VARIABLE BLOCK SIZE LAPPED TRANSFORMS

Ricardo L. de Queiroz

Xerox Corporation
800 Phillips Rd. M/S 128-27E
Webster, NY, 14580
queiroz@wrc.xerox.com

K. R. Rao

Electrical Engineering Department
University of Texas at Arlington
Box 19016, Arlington, TX, 76019
rao@ee.uta.edu

ABSTRACT

A structure for implementing lapped transforms with time-varying block sizes is discussed. It allows full orthogonality in the transitions based on a factorization of the transfer matrix into orthogonal factors. Such an approach can be viewed as a sequence of stages with variable-block-size transforms intermediated by sample-shuffling (delay) stages. Details for a first order system are given and a design example is presented.

1. INTRODUCTION

The choice of a block size in transform coding is a trade-off between time and frequency resolutions. Although not always explicitly pointed, a variable block size is a local adaptation aiming to obtain a better projection of the signal on the time-frequency plane. The block size dictates the trade-off. Larger blocks mean coarser time resolution and a larger number of narrower subbands, while small blocks mean better time localization and poorer frequency resolution. Block transforms of variable size can be easily applied to images as the problem is simplified to a tiling of the image into rectangular regions. We address the issues of perfect reconstruction (PR) and orthogonality using a lapped transform (LT) [1], i.e. a paraunitary uniform FIR filter bank (PUFB) [2], instead of a block transform. A number of papers recently dealt with time-varying filter banks [3]–[6]. However, few addressed a general framework to change the number of channels with orthogonality in the transitions and often the mechanism for change is applied to the filter coefficients. Here, as in [3], the change occurs in the structural factors that compose the filter bank. Hence, the transforms and the transitions are inherently orthogonal.

2. LAPPED TRANSFORMS

M-channel LTs can be described through their respective polyphase transfer matrices (PTM) [2] (the transfer matrix relating the M polyphase components of the input signal with the M subbands). The block size is M and the length of the filters is assumed to be $L = NM$. The entries of the PTM have terms of order $N-1$, i.e. polynomials up to z^{-N+1}. The PTM $\mathbf{E}(z)$ is assumed paraunitary such that $\mathbf{E}^{-1}(z) = \mathbf{E}^T(z^{-1})$. In other words, it is unitary in the unit circle ($z = e^{j\omega}$), and the analysis system leading the polyphase components to the subbands is lossless. For $N = 1$ (and real coefficients) it is clear that the PTM is a regular orthogonal matrix. The PTM can be parameterized using a cascade of delays and orthogonal stages as

$$\mathbf{E}(z) = \mathbf{B}_0 \prod_{i=1}^{N_s-1} (\mathbf{\Lambda}(z)\mathbf{B}_i), \quad (1)$$

where all stages $\mathbf{B}_i$ are allowed to be arbitrary $M \times M$ orthogonal matrices, N_s is the number of stages, and $\mathbf{\Lambda}(z)$ is a paraunitary matrix containing only delays. We assume that $\mathbf{\Lambda}(z)$ is a full-rank matrix with only one non-zero element per row or column, where each non-zero element can be 1 or z^{-1}. A general factorization of all PUFBs is the one where N_s is the McMillan degree of $\mathbf{E}(z)$ (i.e. the order of the determinant of $\mathbf{E}(z)$) and where $\mathbf{\Lambda}(z) = diag\{z^{-1}, 1, 1, \ldots, 1\}$ [2]. The symmetric delay factorization (SDF) [7] is the case where M is even, $N_s = N$, and

$$\mathbf{\Lambda}(z) = \begin{bmatrix} z^{-1}\mathbf{I}_{M/2} & 0 \\ 0 & \mathbf{I}_{M/2} \end{bmatrix} \quad (2)$$

with $\mathbf{I}_n$ as the $n \times n$ identity matrix.

3. TIME-VARYING LAPPED TRANSFORMS

For time-varying filters, however, the notation may not be that simple. For an invariant filter, if $y_i(m) = \sum_{n=0}^{q-1} h_{ij}(n)\ x_j(m-n)$, then the q-tap filter $h_{ij}(n)$ has z-transform $H_{ij}(z) = \sum_{n=0}^{q-1} h(n) z^{-n}$ and $Y_i(z) = H_{ij}(z)X_j(z)$. Let the filter have time-varying coefficients so that its input-output relation is

$$y_i(m) = \sum_{n=0}^{q-1} h_{ij}(m,n) x_j(m-n). \quad (3)$$

0-8186-7310-9/95 $4.00 © 1995 IEEE

We cannot say that $Y_i(z) = H_{ij}(z,m)X_j(z)$, neither use $h_{ij}(m,n)$ to replace the coefficients of $H_{ij}(z,m)$. However, because of the usual assumption of invariant filters, $H_{ij}(z)$ is often viewed as a description of an implementation algorithm and we simplify the notation by using $H_{ij}(z,m)$ to represent the filter coefficients at instant m. Therefore, z^{-1} may be viewed strictly in a systemic approach meaning a delay of the input, i.e. retrieve output from buffer and place input on buffer.

The key to get time-varying filter banks remaining instantaneously paraunitary [3] is to change the matrices $\mathbf{B}_i$ along the time index, but to keep them orthogonal matrices at all times. Thus, for PR orthogonal time-varying LTs (PUFBs) we can rewrite (1) as

$$\mathbf{E}(z,m) = \mathbf{B}_0(m) \prod_{i=1}^{N_s-1} (\mathbf{\Lambda}(z)\mathbf{B}_i(m)) \tag{4}$$

where $\mathbf{B}_i(m)$ changes with time but remains an $M \times M$ orthogonal matrix. $\mathbf{E}(z,m)$ remains paraunitary for all m, since

$$\mathbf{E}^T(z^{-1},m)\mathbf{E}(z,m) = \mathbf{\Lambda}^T(z^{-1})\mathbf{\Lambda}(z) = \mathbf{I}_M \tag{5}$$

i.e. $\mathbf{E}(z,m)$ is paraunitary because $\mathbf{B}_i(m)$ and $\mathbf{\Lambda}(z)$ (as we have defined it) are also paraunitary. From the above discussion on the z-transform of a time-varying filter, we apply the term *instantaneously paraunitary* [3, 7] to describe the property of $\mathbf{E}(z,m)$. Furthermore, the linear transform (not its state space representation) that leads all input samples into the subbands is an orthogonal transform [1, 3, 7].

4. CHANGING THE BLOCK SIZE (NUMBER OF CHANNELS)

With time-varying filter banks as described, the overall transform is orthogonal. In a transform point of view we can think of it as a linear transform $\mathbf{T}$ leading all input samples in vector $\mathbf{x}$ to the subband samples in vector $\mathbf{y}$.

$$\mathbf{y} = \mathbf{T}\mathbf{x} \quad \mathbf{x} = \mathbf{T}^T\mathbf{y} \tag{6}$$

where

$$\mathbf{T} = \tilde{\mathbf{B}}_0 \prod_{i=1}^{N-1} \mathbf{W}_i\tilde{\mathbf{B}}_i \tag{7}$$

$$\tilde{\mathbf{B}}_i = diag\{\cdots, \mathbf{B}_i(m-1), \mathbf{B}_i(m), \mathbf{B}_i(m+1), \cdots\} \tag{8}$$

and $\mathbf{W}_i$ is just a permutation matrix representing all the delays in a stage. As $\mathbf{B}_i(m)$ is orthogonal, so is $\tilde{\mathbf{B}}_i$, and $\mathbf{T}$ is orthogonal iff $\mathbf{W}_i$ is also orthogonal. As long as $\mathbf{W}_i$ implies simple permutations, it is obviously an orthogonal matrix. Therefore, the dimensions of $\mathbf{B}_i(m)$ have no effect on orthogonality of $\mathbf{T}$ as long as $\mathbf{B}_i(m)$ and $\mathbf{W}_i$ remain orthogonal. Clearly, we can change the number of channels of the filter bank by changing the sizes of $\mathbf{B}_i(m)$. Therefore, a PUFB with variable number of channels can be implemented through a sequence of stages with variable block size orthogonal transforms intermediated by shuffling (delay) stages. It is, in other words, a variable block size lapped transform. In a state-space approach, we can rewrite (4) as:

$$\mathbf{E}(z,m) = \mathbf{B}_0(m) \prod_{i=1}^{N_s-1} (\mathbf{\Lambda}(z,i,m)\mathbf{B}_i(m)) \tag{9}$$

where $\mathbf{\Lambda}(z,i,m)$ is the delay block at stage i and instant m. As in the case of (1) we also consider it a full rank matrix with only one non-zero element per row or column. An entry z^{-1} in $\mathbf{\Lambda}(z,i,m)$ at the k-th row and l-th column means: the k-th element of $\mathbf{B}_{i-1}(m)$'s input vector is the l-th element of $\mathbf{B}_i(m-1)$'s output vector. Similarly, the same occurs to a "1" entry with relation to $\mathbf{B}_i(m)$'s output vector. For the class of delay matrices described, in order to obtain orthogonality, it is necessary and sufficient to make $\mathbf{W}_i$ a permutation matrix. Necessary because otherwise $\mathbf{W}_i$ would be rank-deficient, and sufficient because $\mathbf{W}_i$ orthogonal leads to $\mathbf{T}$ orthogonal as we saw. Now let us examine the translation of $\mathbf{W}_i$ (as a permutation matrix) to the system described in (10). It is obvious that the set of columns of $\mathbf{\Lambda}(z,i,m)$ with the entry z^{-1} has to match all columns of $\mathbf{\Lambda}(z,i,m-1)$ that are all 0's as well as those which contain a z^{-1} entry. Let

$$\mathbf{E}(z,m) = \mathbf{E}_0(z,m)\mathbf{\Lambda}(z,i,m)\mathbf{E}_1(z,m) \tag{10}$$

where $\mathbf{E}_0(z,m)$ is a $k \times k$ paraunitary matrix while $\mathbf{E}_1(z,m)$ is an $l \times l$ one. Obviously, $\mathbf{\Lambda}(z,i,m)$ is a $k \times l$ matrix. If $k < l$,

$$\mathbf{\Lambda}^T(z^{-1},i,m)\mathbf{\Lambda}(z,i,m) = \mathbf{I}_k \tag{11}$$

$$\mathbf{E}^T(z^{-1},m)\mathbf{E}(z,m) = \mathbf{I}_k. \tag{12}$$

On the other hand, if $k > l$, these relations do not hold. The stationary relations do not hold because the sizes of the matrices involved also change. Assume a transition from M_1 to M_2 channels (block size changes from M_1 to M_2 samples). With a quick transition in mind, we assume that $\mathbf{E}_0(z,m-1)$ and $\mathbf{E}_1(z,m-1)$ are paraunitary matrices of size $M_1 \times M_1$, while $\mathbf{E}_0(z,m+1)$ and $\mathbf{E}_1(z,m+1)$ are paraunitary matrices of size $M_2 \times M_2$. Let d'_m be the number of delays in $\mathbf{\Lambda}(z,i,m)$ and d_m be the respective number of "1" entries. Thus,

$$\begin{aligned} d_{m-1} + d'_{m-1} &= M_1 \quad d_m + d'_m = k \\ d_{m+1} + d'_{m+1} &= M_2 \quad d_m + d'_{m+1} = l \end{aligned} \tag{13}$$

With the assumption that

$$d_{m+1} - d_{m-1} = d'_{m+1} - d'_{m-1} \tag{14}$$

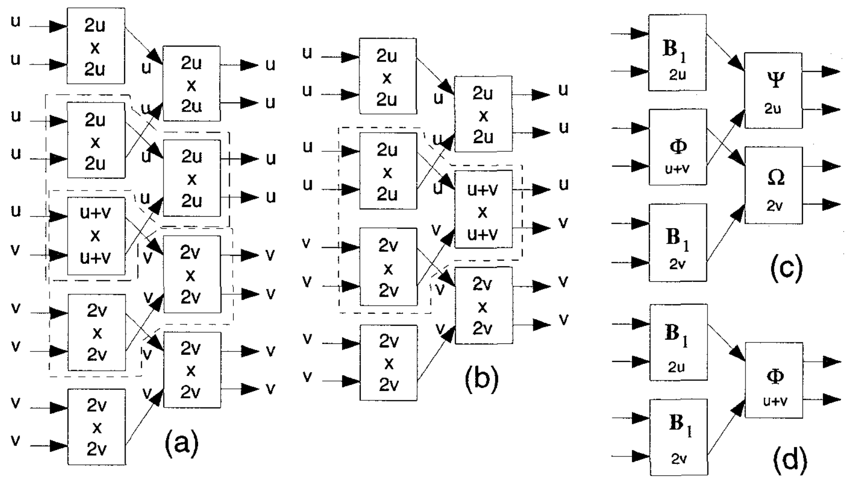

Figure 1: Flow-graph for variable block size lapped transform switching from $M_1 = 2u$ channels to $M_2 = 2v$ channels. (a) Direct method. There are 2 transition filter banks: one with $2u$ and the other with $2v$ channels. (b) Slow switch stepping through an intermediate filter bank of $u + v$ channels. (c)(d) Details of the transitions in (a)(b). $\mathbf{\Phi}$, $\mathbf{\Omega}$ and $\mathbf{\Psi}$ are the matrices to be optimized.

we have that

$$l - k = (M_2 - M_1)/2. \tag{15}$$

Hence, we see that $k \neq l$ is a necessary restriction, as long as $M_1 \neq M_2$. Furthermore, we cannot make l and k equal to M_1 or M_2. In other words, when changing the number of channels from M_1 to M_2, one cannot implement the transition using only block transforms of size M_1 or M_2. The process has to step over intermediary sizes for $B_i(m)$. The same conclusion can be reached using the general factorization, i.e. using $d'_{m-1} = d'_{m+1}$ instead of (14).

5. FIRST ORDER EXAMPLE

For a first order algorithm using SDF we have

$$\mathbf{E}(z) = \mathbf{B}_0 \mathbf{\Lambda}(z) \mathbf{B}_1 \tag{16}$$

where $\mathbf{\Lambda}$ is as in (2). The filters have length which is twice the block size. Then, we change the block size (number of channels) from $M_1 = 2u$ to $M_2 = 2v$. As we discussed, the transition has to contain blocks with size different than M_1 or M_2, and we present two distinct switching methods in Fig. 1. In this figure, the number of samples carried in each branch is indicated. In Fig. 1(a) it is shown the direct switch method. Note that there are 2 transition filter banks one with $2u$ channels and the other with $2v$ channels. This is because we used a block transform of size $(u + v) \times (u + v)$ in the first stage. Thus, is does not affect the number of channels, but the length of the filters. The filters have length $3u + v$ and $3v + u$ respectively. In Fig. 1(b), it is shown the second alternative which is a slow switch stepping through an intermediate filter bank of $u + v$ channels (with filters of $2(u + v)$ taps) which would be the only transitory filter bank. In Fig. 1 the matrices to be optimized are $\mathbf{\Phi}$, $\mathbf{\Omega}$, and $\mathbf{\Psi}$.

The framework in Fig. 1 is general for lapped transforms with $L = 2N$. In a design example, we use (as stationary filter bank) the modulated lapped transform (MLT) [1]. The MLT is a special cosine-modulated filter bank possessing filters with good stopband atenuation allied to a fast implementation algorithm [1]. An algorithm for a variable block size MLT was already presented in [4]. The approach in [4] requires explicit changes in the basis functions, while our approach is based on designing the orthogonal factors composing a lapped transform [3, 7]. Hence, we work directly with a fast implementation algorithm, which is maintained. The MLT is just an example and the generality of the framework allow us to replace it by any other LT.

Design examples are shown in Fig. 2 showing the transition filters for a switch from a 4-channel MLT to an 8-channel MLT. In the direct (hard) switch method there are 2 transition filter banks, one with 4 channels and 10 taps and the other with 8 channels and 14 taps. If the slow (soft) switch is applied there is only one transition filter bank with 6 channels and 12 taps. In both cases, the transition filters were optimized for maximum coding gain of the instantaneous filter bank [3][7] for an AR(1) spectrum with correlation coeffcient 0.95. Note that the transitory LTs are not required to be cosine-modulated filter banks. However, the lower-frequency basis functions of these transforms resemble the bases of the MLT, while the higher frequency bases (which have less significance to the coding gain) do not.

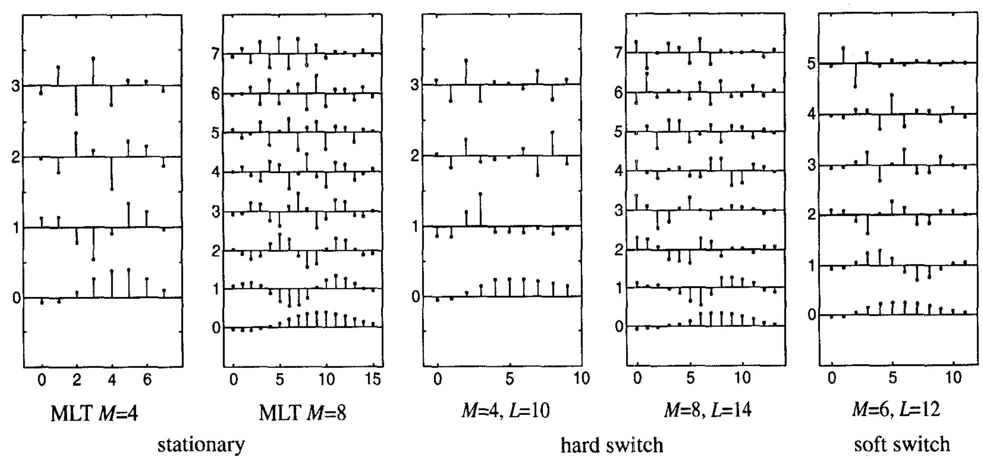

Figure 2: Impulse response of the filters of the time-varying MLT, switching from 4 to 8 channels. It is shown the normal MLTs for $M = 4$ ($L = 8$) and $M = 8$ ($L = 16$), the two transition filter banks for the direct switch and the 6-channels filter bank in the case of a slow switch. The transition filters were optimized for maximum coding gain. L is the length (in taps) of the filters.

6. REMARKS

Time-varying wavelet packet trees - Filter banks can be associated in such a way to construct discrete wavelet and wavelet packets transforms. The association of these filter banks follow the paths of a tree. By dynamically pruning or expanding branches of the tree one can continuously reshape the tree-paths. Such an approach is a time-varying wavelet packet transform. In this fashion, one can obtain a better tiling of the time-frequency plane. Often, the process is applied to a fixed binary tree. Each node is either a leaf or it is split into two branches generating two child nodes. The extension to M-ary trees is trivial, but the problem is still reduced to an on-off decision (prune or expand branches). Prunning a branch can be viewed as applying a 1-channel filter bank. Using LTs with time-varying block size (number of channels) we can achieve a much enhanced tiling of the time frequency plane. This is done by not only deciding upon prunning or not a branch but by deciding the number of channels to assign to each node in the tree (including the 1-channel case). Therefore, for each tree node, the decision is no longer a binary problem and an enhanced tiling can be achieved.

Variable block size transform coding - In a one-node version of the time-varying wavelet-packet we can use a single variable size transform for image compression [8, 9]. Variable block size coding and quadtrees have been successfully used to encode images [8]–[11]. The approach described here can be readily applied to (i) replace the DCT in [8, 9] or to (ii) compress images using variable block size as in [10, 11], but on LT domain.

7. REFERENCES

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

[2] P.P. Vaidyanathan, *Multirate Systems and Filter Banks.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[3] R.L. de Queiroz and K. R. Rao, "Time-varying lapped transforms and wavelet packets," *IEEE Trans. on Signal Processing,* vol. 41, pp. 3293–3305, Dec. 1993.

[4] 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. on Signal Processing,* vol. 41, pp. 3341–3359, Dec. 1993.

[5] I. Sodagar, K. Nayebi and T. P. Barnwell, "Time-varying filter banks and wavelets", *IEEE Transactions on Signal Processing,* vol. 42, pp. 2983–2996, Nov. 1994.

[6] R. Gopinath and C. S. Burrus, "Factorization approach to unitary time-varying filter banks trees and wavelets," *IEEE Trans. on Signal Processing,* vol. 43, pp. 666-680, Mar. 1995.

[7] R. L. de Queiroz, *On Lapped Transforms,* Ph.D. Dissertation. University of Texas at Arlington, Aug. 1994.

[8] C. Chen, "Adaptive transform coding via quadtree-based variable blocksize DCT," *Proc. of ICASSP,* May 1989.

[9] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate distortion sense," *IEEE Trans. on Image Processing,* vol. 2, pp. 160–175, Apr. 1993.

[10] G. Sullivan and R. Baker, "Efficient quadtree coding of images and video," *IEEE Trans. on Image Processing,* vol. 3, pp. 327–331, May 1994.

[11] J. Vaisey and A. Gersho, "Image compression with variable block size segmentation," *IEEE Trans. on Signal Processing,* vol. 40, pp. 2040–2060, Aug. 92.