ARBITRARY ORTHOGONAL TILINGS OF THE TIME-FREQUENCY PLANE *

Cormac Herley [1], *Jelena Kovačević* [2], *Kannan Ramchandran* [1] *and Martin Vetterli* [1]

[1]Department of EE and CTR, Columbia University, New York, NY 10027
[2]AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974

ABSTRACT

In this paper we consider expansions which give arbitrary orthonormal tilings of the time-frequency plane. These differ from the short-time Fourier transform, wavelet transform, and wavelet packets tilings in that they change over time. We show how this can be achieved using time-varying orthogonal tree structures, which preserve orthogonality, even across transitions. One method is based on lapped orthogonal transforms, which makes it possible to change the number of channels in the transform. A second method is based on the construction of boundary filters, and gives arbitrary tilings. We present an algorithm which for a given signal decides on the best binary segmentation, and which tree split to use for each segment, and is optimal in a rate-distortion sense. We present the results of experiments on test signals.

1 INTRODUCTION

Recently there has been a renewal of interest in linear expansions of signals, particularly using wavelets and some of their generalizations (see, for example, [5] and references therein). It is well known that the classical short-time Fourier tranform or Gabor transform, and the more recent wavelet transform are just two of many possible tilings of the time-frequency plane. These are illustrated in Figures 1(a) and (b). We use the term "time-frequency tile" of a particular basis function to designate the region in the plane which contains most of the function's energy. An elegant generalization that contains, at least conceptually, Gabor and wavelet transforms as special cases, is the idea of wavelet packets [8] or arbitrary subband coding trees. An example of a wavelet packet tiling is given in Figure 1(c). While the wavelet packet creates an arbitrary slicing of frequencies (with associated time resolution), it does not change over time. Often a signal is first segmented, and the wavelet packet decomposition is performed on each segment independently. An obvious question is whether we can find a wavelet packet decomposition that changes over time, that is, an arbitrary orthogonal tiling of the time-frequency plane. An example of such a generalized tiling is shown in Figure 1(d). We use the term "arbitrary" somewhat casually, since the tiling is restricted to those produced by binary tree structures. However, the wavelet packet construction is generalized sufficiently to warrant the term.

Some of the key questions to be answered are:

- How general can the basis functions be ? Can, for example, adjacent wavelet packet decompositions overlap in order to avoid discontinuities due to segmentation ?
- For a given signal, what is the optimal tiling, and can it be found efficiently ?
- What relationship exists between discrete-time and continuous-time constructions ? In other words, is it possible to construct continuous-time bases from discrete-time ones as in the wavelet case ?

The goal of this paper is to answer these questions. We propose bases, develop algorithms and carry out experiments to verify performance.

The paper is structured as follows. We begin by constructing adaptive orthonormal systems. The first, in Section 2, is an adaptive lapped orthogonal transform (LOT), allowing one to change the number of channels on the fly, with overlapping basis functions. The second, based on orthogonal boundary filters treated in Section 3, allows us to start a frequency segmentation at any point in time, and thus can be used for arbitrary tilings. Given such arbitrary bases, it is necessary to find efficient algorithms to choose the best basis for a given input signal. This is done in Section 4, where the best tiling in a rate-distortion sense is found using an extension of the wavelet packet algorithm [4]. Section 5 shows the results of some experiments on speech signals.

2 VARIABLE SIZE LOT'S

We start by examining the first question, namely whether it is possible for the basis functions of the adjacent decompositions to overlap in order to avoid discontinuities due to segmentation. In this section, we present a particular construction achieving the overlap but leading to very restricted tiling given in Figure 2(a). It will be referred to as "variable size LOT's".

LOT's are a special class of perfect reconstruction filter banks (see, for example, [3]), using a single prototype filter of length $2N$ (where N is the number of channels) to construct all of the filters $h_0, \ldots, h_{N-1}$, by modulation as follows:

$$h_k(N,n) \;=\; \frac{h_{pr}(N,n)}{\sqrt{N}} \cdot \cos\Big(\frac{(2k+1)}{4N}(2n-N+1)\pi\Big), \quad (1)$$

with $k=0,\ldots,N-1$, $n=0,\ldots,2N-1$ and where the prototype lowpass filter $h_{pr}(N,n)$ is usually symmetric. The question we want to answer now is whether it is possible to switch between different LOT's with overlapping basis functions, while preserving perfect reconstruction and orthogonality. The answer is positive and here we summarize it. For more details, refer to [2].

Theorem 2.1 *One can switch from an N_1-channel LOT to an N_2-channel LOT (where $N_1 < N_2$ and N_1, N_2 even) as follows:*
1. Set

$$h_{pr}(N_2,k,n) \;=\; \begin{cases} 0 & n \in [0, \frac{N_2-N_1}{2}-1], \\ 0 & n \in [\frac{3N_2+N_1}{2}, 2N_2-1], \\ \sqrt{2} & n \in [\frac{N_2+N_1}{2}, \frac{3N_2-N_1}{2}-1] \end{cases} \quad (2)$$

*WORK SUPPORTED IN PART BY THE NATIONAL SCIENCE FOUNDATION UNDER GRANTS ECD-88-11111 AND MIP-90-14189.

2. Choose

$$h_k(N_1,n) \;=\; \sqrt{\frac{N_2}{N_1}}h_{i_k}(N_2,n+\frac{N_2-N_1}{2}), \qquad (3)$$

for $n \in [0, N_1-1]$ *and*

$$h_k(N_1,n) \;=\; \sqrt{\frac{N_2}{N_1}}h_{i_k}(N_2,n+\frac{3N_2-N_1}{2}), \qquad (4)$$

for $n \in [N_1, 2N_1-1]$ *and where* $i_k \in [0, N_1-1]$ *and* $i_k \neq i_j, \forall k, j$.

The proof is very simple and can be found in [2]. The idea behind the construction is the following: start with an N_2-channel LOT and construct an N_1-channel LOT by taking N_1 basis functions from the N_2-channel LOT and discarding the middle coefficients where the prototype filter is constant. Therefore, effectively we have constructed an N_1-channel LOT from the tails of the N_2-channel one (a similar scheme can be found in [8], and this also resembles what the signal processing community calls MDCT with a 50% overlap). As an example, consider switching from a 4-channel to an 8-channel LOT. It yields the tiling of the time-frequency plane as given in Figure 2(a). As mentioned earlier, the types of tilings we can obtain in this fashion are restricted. Another point is that the amount of overlap is fixed to the size of the smaller LOT (in this example, it would be 4).

Using the same type of a reasoning, one can construct even more sophisticated schemes. For example, one could switch from a 2-channel to a 4-channel to an 8-channel LOT. The type of tiling one could produce in this manner is given in Figure 2(b). For the details of the construction, refer to [2].

3 SWITCHING ORTHOGONAL TREES

We now consider a different approach, which will allow us to realize the most arbitrary tree-based tiling, as depicted in Figure 1(d).

Consider the following simple example by way of illustration: suppose we wish to split a signal $x(n)$ using length 4 orthogonal analysis filters $H_0(z)$ and $H_1(z)$, but only for time $n_0 \leq n$. Note that the relation between the analysis filters is $H_1(z) = H_0(-z^{-1})$ [7].

The matrix corresponding to the filter bank operating on an infinite signal can be written as a doubly infinite unitary block Toeplitz matrix [7]. In restricting the matrix to the half infinite range $n_0 \leq n$, however, one of two things happens: either the matrix fails to be unitary, or the matrix becomes rank deficient. Orthogonality is required by the fast tree pruning algorithm which we use in the next section, and if the matrix is rank deficient it has a null-space, which means that certain non-zero signals would be "invisible" to the system. We wish to avoid both of these undesirable effects.

We overcome both problems by finding boundary filters. For example, we will show how to find the filters $h_0'(n), h_1'(n)$ such that the operator $\mathbf{T}_1$ in (5) is unitary without a null-space.

$$\mathbf{T_1} = \begin{bmatrix} h_0'(0) & h_0'(1) & h_0'(2) & 0 & 0 & \cdots \\ h_1'(0) & h_1'(1) & h_1'(2) & 0 & 0 & \cdots \\ 0 & h_0(0) & h_0(1) & h_0(2) & h_0(3) & \cdots \\ 0 & -h_0(3) & h_0(2) & -h_0(1) & h_0(0) & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \qquad (5)$$

Constructions of this kind will allow us to change between one orthogonal tree and another, without losing the orthogonality property. For example, suppose we wished to grow a tree using the filters $\{h_0(n), h_1(n)\}$ starting at some time n_0. We could let $\mathbf{I}_0$ be the half infinite identity matrix containing ones on the diagonal as far as $n_0 - 1$, and then use the operator

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_1 \end{bmatrix},$$

which gives the appropriate transition at time n_0. In this way we will be able to change between different trees, and grow branches at will, provided the appropriate boundary filters are used. Pruning a tree is done similarly, boundary filters are found which allow us to cease splitting a signal with a certain set of filters.

Solutions

To illustrate the procedure we solve the particular case already introduced, the length 4 example (see [2] for the more general case). We claim that a solution can be found, when the boundaries are as shown in (5); thus we must find the coefficients of the boundary filter pair at the top of the matrix in terms of those of $h_0(n)$ and $h_1(n)$. For convenience denote by r_i the i-th row of the matrix. To make the matrix unitary we must hence ensure $r_i r_j^T = \delta_{ij}$.

To force r_1 and r_2 to be orthogonal to r_3 and r_4 gives

$$\begin{pmatrix} h_0'(1) & h_0'(2) \\ h_1'(1) & h_1'(2) \end{pmatrix} = \begin{pmatrix} k_0 & 0 \\ 0 & k_1 \end{pmatrix} \cdot \begin{pmatrix} h_0(1) & -h_0(0) \\ h_0(1) & -h_0(0) \end{pmatrix}, \qquad (6)$$

for any k_0 and k_1. Orthogonality of r_1 and r_2 requires

$$h_0'(0)h_1'(0) + k_0k_1h_0(1)^2 + k_0k_1h_0(0)^2 \;=\; 0. \qquad (7)$$

Hence we can arbitrarily choose $h_0'(0)$ and then determine $h_1'(0)$ from (7), and the remaining coefficients from (6). Normalization of $h_0'(n)$ and $h_1'(n)$ gives k_0 and k_1.

Boundary scaling functions and wavelets

In [1] it was shown how discrete-time orthonormal bases could be used to generate continuous-time ones. The above construction shows how to find time-varying discrete-time orthonormal bases. We shall apply an approach similar to that of [1] to our novel time-varying bases, and suggest how this may be used to find compactly supported wavelet bases for functions which are supported on an interval.

We again use the length 4 case for illustration. Consider the half-infinite block Toeplitz matrix $\mathbf{H}_0$ which contains as rows the shifted impulse response of $h_0(n)$, and has the boundary filter in the first row.

If $L_{ik}(z)$ denotes the z-transform of the coefficients of the i-th row of $(\mathbf{H}_0)^k$, then it can be shown that

$$L_{ij}(z) = z^{-1-2^j(i-2)} \prod_{k=0}^{j-1} H_0(z^{2^k}) \quad i > 1. \qquad (8)$$

The function $L_{2j}(z)$ can easily be recognized as the z-transform of the "graphical iteration" [1] to find the scaling function of a compactly supported wavelet scheme (see [7]). That is, if we define from $L_{ij}(z)$ a continuous-time function

$$f^{(j)}(x) = l_{2j}(n) \qquad n/2^j \leq x < (n+1)/2^j,$$

it can be shown that $f^{(j)}(x)$ converges to the scaling function $\phi(x)$ as $i \to \infty$ (under some constraints on $h_0(n)$).

Because of the boundary, a special form is required to calculate $L_{1j}(z)$

$$L_{1j}(z) = h_0'(0)L_{1,j-1}(z) + z^{-1}E(z^{2^{j-1}}) \prod_{k=0}^{j-2} H_0(z^{2^k}), \qquad (9)$$

where $E(z) = h_0'(1) + h_0'(2)z^{-1}$. As for $f^{(j)}(x)$ we define $b^{(j)}(x)$, as a continuous-time function derived from $L_{1j}(z)$. If $b^{(j)}(x)$ converges as $j \to \infty$ we call it the boundary function. That it is orthogonal to $\phi(x-k)\ \forall\ k \in Z$ follows from orthogonality of the filter set. Similarly it is compactly supported since the filters are FIR.

By way of illustration, Figure 3 shows the wavelets generated by the Daubechies length 4 filters [1], with appropriate boundary functions, for an interval of length 4. The figure shows $\psi(x), \psi(x-1)$ together with the boundary functions at the $x = 0$ and $x = 4$ boundaries.

It is worth emphasizing that we have given the analysis for the length 4 case; the number of boundary functions depends on the length of the filters used. Also note that the boundary function, which is determined by (9) does not obey a two-scale difference equation, and that the question of convergence is less clear.

Orthogonality of the boundary scaling function and wavelet with respect to integer translates follows from the orthogonality of the rows of $\mathbf{H}_0$ and $\mathbf{H}_1$. Orthogonality across scales is verified similarly. This is the machinery we need to derive wavelet bases for the interval. A more detailed account is given in [2].

4 R–D OPTIMAL TILING

From a coding perspective, the optimal orthonormal (ON) tiling of the time–frequency plane should be in the operational rate–distortion (R–D) sense. Thus, an optimal tiling for one coding application corresponding to a particular wavelet kernel, quantizer set and coding scheme need not be the same as that for another. In the R–D framework, the optimal tiling is that which minimizes total distortion subject to a maximum total bit rate constraint, or conversely, which minimizes the bit rate subject to a maximum distortion constraint.

Optimal bit allocation

The topic of optimal bit allocation, a constrained optimization problem (COP) where the distortion is minimized subject to a target bit budget constraint (or vice versa), has received exhaustive study [6]. For an ON decomposition and an additive distortion measure (such as the mean square error – MSE), the "hard" constrained optimization problem can be solved by converting it to an "easy" equivalent unconstrained optimization problem (UOP) via the Lagrange multiplier λ which "trades off" rate for distortion [4, 6]. The UOP is the minimization, over all permissible operational (R,D) points, of the Lagrangian cost $D + \lambda R$, where D is the distortion and R the rate. As shown in [4], the UOP solution for the "correct" value of λ, λ^*, is the desired convex–hull operating point for the original COP as well. The optimal operating slope λ^* is obtained by performing a fast convex recursion in λ. See [4] for details.

Brief Review of Wavelet Packet Algorithm

Optimal bit allocation in a wavelet packet (WP) framework has been solved in [4] using an R–D criterion to find the "best WP basis" decomposition for a given signal. The basic idea is that the full-depth WP tree is populated with the Lagrangian cost $D(node) + \lambda R(node)$ for each internal tree node. Then, a fast pruning algorithm, based on Bellman's optimality principle, is used to prune the full-depth tree into that subtree which has minimum total sum-of-leaves cost. The basic pruning criterion applied at each node is that of deciding in favor of the parent or its children based on which has the lower Lagrangian cost (for a fixed quality factor λ).

Arbitrary tiling: the double tree algorithm

In order to solve the problem of finding the R–D optimal tiling of the time–frequency plane, we extend the fast WP algorithm outlined above to a "double tree" algorithm. This is easiest explained through an example. Consider Figure 4. Assume a length-4 input signal [1,2,3,4] and a Haar basis as the wavelet kernel. To find the optimal split, optimal WP subtrees are found for all possible binary signal subsets: {[1,2,3,4],[1,2],[3,4],[1],[2],[3],[4]} as shown in Figure 4(a). A scalar quantizer of step size 4 has been picked for this example to quantize all WP coefficients. As before, a Lagrangian cost criterion is used for the optimal tree pruning operation ($\lambda = 0$ shown in the figure). Then, the costs associated with the best bases determined in the first step are used to populate a second tree called *the splitting tree* as shown in Figure 4(b). The root of the splitting tree is populated with the cost associated with the best basis WP for the [1,2,3,4] signal split, the first tree level with the two costs corresponding to the [1,2] and [3,4] splits respectively, etc.. The splitting tree is pruned using the identical fast algorithm as that used to find the WP trees whose costs populate its nodes. The optimal operating slope λ^* is found as in [4].

5 EXPERIMENTAL RESULTS

The double tree algorithm was used to find the optimal tiling for several test signals. The optimal split for a 512 point speech segment to a maximum depth of 7 using the Daubechies length 4 filter with its boundary filters is shown in Figure 5. A quantizer step of 20 was used and first order entropy and MSE were used as the rate and distortion measures. The optimal split achieves 1.99 bits per sample (bps) with 27.98 MSE, while the best basis WP tree needs 2.44 bps with 28.33 MSE, highlighting the usefulness of efficient binary tiling. Note that the overhead of sending the splitting map has been included.

REFERENCES

[1] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, XLI:909–996, 1988.

[2] C. Herley, J. Kovačević, K. Ramchandran, and M. Vetterli. Orthogonal tilings of the time-frequency plane. *IEEE Trans. Signal Proc., Special Issue on Wavelets*, 1992. To be submitted.

[3] H.S. Malvar. *Signal Processing with Lapped Transforms.* ARTECH House, 1992.

[4] K. Ramchandran and M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Trans. on Image Proc.*, 1992. Submitted.

[5] O. Rioul and M. Vetterli. Wavelets and signal processing. *IEEE Signal Proc. Mag.*, Oct. 1991.

[6] Y. Shoham and A. Gersho. Efficient allocation for an arbitrary set of quantizers. *IEEE Trans. Acoust., Speech, Signal Proc.*, 36:1445–1453, 1988.

[7] M. Vetterli and C. Herley. Wavelets and filter banks: theory and design. *IEEE Trans. on Signal Proc.*, Sept. 1992. To appear.

[8] M. V. Wickerhauser. INRIA lectures on wavelet packet algorithms. Technical report, Dept. of Mathematics, Yale University, Mar. 1991.

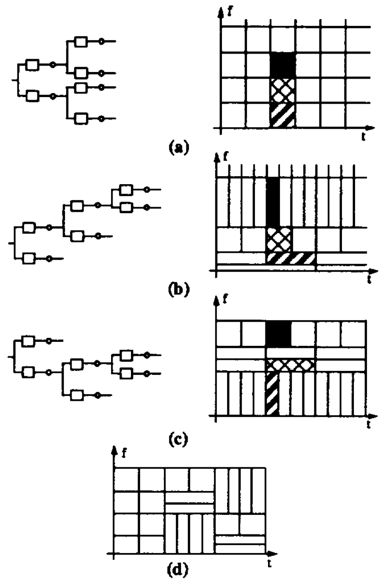

Figure 1: Tilings of the time-frequency plane. (a) Short-time Fourier transform tiling. (b) Wavelet tiling. (c) Wavelet packet tiling. (d) Generalized tiling which adapts in time as well as in frequency.

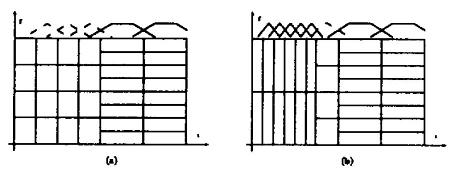

Figure 2: Variable size LOT's. Thick lines denote the switching point. On top of each tiling, basis functions with appropriate overlaps are given. (a) Switching from a 4-channel to an 8-channel LOT. (b) Switching from a 2-channel to a 4-channel to an 8-channel LOT.

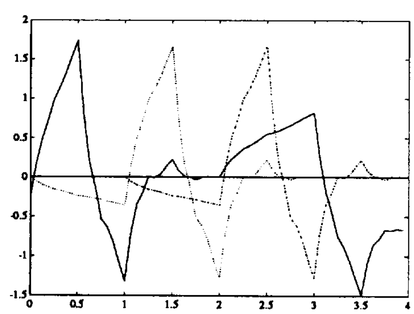

Figure 3: Boundary functions for wavelet basis for the interval $[0,4)$. Left and right boundary functions are supported on $[0,2)$ and $[2,4)$ respectively.

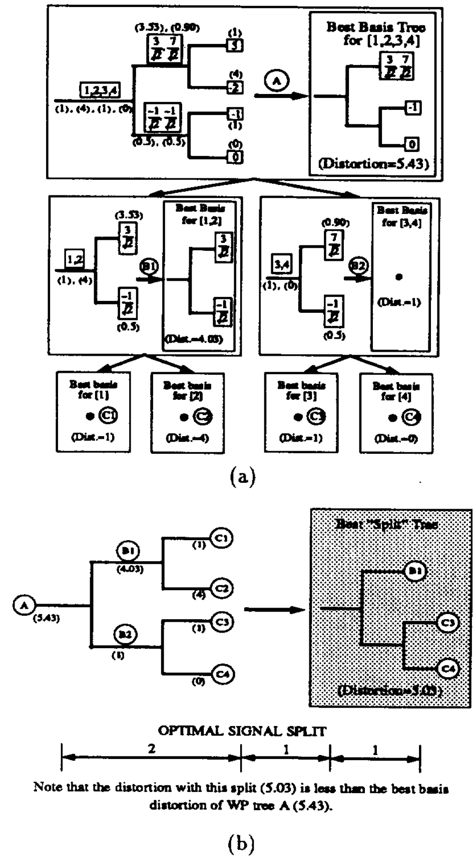

Figure 4: The double tree optimal ON splitting algorithm on the input signal [1,2,3,4] in R^4 for the Haar kernel and a scalar quantizer of step size 4. Lagrangian costs are shown in brackets ($\lambda = 0$ used here). (a) The best basis WP subtrees corresponding to all feasible signal subsets. (b) The splitting tree whose nodes are populated from the best basis costs of (a).

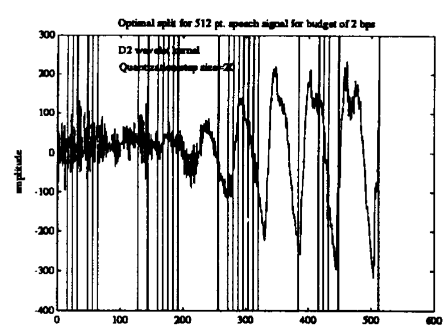

Figure 5: Optimal split for a 512 point speech signal for the D2 filter, a scalar quantizer of step size 20, and maximum depth 7. First order entropy and MSE are used as the rate and distortion measures.