldsymbol{B}_n(k+1), \cdots\} \tag{27}$$

The elements in Fig. 3 are now varied with time, and another example is shown in Fig. 4. The relation (9) remains unchanged, but we can rewrite (13) as

$$\boldsymbol{P}(k) = \boldsymbol{Z}(k)\boldsymbol{D}_0(k)\boldsymbol{B}_0'(k)\boldsymbol{D}_1(k)\boldsymbol{B}_1'(k) \cdots \cdot \boldsymbol{D}_{N_2-2}(k)\boldsymbol{B}_{N_2-2}'(k). \tag{28}$$

In the example in Fig. 4, it is easy to see that a particular choice of $\boldsymbol{Z}(k)$ would influence solely $\boldsymbol{P}(k)$, but a choice of $\Theta_0(k)$ would influence $\boldsymbol{P}(k)$ and $\boldsymbol{P}(k-1)$. Similarly, any particular choice of $\Theta_1(k)$ would affect $\boldsymbol{P}(k-2)$ through $\boldsymbol{P}(k+1)$. This suggests that any change of the filter bank is preceded and followed by transition regions [11].

The time-domain PR equations in (3) assume that $\boldsymbol{P}$ remains unchanged along time-index. Therefore, the analysis can be more easily presented due to the periodic nature of the problem [10]. In fact, this was also done in [3], [4]. In time-varying systems, we have to choose an index k and find the PR equations for it, noting that (3) is no longer valid. Roughly speaking, the PR conditions for steady systems must state the orthogonality of the basis functions and the so-called orthogonality of the "tails." This implies that aliased (shifted) versions of $\boldsymbol{P}$ would cancel in the synthesis process [3], [4]. Since $\boldsymbol{P}(k)$ is no longer constant with k, (3) must be rewritten to ensure the orthogonality among $\boldsymbol{P}(k)$ and its neighbors $\boldsymbol{P}(k \pm 1)$, $\boldsymbol{P}(k \pm 2)$, $\boldsymbol{P}(k \pm 3) \cdots$. Then, it is easy to show that the PR conditions for time-varying lapped transforms (paraunitary filter banks) are

$$\sum_{m=0}^{N-1-l} \boldsymbol{P}_m(k)\boldsymbol{P}_{m+l}^T(k-l) = \sum_{m=0}^{N-1-l} \boldsymbol{P}_{m+l}(k)\boldsymbol{P}_m^T(k+l) = \delta(l)\boldsymbol{I}_M \tag{29}$$

for $l = 0, 1, \cdots, N-1$, yielding $2N-1$ independent matrix equations.

As a remark, the term lapped transform was maintained (although (3) is no longer valid) because $\tilde{\boldsymbol{P}}$ remains orthogonal and for each instant k, the synthesis filters are time-reversed versions of the analysis ones. A filter bank is said *instantaneously paraunitary* if all factors in (26) and (27) are orthogonal matrices. In this case, (29) is satisfied, but not necessarily the steady conditions in (3) or (4) are met. In this situation, the filter bank is said *tran-*

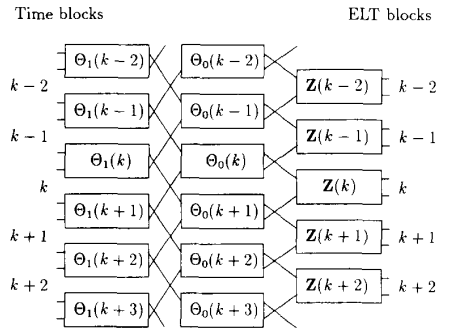

Fig. 4. Flow graph for the time-varying ELT with $K = 2$. Now the orthogonal matrices in the boxes can change and are a function of a time index k. As long as these matrices maintain orthogonality the flow-graph for synthesis follows the paths from right to left. Time and ELT domain samples are grouped into blocks of M samples.

sitory. Of course, all equations are satisfied when the filter bank is not variable, i.e., the factors in (26) and (27) are the same along the diagonal.

Although possible, a continuous change of coefficients, like in adaptive filtering, is not likely to be beneficial, using this structure. Maybe the most practical idea would be to switch between two filter banks at a time. Each one would have its own characteristics (and factorization) well defined

$$\cdots \rightarrow \text{filter bank 1} \rightarrow \text{filter bank 2} \rightarrow$$
$$\text{filter bank 3} \rightarrow \cdots$$

It is, thus, possible to switch between any two filter banks, using the general factorization structure. But what are the limitations of the filter banks to be used in this approach? First, we cannot change the number of channels, unless we first switch to a "transparent state" as we will discuss later. Second, the filters' length is limited to lie between M and L. In fact, we can switch between block and lapped transforms, since the former is a special case of the latter [3], [4]. To see this, let the $M \times l$ matrix $\boldsymbol{Q}$ represent the transform matrix of a filter bank with M channels and filters' length $l \leq L$. If we pad with zeros on both sides of $\boldsymbol{Q}$ in a balanced way such that $\boldsymbol{Q}$ is still centered and

$$\boldsymbol{P} = [\boldsymbol{0} \quad \boldsymbol{Q} \quad \boldsymbol{0}],$$

then, as long as (3) is satisfied, $\boldsymbol{P}$ is a valid lapped transform. Of course, $l \geq M$, otherwise $\boldsymbol{P}$ becomes rank deficient.

Transparent States and Segmentation

A special case of interest is when $\boldsymbol{Q}$ is the identity matrix. In this case the input samples are solely copied to output and the transform is bypassed. However, the transitions between this state and any other state will have PR assured by the algorithm. This is what we call the *transparent* or *bypass* state of the filter bank. Furthermore, if one switches to a transparent state, after the transition is over, no transform is applied to the input. Therefore, these samples are independent from the samples of the region where either the filter bank was active or where there was a transition. These two regions would be independent segments of the signal. If right at the beginning of the second segment we force a transition from the transparent state to a second filter bank, then we can switch transforms and even change the number of channels M, since the segments are independent. We can have

$$\text{filter bank 1 } (M_1 \text{ channels}) \rightarrow \text{transparent state} \rightarrow$$
$$\text{filter bank 2 } (M_2 \text{ channels})$$

or, directly

$$M_1 \text{ channels} \rightarrow$$
$$\text{transition to bypass} \rightarrow \text{transition from bypass} \rightarrow$$
$$M_2 \text{ channels}$$

In the next section, we will describe this switching procedure in the case of the ELT. See [17]–[19] for further results in segmentation and processing of finite-length signals.

IV. Time-Varying ELT

The filter bank is supposed *instantaneously paraunitary* and we may focus our attention on the analysis sections, since the synthesis filters would follow the analysis structure with PR assured. We will now use the ELT for considering adaptation issues, reminding the reader that, so far, with the actual factorization, the only parameter that can be changed is the frequency response of the filters.

The ELT structure is based on the modulation of a cosine train by a window. This window has $2KM$ elements and is defined by the choice of the $KM/2$ plane rotation angles (see Section I-B). The manipulation of these angles permits to shape the filters to some limited extent. At block k, the matrix $\boldsymbol{P}(k)$ with elements $p_{mn,k}$ ($m = 0, 1, \cdots, M-1$ and $n = 0, 1, \cdots, L-1$) can be defined, as in [3], but with a time-varying window, by

$$p_{mn,k} = h(n, k) \cos\left[\left(m + \frac{1}{2}\right)\left(\left(n - \frac{L-1}{2}\right)\frac{\pi}{M} + (N+1)\frac{\pi}{2}\right)\right] \tag{30}$$

$h(n, k)$ is the window for instant k. The M impulse responses of the analysis and synthesis filters, denoted as $f_m(n, k)$ and $g_m(n, k)$, respectively are related by

$$p_{mn,k} = g_m(n, k) = f_m(L-1-n, k). \tag{31}$$

Thus, for block k, we have the frequency response of the filters (dropping index k from the formula for convenience) given as [3]

$$G_m(e^{j\omega}) = \frac{1}{\sqrt{2M}}\left[e^{-j\eta_m} H(e^{j(\omega-\omega_m)}) + e^{j\eta_m} H(e^{j(\omega+\omega_m)})\right] \tag{32}$$

where $\omega_m \equiv (m + (1/2))\pi/M$, $\eta_m \equiv \omega_m(M+1)/2$, and $H(e^{j\omega})$, $G_m(e^{j\omega})$ are the frequency responses of $h(n, k)$ and $g_m(n, k)$, respectively.

As discussed previously, the transparent state would cause the input samples to be copied to output without any transformation. To achieve this state we have to set all angles as $\pi/2$ and replace $\boldsymbol{Z}$ by

$$\boldsymbol{Z} = \begin{pmatrix} \boldsymbol{J}_{M/2} & \boldsymbol{0}_{M/2} \\ \boldsymbol{0}_{M/2} & \boldsymbol{J}_{M/2} \end{pmatrix}^K \begin{pmatrix} \boldsymbol{0}_{M/2} & \boldsymbol{I}_{M/2} \\ \boldsymbol{I}_{M/2} & \boldsymbol{0}_{M/2} \end{pmatrix}^{K+1}. \quad (33)$$

$$\boldsymbol{P}(k) = \boldsymbol{DCT}^{\mathrm{IV}} \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \Theta_0(k) & \boldsymbol{0} \\ \boldsymbol{0} & \Theta_0(k+1) \end{pmatrix} \times \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I} \end{pmatrix}$$

$$\times \begin{pmatrix} \Theta_1(k-1) & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \Theta_1(k) & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \Theta_1(k+1) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \Theta_1(k+2) \end{pmatrix} \quad (36)$$

The reader can confirm this statement either by substituting the values of the angles in the ELT matrix factorization, or by inspection of the flow-graph shown in Fig. 5. If we switch from a regular ELT to a transparent state, or vice-versa, there will be a transition region, which will be focused later. As the transition region is over, it has no sense in following the paths in Fig. 5 and samples can be just copied to output. The notion of the transparent state is only necessary to maintain PR in the transition regions.

A. Cases $K = 1$ and $K = 2$

We shall focus our attention on $h(n, k)$, since the window defines the filter bank, being a low-pass prototype which is modulated by the cosine sequence. Regrettably, there is no direct analytical relation among the angles and the window for all K. We will present only the cases of K equals to 1 and 2 under time variations, which is only an extension of the steady case presented in [3]. Assume $\boldsymbol{I}$ and $\boldsymbol{0}$ are of size $M/2 \times M/2$. For $K = 1$,

$$\boldsymbol{P}(k) = \boldsymbol{DCT}^{\mathrm{IV}} \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \cdot \begin{pmatrix} \Theta_0(k) & \boldsymbol{0} \\ \boldsymbol{0} & \Theta_0(k+1) \end{pmatrix} \quad (34)$$

with Θ_0 given by (19), but with time varying angles. The window can be expressed as

$$\begin{aligned} h(l, k) &= -\cos(\theta_{l,0}(k)) \\ h(M-1-l, k) &= -\sin(\theta_{l,0}(k)) \\ h(M+l, k) &= -\sin(\theta_{l,0}(k+1)) \\ h(2M-1-l, k) &= -\cos(\theta_{l,0}(k+1)) \\ & \qquad l = 0, 1, \cdots, M/2-1. \end{aligned} \quad (35)$$

For $K = 2$, with Θ_1 also given by (19), we have

with the window given by

$$\begin{aligned} h(l, k) &= \cos(\theta_{l,0}(k))\cos(\theta_{l,1}(k-1)) \\ h(M-1-l, k) &= \cos(\theta_{l,0}(k))\sin(\theta_{l,1}(k-1)) \\ h(M+l, k) &= \sin(\theta_{l,0}(k+1))\cos(\theta_{l,1}(k)) \\ h(2M-1-l, k) &= -\sin(\theta_{l,0}(k+1))\sin(\theta_{l,1}(k)) \\ h(2M+l, k) &= -\sin(\theta_{l,1}(k))\sin(\theta_{l,0}(k+1)) \\ h(3M-1-l, k) &= \sin(\theta_{l,0}(k))\cos(\theta_{l,1}(k+1)) \\ h(3M+l, k) &= \cos(\theta_{l,0}(k+1))\sin(\theta_{l,1}(k+2)) \\ h(4M-1-l, k) &= \cos(\theta_{l,0}(k+1))\cos(\theta_{l,1}(k+2)) \\ & \qquad l = 0, 1, \cdots, M/2-1. \end{aligned} \quad (37)$$

B. Transition

A transition region can be defined as containing those indexes k where $\boldsymbol{P}(k)$ is formed by different $\Theta_i(n)$, for a fixed i. In other words, it is composed by *transitory* filter banks, as defined in a previous section. These transitory filter banks can have undesired frequency responses and the transition region, for the ELT, will last for $K+1$ blocks.

As we saw in a previous section, the analysis or synthesis processes ($\tilde{\boldsymbol{P}}$) can be viewed as a layered implementation of block transforms and permutations. If, in the middle of the transition, the $\boldsymbol{B}$ matrices are changed, this will further prolong it. On the other hand, the $\boldsymbol{Z}$ matrix in $\tilde{\boldsymbol{Z}}$ is directly connected to the output and $\boldsymbol{Z}(k)$ just influ-

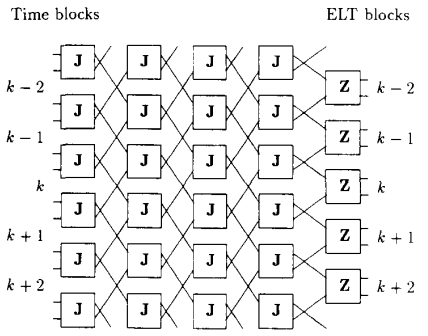

Fig. 5. Flow graph for the ELT with $K = 4$ under the transparent state, where input is solely copied to output.

ences $\boldsymbol{P}(k)$ (see Fig. 4). A time-varying block transform is trivial and

$$\boldsymbol{P}(k) = \boldsymbol{Z}(k)\,\boldsymbol{W}(k) \tag{38}$$

$$\boldsymbol{W}(k) = \boldsymbol{D}_0(k)\boldsymbol{B}_0(k)\boldsymbol{D}_1(k)\boldsymbol{B}_1(k)\cdots\boldsymbol{D}_{K-1}(k)\boldsymbol{B}_{K-1}(k) \tag{39}$$

We can find a "make-up" matrix to use as $\boldsymbol{Z}(k)$ and replace the matrix defined in (16). This approach has the intention to improve the filters' characteristics. If the input signal has autocorrelation matrix $\boldsymbol{R}_{xx}$, then the signal after using the transforms $\boldsymbol{W}(k)$ and $\boldsymbol{P}(k)$ have autocorrelation matrices $\boldsymbol{R}_{ww}$ and $\boldsymbol{R}_{yy}$, respectively, given by

$$\boldsymbol{R}_{ww}(k) = \boldsymbol{W}(k)\boldsymbol{R}_{xx}\boldsymbol{W}^T(k) \tag{40}$$

$$\boldsymbol{R}_{yy}(k) = \boldsymbol{Z}(k)\boldsymbol{R}_{ww}\boldsymbol{Z}^T(k). \tag{41}$$

In compression-coding applications, we can find a matrix $\boldsymbol{Z}(k)$ that would decorrelate the signal and make $\boldsymbol{R}_{yy}$ a diagonal matrix. It is well known that an orthogonal matrix for decorrelating $\boldsymbol{y}$ would have its rows as the M eigenvectors of $\boldsymbol{R}_{ww}$. Hence, $\boldsymbol{Z}(k)$ can be expressed as

$$\boldsymbol{Z}^T(k) = [\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_M] \tag{42}$$

where $\boldsymbol{q}_1 \cdots \boldsymbol{q}_M$ are the eigenvectors of $\boldsymbol{R}_{ww}$. This concept was used in the development of the LOT in [9].

On the other hand, if one is switching between a transform and a transparent state, the goal can be exactly the opposite, i.e., no decorrelation, and the filters' frequency response should be as flat as possible. We just highlighted one point (increasing decorrelation) and each case may be treated separately.

C. Some Applications

1) Variable Overlap Transform: In image coding, block transforms such as the DCT [15] are widely used for several reasons. However, block transforms lead to the so-called blocking effects [15] at low bit-rates, and the LOT [9] was designed to eliminate them by extending the basis functions across the traditional block boundaries. It does not completely eliminate all blocking structure, neither do the ELT's, although this effect is largely reduced. Some blocking is still noticeable if we decrease the bit-rate further. This is because the end of the basis functions for one block are still neighbors of the beginning of the basis functions for another block, forming a well defined boundary. On the other hand, we would like to keep the filters (basis functions) as long as possible to increase the filters selectivity and, therefore, the coding gain, through better spectral energy compaction. As the amount of distortion in a particular region is kept small, the blocking tends to be eliminated and we can use the benefits of computational savings of block transforms with fast algorithms. All these facts account to a possible benefit if one uses transforms with variable overlap. One can develop an adaptive coder which would employ a variable overlap and even mix lapped transforms with block transforms. With the ELT, at least, we have the means to do it, while maintaining PR.

For $K = 1$, in (35) we can set marginal elements of the window to be zero, by forcing the angles $\theta_{00}, \theta_{10}, \cdots$ to be $\pi/2$. All the basis functions are modulated by this window and would also be set to zero. Actually, the angles can be changed at any point. The transition region will only encompass one block and the filters will have frequency response lying in between the responses of initial and final filter banks. Fig. 6(a) shows the frequency response of an ELT ($K = 1$) for $M = 8$. Fig. 6(b) shows the same plot, but setting $\theta_{00} = \theta_{10} = \pi/2$ (two elements in each extreme are set to zero for each basis function).

A more extreme case would be one where it is desired to switch betwen ELT's and block transforms at any time. All the angles are, then, suddenly changed to $\pi/2$. The overlap disappears because the window becomes rectangular of length M. Thus, we can substitute the $\boldsymbol{Z}$ matrix by any other matrix. The frequency responses in Fig. 6(a) for the steady ELT are repeated in Fig. 7(a) for the transition filter bank and in Fig. 7(b) for the DCT. The flowgraph in Fig. 8 shows a fast algorithm for changing an ELT ($K = 1$) into a DCT. The inverse operation (DCT to ELT) is straight forward by turning the flow graph in Fig. 8 upside down. We can see that PR is maintained and the transitory frequency response still preserves reasonable attenuations. We can also develop the same approach using the concepts of transparent state and region segmentation as discussed earlier.

2) Time-Varying Wavelet Packets: Wavelet packets are hierarchical associations of filter banks following the paths of a binary tree. We propose a way to fully vary the shape of the tree, by pruning and expanding branches. For this we can use the ELT as the filter bank for all branches. When a particular branch is to be pruned, we set the ELT to its transparent state, and when this branch is to be expanded we switch the ELT from its transparent state to its normal state again. The reason for switching to transparent state is because if we suddenly prune the tree by not processing the input samples, the transition will not maintain PR. We will devote the next section to the subject of wavelet packets.

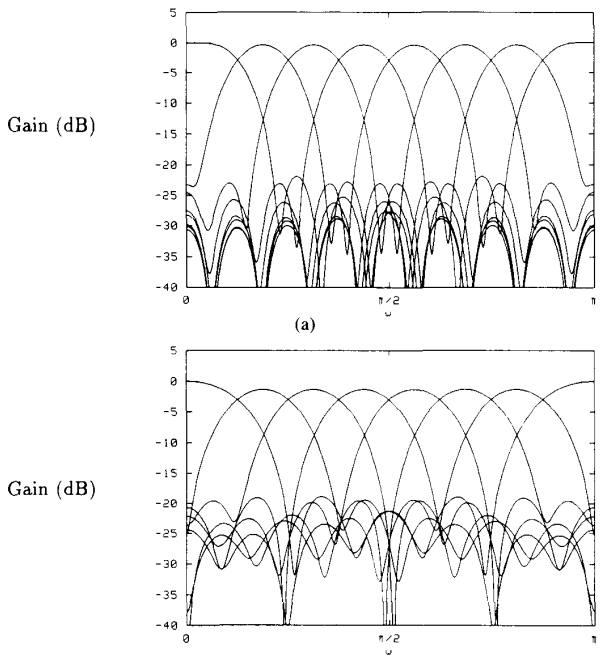

Fig. 6. Frequency response of the ELT with $K = 1$, $M = 8$. Each plot corresponds to the frequency response (magnitude in dB) of each synthesis filter $G_m(e^{j\omega})$. (a) Top: regular design using all angles; (b) bottom: setting 2 (out of 4) elements at each border of the overlapped part of the modulating window to be zero without reoptimizing the remaining elements.

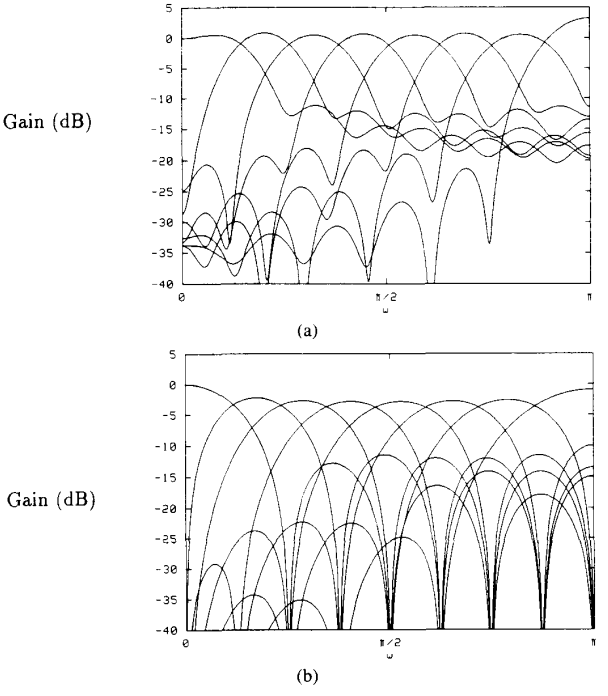

Fig. 7. Frequency response of instantaneous filter banks. Each plot corresponds to the frequency response (magnitude in dB) of each synthesis filter $G_m(e^{j\omega})$. (a) Top: transition filter bank, half of the angles are normal and half of them are forced to be $\pi/2$, resulting in an asymmetrical window; (b) bottom: DCT for $M = 8$.

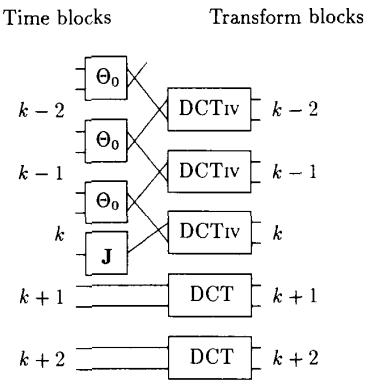

Fig. 8. Flow graph for fast switch between the ELT ($K = 1$) and the DCT, beginning from block k, which belongs to the transition.

V. Time-Varying Wavelet Packets Using Variable Tree-Paths and the ELT

The hierarchical structures are gaining more and more attention due to their improved time-frequency resolution. Furthermore, it is possible to achieve a great variety of nonuniform filter banks. The wavelet transform has been largely used and investigated [5], [6], [20]–[22] and it is well known that in the discrete case, the orthogonal wavelet transform is an association of 2-channel ($M = 2$) filter banks [5]–[7], [23]. The ELT in the 2-channel case is a very good way to implement such transforms [24]. Wavelet transforms and octave analysis have an asymptotic behavior, regarding compaction performance for a stationary AR(1) process [23]. It is easy to find an ELT with greater compaction and faster algorithms than wavelet transforms based on a single 2-channel bank, no matter how many stages are connected [23]. However, for nonstationary patterns the exchange of frequency by time resolution can make the wavelets superior to M-band systems which have large filter responses for all frequencies. The so-called wavelet packet [7], [25]–[27] is a compromise between the wavelet transform and the full-tree and the filter banks association follows arbitrary shapes of the binary tree. The binary tree was only used for illustration purposes. M-ary and mixed trees can be also employed obeying the same basic principles as discussed here.

A. Binary Trees Notation

If the same lapped transform is used as the analysis cell for each stage, it is sufficient to describe the paths of the tree to completely describe the whole analysis–synthesis system. With the aid of Fig. 9 we present a more convenient notation. In Fig. 9(a), we see a 2-band maximally decimated filter bank with its low- and high-pass filters, as well as subsamplers. This system will be represented here by tree nodes and branches [Fig. 9(b)], where the signal flows in the nodes and the branches represent filters and decimators. We denote the number of stages as S and clearly the maximum number of terminal nodes is 2^S which is also the number of nodes in the last stage when we use a full-tree. We label the nodes in the tree as η_{ij}, where i

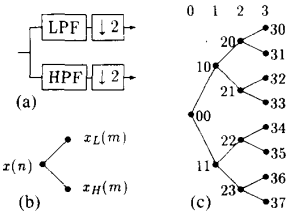

Fig. 9. The binary tree notation. (a) 2-band filter bank; (b) its equivalent representation in a tree; (c) labelling of nodes in the tree.

is the stage number and j is the number of the node in that stage ($0 \le j \le 2^i - 1$), just as indicated in Fig. 9(c), for a 3-stage binary tree. The familiar parenthood notation is used, so that node η_{ij} is parent of nodes $\eta_{i+1,2j}$ and $\eta_{i+1,2j+1}$, and is the child of $\eta_{i-1,j\varnothing 2}$ ($\varnothing$ means integer division). We denote x_{ij} as the signal flowing in η_{ij}, while x_{00} is the original input signal. As a remark, the number of a node in a level does not correspond to an increasing frequency ordering of bands.

The idea is to adaptively reshape the tree. In this case, it is convenient to define an infinite number of stages and an activity map. This map indicates if the node is active (its signal is being processed as an input to a filter bank) or not. Let $a_{ij}(n) = 1$ denote an active node ij, being 0 otherwise. The rightmost node in a path will be called instantaneously virtual end node (IVEN). All nodes with a childhood relation to an IVEN are inactive at instant n, and clearly an IVEN is inactive.

B. Pruning and Expanding Branches

To prune a branch of the tree, it is sufficient to bypass the transform applied to the parent node of an IVEN which, after pruning, will become a new IVEN. The recurrence of this procedure applied to the desired branches will bypass the signal from the resulting IVEN to the right, pruning the tree. The inverse procedure expands the tree, activating nodes and branches. As discussed previously, the use of the transparent state is necessary, in our approach, if we want to maintain PR in the transitions. One further detail should have our attention when the transform is bypassed. The filters have gain of $\sqrt{M} = \sqrt{2}$. Since some samples will be copied to the output and some others will be transformed, after several stages along the tree, the resulting signal can have highly unbalanced range of sample amplitudes. To facilitate the processing of the transformed samples, it is advisable to multiply the samples by $\sqrt{2}$ when they are "transformed" by a transparent state ELT. At the inverse transform these samples should be divided by the same value.

To bypass an ELT, there are K transitory filter banks (blocks). Fig. 10 shows the analysis flow-graph to bypass an ELT for $K = 2$. A similar configuration is followed for any value of K. The inverse transition between transparent and normal states of an ELT are, again, obtained by reversing the signal ordering, i.e., by viewing Fig. 10 upside down. For $K = 2$, the low- and high-pass basis functions are shown for the transition process in Fig. 11.

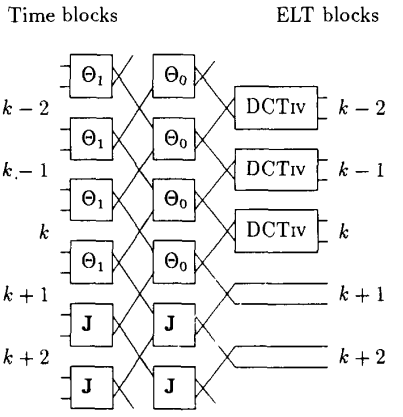

Fig. 10. Flow graph for the ELT with $K = 2$ switching to transparent state, where input is solely copied to output. Blocks k and $k + 1$ belong to the transition. Blocks $k + 2$, $k + 3$ . . . do not need to follow the paths anymore, being directly copied to output.

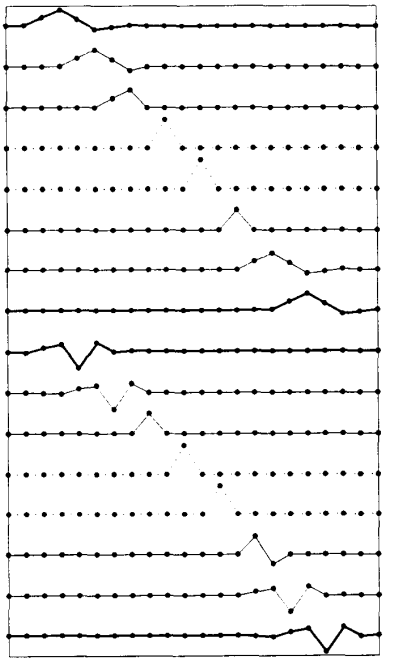

Fig. 11. Illustration of the basis for the signal space during transitions to transparent states, using a relatively short filter bank (ELT, $K = 2$). The filter bank branches are pruned from the tree and after two blocks of two samples, are reexpanded. On the top 8 plots, it is shown the "low-pass" bases, while at the bottom it is shown the 8 "high-pass" ones. Thick lines represent regular bases, thin lines transitions, and dotted lines the trivial bases when the transform is bypassed.

The pairs of impulse responses are shifted in time showing their support as basis of the space of the signal and four pairs are shown: the normal ELT, two transitory basis (impulse responses) and the transparent state impulse responses which are actually impulses. As another ex-

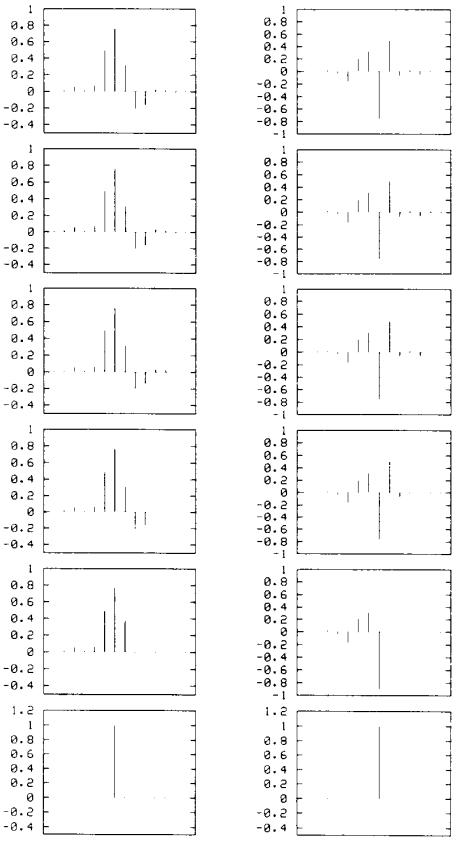

Fig. 12. Basis for the signal space during transition, for $K = 4$. On left it is shown the basis relative to the low-pass filters and on the right those relative to the high-pass one, at all stages of the transition to the bypass state. The topmost refers to regular ELT and at the bottom it is shown the trivial bases (transparent state). The first 4 bases used the DCT-IV as $\mathbf{Z}$ while the last in the transition used $\mathbf{Z} = \mathbf{I}$.

ample, the bases, for $K = 4$, at the transition to the transparent state are shown in Fig. 12.

C. Adaptation and Tree-Shaping

The best time-frequency representation of a signal, or the best wavelet packet, are abstract ideas. Almost all representations and tiling of the time-frequency plane may have their utility, including only frequency resolution (a transform such as DFT applied to the whole signal); only time resolution (no processing is applied to the input signal); fragmentation into blocks and use of M-band lapped transforms; wavelets (octave analysis); etc. . . . In [27], it was developed a method to find the best wavelet packet based on a rate-distortion criterion. In an independent work [19], this criterion was applied to search for an adaptive wavelet packet which would track the best tree shape. The method populates each node of the tree with a Lagrangian cost function $J_{ij}(\lambda) = D_{ij} + \lambda R_{ij}$, where D_{ij} and R_{ij} are the distortion and bit-rate associated with node η_{ij}, respectively, and λ is a Lagrange multiplier. Then, an algorithm is used to find the minimum cost for all possible set of terminal nodes, given a quality factor λ [19], [27]. For variation in time, a "double-tree" algorithm was employed in [19]. This algorithm was also applied to code speech signals, and it can be applied to our case, since we are providing the means to adapt the tree and not the adaptation cost function.

On the other hand, we can also use an ad-hoc simplifying strategy, which is solely based on intuition and concepts. This can simplify the adaptation procedure and even allow backward adaptation. If a stationary model is assumed, for minimum mean square error, the greater energy compaction in fewer coefficients [28] results in less distortion for a given bit-rate. This will generally lead us to choose the full tree which has better frequency resolution [23]. However, as a transform is bypassed, the filter for the resulting subband is shortened, therefore a better temporal localization is attainable. Furthermore, if we use shorter filters, the distortion in a coefficient in a particular subband would spread along a smaller region than if the filters were longer. Therefore, we can seek the maximum time resolution whenever not much energy compaction is provided by the transform. Let $x(m) = x_{ij}(m)$ and $x_L(n) = x_{i+1,2j}(n)$, $x_H(n) = x_{i+1,2j+1}(n)$. For this node, $a(n)$ is the activity signal. Let

$$\sigma_L^2 = E[x_L^2(n)] \quad \sigma_H^2 = E[x_H^2(n)]. \tag{43}$$

The above variances are related to the variance of $x(n)$ since the ELT has filters (with gain of $\sqrt{2}$, for orthonormality) which obey the power complementary property, i.e.,

$$\sigma_L^2 + \sigma_H^2 = 2\sigma_x^2. \tag{44}$$

The signal is not assumed stationary and the variances can be estimated continuously. Further computations over $x(n)$ would lead to more complexity and we can work directly with the decomposed signal. Then, a windowed estimation of the variance using a filter with impulse response $h(n)$ results in

$$\hat{\sigma}_L^2(n) = x_L^2(n) * h(n), \quad \sigma_H^2(n) = x_H^2(n) * h(n). \tag{45}$$

Using the estimated variances for energy compaction computations, we have a measure of the transform coding gain [28] as

$$G(n) = \frac{1}{2} \frac{\hat{\sigma}_L^2(n) + \hat{\sigma}_H^2(n)}{\hat{\sigma}_L(n)\hat{\sigma}_H(n)} \tag{46}$$

and we can compare $G(n)$ to a threshold g in order to decide if we set $a(n) = 1$ or not. Hence,

$$a(n) = u[G(n) - g] \tag{47}$$

where $u(x)$ is the step function.

1) Interrelation of the Nodes: To determine whether all nodes are to be made active or not we can start from the maximum available frequency resolution, i.e., check nodes in a maximum stage S and, then, their parents. At each node, we can evaluate $a_{ij}(n)$ as in (47).

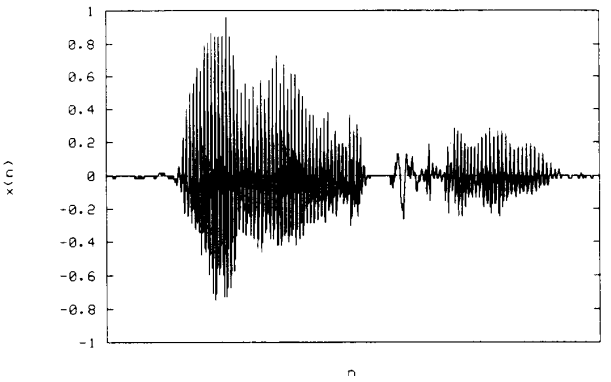

Fig. 13. Sample segment of speech used as test.

Start: $a_{ij} = 0$ (all ij); $n \leftarrow S$.
repeat
  for $m = 0 \cdots 2^n - 1$
    if $a_{nm} = 0$
      evaluate a_{nm}
      if $a_{nm} = 1$
        make $a_{ij} = 1$ for $\eta_{ij} \in$ path $\eta_{00} \rightarrow \eta_{nm}$.
    if all $a_{nm} = 1$ then stop else $n \leftarrow n - 1$.
until $n = 0$

2) Backward Adaptation: In order to prune or expand the same branches in analysis (transmitter) and synthesis (receiver), it is necessary to reconstruct $a_{ij}(n)$ at the receiver. In (46) and (47) we use ELT domain samples. If $G(n)$ in (46) uses quantized samples $\hat{x}_L(n)$, $\hat{x}_H(n)$ to estimate the variances and if, for a node η_{ij}, we use

$$a_{ij}(n+1) = u[G_{ij}(n) - g_{ij}] \tag{48}$$

then the receiver can recover a_{ij} without transmission of side-information, because it has available the past quantized ELT or time-domain samples and g_{ij} can be a fixed threshold. Both receiver and transmitter have to be synchronized, such that the transmitter has to use quantized values also when the transform is bypassed. Furthermore, the same nodal interrelation algorithm has to be used, always preparing future values of $a_{ij}(n+1)$. The filter $h(n)$ has to be causal and a simple first or second order IIR filter can be adequate, having a narrow low-pass bandwidth to avoid frequent transitions. Whenever a transition occurs, the states of the IIR filter may be reset, interrupting filtering until the transition is over. After that, filtering is resumed. This is because at a transition, the receiver will not have ELT samples or time-domain ones. Therefore, it will not be able to perform filtering on $\hat{x}_L^2(n)$, $\hat{x}_H^2(n)$ unless the time-domain samples are recovered and transformed again. Setting the filters to avoid frequency changes may help in this case.

3) Forward Adaptation: In case the activity map with all a_{ij} is sent in parallel, then things get much easier. First, there is no need to calculate activity on the receiver side. Second, one can use any means to determine activity of the nodes, including non-causal filters. A noncausal $h(n)$ is naturally preferred since it will avoid very short changes. The binary signal $a_{ij}(n)$ can be processed to avoid short bursts and locally oscillatory behaviors. A possible solution is a recursive median filter, which, in the binary case, is easily computed using tables. The formula for this is

$$a_{ij}(n) = \text{round [mean } (a_{ij}(n-k) \cdots a_{ij}(n+k))]. \tag{49}$$

This would prevent bursts of up to k isolated values of a_{ij} and would not oscillate if the input is an alternation of 0's and 1's. When an oscillation is encountered, the state just before it is preserved. The order in which the node activity is evaluated may be found using the algorithm described earlier.

The disadvantage is due to the transmission of side information, since all the activity map has to be transmitted. Assuming 1 bit per sample and a maximum stage number S, it would require $S/2$ bits per sample as overhead. (Remember that if node η_{00} works at a sampling rate f_s, node η_{ij} works in a sampling rate $2^{-i}f_s$.) However, assuming the filters would prevent very frequent changes, run-length coding can be applied to largely compress this map. Furthermore, as a node is active, all nodes connecting it to the root node will also be active. Therefore, information for them is not necessary.

Tests were made using the forward adaptation algorithm to code a segment of 8192 samples of a speech signal shown in Fig. 13, based on the ELT ($K = 2$, $M = 2$) and $S_{\max} = 6$. The transformed samples were coded using a uniform quantizer whose step size was varied. The entropy of the quantizer output plus tree-information was evaluated as a measure of the rate obtained. The measure of distortion was

$$D = \left(\frac{1}{8192} \sum_{n=0}^{8191} (x(n) - \hat{x}(n))^2 \right)^{1/2}$$

where $\hat{x}(n)$ is the reconstructed signal after synthesis. The plots of distortion versus entropy (DH) are shown in Fig. 14 for several threshold values g. Once the threshold g is

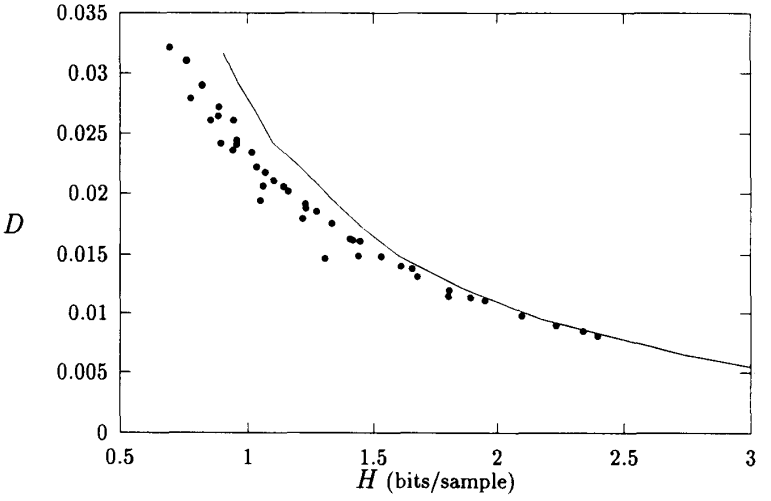

Fig. 14. Distortion (D) versus Entropy (H), in bits-per-sample, plot of the results simulating the adaptive wavelet packets, using several values of threshold and quantizer steps. The filter bank used was the ELT for $K = 2$ and $M = 2$. Solid line shows the $D \times H$ plot for the wavelet transform, using the same filter bank.

chosen (the same value g was applied to all nodes), we varied the quantizer step to obtain curves in the DH plane. The curves are not shown (only the points are shown), because we have tried several threshold values between $g = 2.5$ and $g = 4.0$ for several step sizes and the plot of each curve would be confusing. On solid line we have the same results obtained with the use of a regular wavelet transform. We can see the concentration of points below the solid line indicating the best performance of the adaptive scheme.

VI. Conclusions

We have presented time-varying structures which guarantee distortionless processing. Our main concern here does not lie on applications but on PR analysis–synthesis under time variations and structures for this purpose. The concept of an instantaneous filter bank was developed, along with the study of the transforms in their factorized forms. A general factorization was presented, allowing the use of time-varying systems which can switch between any paraunitary filter bank (following the restrictions of filters' length and maintaining the same number of channels), by a change of rotation angles. Thus, one can make the adaptation in order to perform distinct transforms for different regions of the signal. Also segmentation of the signal is possible and would allow to switch between different number of channels, paying the price of a longer transition region. Other possible strategies can be to switch between block and lapped transforms or change the overlap factor. This strategy was developed for the ELT ($K = 1$). The use of the ELT factorization gives us insights over varying systems, although restricting the filter bank responses and the design flexibility.

The adaptive wavelet packet structure is reasonably robust and the adaptation strategy, here presented, can be applied in the analysis of signals with variable statistics. An adaptive tree-path may be used when signal statistics vary from time to time and new conditions remain for a certain period, because of the transition regions. Clearly, each application would require a detailed study and extensive simulations. As a remark, there was no concern with regularity in transitions. We hope this method could be useful in applications not requiring too many stages in the wavelet packet tree. In this case, regularity would not be so crucial as it is when several stages are cascaded. The addition of regularity constraints in the transitions would surely pose a challenging problem.

Remarks

After the first version of this paper was submitted, we have learned that other researchers were obtaining similar results in the field of time-varying filter banks. This recent trend continuously produces new results and it has become impossible to track them while this paper was under review. However, the reader may find among the references interesting and different viewpoints. The work with time-varying filter banks in [11] gave us incentive to work in the field, while the finite-length solution for ELTs [3] inspired the use of time-varying (orthogonal) sparse factors. The concept of segmenting the signal through boundary filter banks is also reported in [19], [29]. While in [19], [29] the coefficients of the filters are adjusted, in a Gram–Schmidt orthogonalization procedure, here the parameters arising from a factorization of the filter bank in a lattice are changed with time. In [30], the same bypassing-segmentation approach was achieved

using another factorization of paraunitary filter banks, which is not based on plane rotations. Also, in [19], [29], there is a different approach to adapt the tree-shape, as we mentioned earlier. Lattice factorization for time-varying two-channel filter bank can be found in [31] to reduce Gibbs phenomena when coding image edges, in a work parallel to the one presented here and in [16]. Finally, time-varying filter banks for wavelet transforms are also explored in [32] for FIR filters and in [33] for IIR filters.

REFERENCES

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

[2] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks and applications: A tutorial," *Proc. IEEE*, vol. 78, pp. 56-93, 1990.

[3] H. S. Malvar, *Signal Processing with Lapped Transforms*. Norwood, MA: Artech House, 1992.

[4] M. Vetterli and D. Le Gall, "Perfect reconstruction filter banks: some properties and factorizations," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, pp. 1057-1071, July 1989.

[5] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Trans. Signal Processing*, vol. 40, pp. 2207-2232, Sept. 1992.

[6] A. N. Akansu, R. A. Haddad, and H. Caglar, "The binomial QMF-wavelet transform for multiresolution signal decomposition," *IEEE Trans. Signal Processing*, vol. 41, pp. 13-19, Jan. 1993.

[7] A. K. Soman and P. P. Vaidyanathan, "Paraunitary filter banks and wavelet packets," *Proc. ICASSP*, San Francisco, CA, 1992, vol. IV, pp. 397-400.

[8] H. S. Malvar, "Extended lapped transform: Fast algorithms and applications," *Proc. ICASSP*, Toronto, Canada, 1991, pp. 1797-1800.

[9] H. S. Malvar and D. H. Staelin, "The LOT: Transform coding without blocking effects," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 553-559, Apr. 1989.

[10] K. Nayebi, T. P. Barnwell, and M. J. Smith, "The time domain analysis and design of exactly reconstructing FIR analysis/synthesis filter banks," *Proc. ICASSP*, Albuquerque, NM, 1990, pp. 1735-1738.

[11] K. Nayebi, T. P. Barnwell, and M. J. Smith, "Analysis-synthesis systems with time-varying filter bank structures," *Proc. ICASSP*, San Francisco, CA, vol. IV, 1992, pp. 617-620.

[12] Z. Doğanata, P. P. Vaidyanathan, and T. Q. Nguyen, "General synthesis procedures for FIR lossless transfer matrices, for perfect reconstruction multirate filter banks applications," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 1561-1574, Oct. 1988.

[13] 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.

[14] T. Q. Nguyen, "A quadratic constrained least-square approach to the design of digital filter banks," *ISCS*, vol. 3, pp. 1344-1347, San Diego, CA, May 1992.

[15] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*. San Diego, CA: Academic Press, 1990.

[16] R. L. de Queiroz and K. R. Rao, "A structure for time-varying paraunitary filter banks with perfect reconstruction," *Electron. Lett.*, vol. 29, pp. 217-218, Jan. 1993.

[17] R. L. de Queiroz, "Subband processing of finite length signals without border distortions," *Proc. ICASSP*, San Francisco, CA, vol. IV, 1992, pp. 613-616.

[18] R. L. de Queiroz and K. R. Rao, "Reconstruction methods for processing finite-length with paraunitary filter banks," preprint.

[19] C. Herley, J. Kovacevic, K. Ramchandran, and M. Vetterli, "Arbitrary orthogonal tiling of the time-frequency plane," *Proc. Int. Symp. Time-Frequency and Time-Scale Anal.*, Victoria, BC, Canada, Oct. 1992.

[20] S. G. Mallat, "Multifrequency channel decomposition of images and wavelet models," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 2091-2110, Dec. 1989.

[21] I. Daubechies, "Orthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. XLI, pp. 909-996, 1988.

[22] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, pp. 14-38, Oct. 1991.

[23] R. L. de Queiroz and H. S. Malvar, "On the asymptotic performance of hierarchical transforms," *IEEE Trans. Signal Processing*, vol. 40, pp. 2620-2622, Oct. 1922.

[24] H. S. Malvar, "Fast computation of wavelet transforms with the extended lapped transform," *Proc. ICASSP*, San Francisco, CA, vol. IV, 1992, pp. 393-396.

[25] R. Coifman, Y. Meier, D. Quaker, and V. Wickerhauser, "Signal processing and compression with wave packets," preprint.

[26] R. Coifman and Y. Meier, "Orthonormal wave packet bases," preprint.

[27] K. Ramchandran and M. Vetterli, "Best wavelet packet bases using rate-distortion criteria," *Proc. ISCS*, San Diego, CA, vol. II, Apr. 1992, pp. 971-974.

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

[29] C. Herley, J. Kovacevic, K. Ramchandran, and M. Vetterli, "Tilings of the time-frequency plane: Construction of arbitrary orthogonal bases and fast tiling algorithms," *IEEE Trans. Signal Processing*, this issue.

[30] R. A. Gopinath and C. S. Burrus, "Factorization approach to unitary time-varying filter banks," Rice University, Houston, TX, Tech. Rep. CML-TR92-93, Dec. 1992.

[31] J. L. Arrowood and M. J. T. Smith, "Exact reconstruction analysis-synthesis filter banks with time-varying filters," *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Minneapolis, MN, vol. III, Apr. 1993, pp. 233-236.

[32] I. Sodagar, K. Nayebi, and T. P. Barnwell, "A class of time-varying wavelet transforms," *Proc. ICASSP*, Minneapolis, MN, vol. III, Apr. 1993, pp. 201-204.

[33] W. C. Chung and M. J. T. Smith, "Spatially-varying IIR filter banks for image coding," *Proc. ICASSP*, Minneapolis, MN, vol. V, Apr. 1993, pp. 570-573.

Ricardo L. de Queiroz (S'93) received the B.S. degree from Universidade de Brasilia in 1987 and the M.S. degree from Universidade Estadual de Campinas, Brazil, in 1990, both in electrical engineering.

In 1990-1991, he was with the DSP research group at Universidade de Brasilia as a Research Associate. He is currently enrolled in the Ph.D. program at the University of Texas at Arlington. His research interests are multirate signal processing, filter banks, image and audio compression, and image databases.

K. R. Rao (M'67-SM'73) received the Ph.D. degree in electrical engineering from the University of New Mexico, Albuquerque, in 1966.

Since 1966, he has been with the University of Texas at Arlington, where he is currently a Professor of electrical engineering. He has published extensively in reviewed technical journals in the areas of discrete orthogonal transforms and digital image coding. He, along with two other researchers, introduced the discrete cosine transform in 1975, which has since become very popular in digital signal processing. He has organized and conducted short courses and conferences on thermoelectric energy conversion from 1969 to 1992. He is the coauthor of the books *Orthogonal Transforms for Digital Signal Processing* (Springer-Verlag, 1975), *Fast Transforms: Analyses and Applications* (Academic Press, 1982), and *Discrete Cosine Transform—Algorithms, Advantages, and Applications* (Academic Press, 1990), as well as coauthor or editor on other works. He has conducted workshops on digital image coding worldwide.