TIME-VARYING ORTHONORMAL 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

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 orthogonal boundary filters; these allow us to construct essentially arbitrary tilings. We present a double-tree algorithm which for a given signal decides on the best binary segmentation, and which tree split to use for each segment. That is, it is a joint optimization of time and frequency splitting. The algorithm is optimal for additive cost functions (e.g. rate-distortion). This gives best time-varying bases. Results of experiments on test signals are shown.

1 INTRODUCTION

There has been a renewal of interest in linear expansions of signals, particularly using wavelets and some of their generalizations (see, for example, [7] 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 [1] 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.

The goal of this paper is to futher develop such arbitrary tilings, as well as to describe the double tree algorithm to find an optimal tiling for a given signal.

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

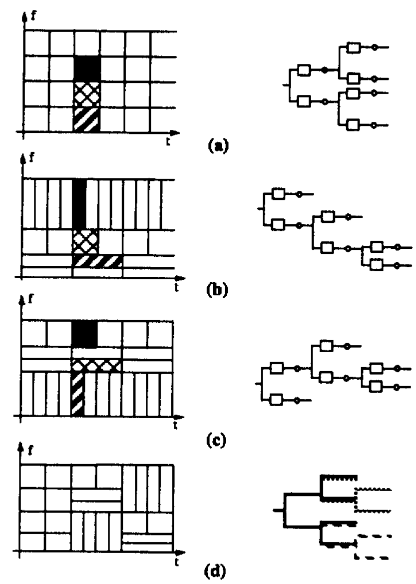


Figure 1: Tilings of the time-frequency plane. (a) Short-time Fourier transform tiling. (b) Wavelet tiling. (c) Wavelet packet tiling. (d) Generalized tiling which adapts in time as well as in frequency.

2 ADAPTIVE MODULATED LAPPED TRANSFORMS

We mentioned previously, that the wavelet packets (see Figure 1(c)) produce "arbitrary" slicing of frequencies, but they obviously do not change over time. Consider what would happen if we would exchange axes in Figure 1(c). In this new tiling, the time is sliced up, while the frequency division is uniform, which can be seen as the dual of wavelet packets. Now which system can produce such a tiling? The uniform division of the spectrum can be obtained with many different filter banks structures, one of them being the so-called *modulated lapped transforms (MLT)* [4]. Moreover, we will show that it is possible to switch between different MLT's over time, giving the dual of wavelet packets discussed above. We will call these constructions *adaptive modulated lapped transforms.*

MLT's are a special class of perfect reconstruction filter banks (see, for example, [4]), 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:

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

with $k = 0, \ldots, N-1$, $n = 0, \ldots, 2N-1$. Here the prototype lowpass filter $h_{pr}(N,n)$ is usually symmetric and has to satisfy $h_{pr}^2(N,n) + h_{pr}^2(N, N-1-n) = 2$. This last condition, imposed on the window, ensures that the resulting MLT is orthogonal. The two symmetric halfs of the window are called "tails". We want to show how to construct adaptive MLT's. Note that something similar, without details of the construction or the proof, was mentioned in [1]. The basic idea is that since the tails of the larger, N_2-channel MLT are already orthogonal to each other, why not use them to construct the filters in the smaller, N_1-channel MLT. Thus, we choose N_1 filters out of N_2 possible ones, discard the middle coefficients and just keep the tails. Then by construction these tails are going to be orthogonal to each other, as well as across overlaps.

We now briefly present the construction (for more details, refer to [2]). We want to show that one can switch from an N_1-channel MLT to an N_2-channel MLT (where $N_1 < N_2$ and $N_1 = 2^{n_1}, N_2 = 2^{n_2}$) as follows:

(a) Since the maximum overlap between the two MLT's is N_1, then one has to adjust the size of the window of the N_2-channel MLT accordingly, that is, its length has to be reduced to $N_2 + N_1$, or the outer $N_1/2$ coefficients of each window tails have to be zero, while the middle N_1 ones have to equal $\sqrt{2}$.

(b) Then, one has to choose N_1 basis functions out of N_2 possible ones. It can be shown that the choice of indices of the form $i \cdot 2^{n_2-n_1+i}$ or $i \cdot (2^{n_2-n_1+i} - 1)$ will yield the appropriate basis functions leading to an orthogonal N_1-channel MLT.

(c) Finally, the resulting basis functions of the N_1-channel MLT are $h_k(N_1,n) = \sqrt{N_2/N_1} h_{i_k}(N_2, n + (N_2 - N_1)/2)$, for $n \in [0, N_1 - 1]$, and $h_k(N_1, n) = \sqrt{N_2/N_1} h_{i_k}(N_2, n + 3(N_2 - N_1)/2)$, for $n \in [N_1, 2N_1 - 1]$, where i_k belongs to the set of valid indices.

The proof is very simple. Here we just sketch it, for more details, refer to [2]. One has to prove that the orthogonality of overlapping tails between the two banks holds (true by construction). By the same token, the orthogonality of tails for the N_1-channel MLT holds. The two facts left to show are that the resulting vectors from the N_1-channel MLT are unitary, as well as that they are mutually orthogonal. After some algebraic manipulations, these follow easily.

In [2], a few more interesting tilings are shown, demonstrating how to switch among more than two MLT's. It is also discussed how to obtain boundary MLT's, that is, MLT's that are switched without overlaps.

Since the MLT can be seen as the dual of the wavelet packet tree [6, 1], MLT can be pruned optimally using the single tree algorithm. The fast dynamic programming based pruning algorithm used to find the best basis in the wavelet packet application, is also applicable here due to the independence in the optimal orthonormal decomposition splits for adjacent signal segments of the MLT tree. Note that this is true only when the tails of all the MLT's involved are the same, otherwise changing MLT's become a "dependent" problem. For more details on these issues, refer to [2].

3 SWITCHING ORTHOGONAL TREES

We now consider the problem of changing between orthogonal trees based on two-channel filter banks. If we can do this we will be able to construct the most arbitrary tree-based tilings as indicated in Figure 1 (d). We will make extensive use of the time-domain operator notation for filter banks [5, 8]. For example, if we wish to grow some section of a subband tree, but only for a finite portion of the signal, we can apply a finite duration orthogonal filter bank for this segment. That is we apply the operator

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad (2)$$

to the outputs of the branch that we wish to grow. $\mathbf{M}$ is the matrix that represents the finite duration orthogonal filter bank, so that we get the additional frequency resolution for the duration in question, but before and afterwards we process with the identity operator, *i.e.* leave the branch alone. In this way we can change between arbitrary filter bank topologies (and corresponding arbitrary tilings of the time-frequency plane) by taking appropriate cascades of orthogonal filter bank operators.

3.1 Orthogonalization procedure

The problem of applying an orthogonal filter bank over a finite signal segment involves finding an appropriate way of treating the boundaries. If we take, for example, the case of length-4 filters, applied for the segment $0 \le n \le n_1$, consider the following truncation of the time-domain operator

$$\mathbf{M} = \begin{bmatrix} h_0(1) & h_0(2) & h_0(3) & 0 & 0 & \cdots \\ h_0(2) & -h_0(1) & h_0(0) & 0 & 0 & \cdots \\ 0 & h_0(0) & h_0(1) & h_0(2) & h_0(3) & \cdots \\ 0 & -h_0(3) & h_0(2) & -h_0(1) & h_0(0) & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \quad (3)$$

We have shown the top left corner only, but the bottom right is entirely similar. It is easy to verify that this matrix is square, has full rank, but is no longer unitary. If we denote by M_i the i-th row of $\mathbf{M}$ we find that

$$< M_i, M_j > = 0, \quad i \in \{0,1\}, j \in \{2,3, \cdots n_1 - 1, n_1\}; \quad (4)$$

but $< M_0, M_1 > \neq 0$ and $< M_{n_1-1}, M_{n_1} > \neq 0$. Since the matrix is of full rank, *i.e.* we have a set of linearly independent vectors, we can restore orthogonality using the Gram-Schmidt procedure. To do this start by normalizing the first vector $M_0^{''} = M_0/||M_0||$, and then

$$\begin{aligned} M_1^{'} &= M_1 - < M_1, M_0^{''} > M_0^{''} - \sum_{j=3}^{n_1} < M_1, M_j > M_j \\ &= M_1 - < M_1, M_0^{''} > M_0^{''}. \end{aligned} \quad (5)$$

The simplification is a consequence of (4). Finally set $M_1^{''} = M_1^{'}/||M_1^{'}||$. Note that since M_1 and $M_0^{''}$ each have only three non-zero entries, so does $M_1^{''}$ from (5). The same procedure is applied to the other boundary vectors M_{n_1-1} and M_{n_1}. A new matrix $\mathbf{M}^{''}$ which has rows

$$\{M_0^{''}, M_1^{''}, M_2, M_3, \cdots, M_{n_1-2}, M_{n_1-1}^{''}, M_{n_1}^{''}\}$$

is then obviously unitary. What is important to note is that $\mathbf{M}^{''}$ has exactly the same zero entries as $\mathbf{M}$; *i.e.* the orthogonal boundary filters have the same support as the truncated filters.

This particularly simple example illustrates a much stronger result, which gives that the boundary filters for orthogonal FIR filter banks always have support only in the region of the boundary. We can formally state this as follows.

Proposition 3.1 *The set of boundary filters needed to apply a two-channel orthogonal filter bank, with length-N filters to a finite length signal is a set of $(N-2)/2+d$ vectors at each boundary, each of which has only $N-2+d$ non-zero values. If we define*

$$\mathbf{Q} = [\ \mathbf{0}_{d_l} \quad \mathbf{G} \quad \mathbf{0}_{d_r}\ ], \tag{6}$$

where $\mathbf{G}$ *is the* $2k \times N_0 + 2(k-1)$ *matrix containing the shifted filter impulse responses of the filters* $h_0(n)$ *and* $h_1(n)$, *and* $\mathbf{O}_{d_l}$ *and* $\mathbf{O}_{d_r}$ *are* $2k \times d_l$ *and* $2k \times d_r$ *matrices of zeros, then the boundary vectors are always of the form*

$$\mathbf{e}_i' = (\mathbf{I} - \mathbf{Q}^T\mathbf{Q}) \cdot \mathbf{e}_i,$$

for some $\mathbf{e}_i$.

The outline of the proof is given in [3]. The importance of the result is that it is constructive; for any orthogonal filter bank it tells us how to take care of the boundary. Essentially this proposition merely involves carrying out the Gram-Schmidt procedure in operator notation; the novel factor is that the resulting output vectors have non-zero elements only in the region of the transition. The result of the orthogonalization is, of course, not unique. However, given one solution we can explore the space of all possible orthogonal boundary solutions by premultiplying by the matrix

$$\begin{bmatrix} \mathbf{U}_l & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_1-(N-2)-d_l-d_r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U}_r \end{bmatrix}, \tag{7}$$

where $\mathbf{U}_l$ and $\mathbf{U}_r$ are unitary matrices of size $(N-2)/2+d_l$ and $(N-2)/2+d_r$.

The above orthogonalization procedure is actually a special case. Clearly we can use it to change between orthogonal trees just by calculating an appropriate set of boundary functions on each side of the transition; Proposition 3 tells us how to do so. In this case there will be no overlap across the boundary; more general solutions, where there is in fact overlap, are given in [3]. Further, it is possible to use the discrete-time time-varying bases described above in an iterative scheme to derive continuous-time time-varying bases [2, 3].

4 OPTIMAL TILING

We have so far considered the construction of time-varying bases. The obvious question is how to find the "best basis" [1] for a given input signal. We now describe the machinery needed to *jointly* find the best adaptive time split together with the best basis for each time segment for a given signal. This will be done using an extension of the wavelet packet algorithm of [1, 6] called the *double tree algorithm* which uses the orthogonal boundary filters designed in the previous section. Before we describe the double tree algorithm, we first lay down the necessary background groundwork.

4.1 Rate–Distortion framework

While the double tree algorithm is an extension of the best basis algorithm of [1] and is valid *for any additive cost measure* over the set of sequences considered, we turn here to the rate-distortion (R-D) cost measure, i.e. we seek the optimal binary tiling of the time–frequency plane in the operational R-D sense. The R-D measure is a two-sided cost function encompassing both rate and distortion, having the distortion-only and entropy-only criteria of [1] as special cases, and it is the correct measure for coding applications. Note that 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. For an orthonormal 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 [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 [6], 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 [6] for details.

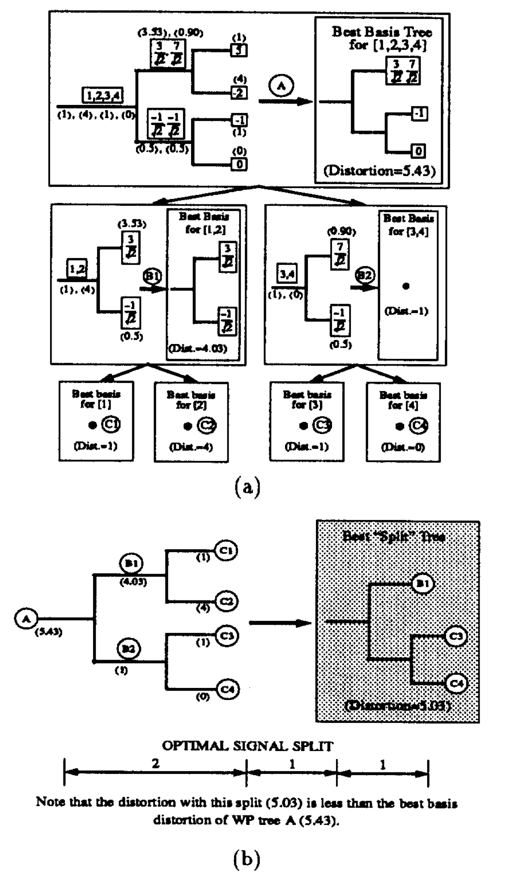
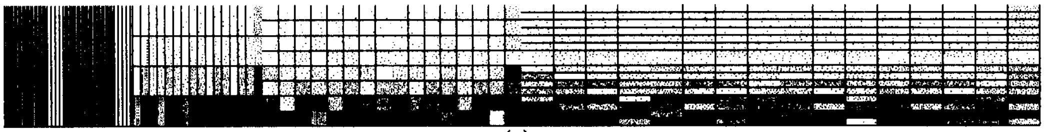


Figure 2: The double tree optimal ON splitting algorithm on the input signal [1,2,3,4] in R^4 for the Haar kernel and a scalar quantizer of step size 4. Lagrangian costs are shown in brackets ($\lambda = 0$ used here). (a) The best basis WP subtrees corresponding to all feasible signal subsets. (b) The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


(a)

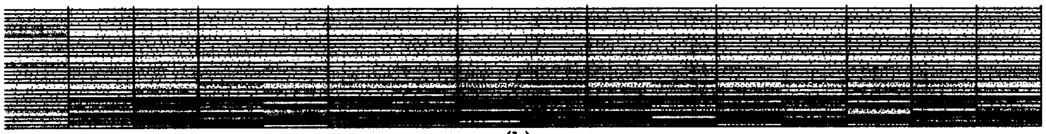

(b)

Figure 3: Optimal tiling for a synthetic test signal consisting of a Dirac and a sinusoid: Dirac is close to the left edge. (a) Optimal split using the double tree algorithm. Note how both the Dirac in time and Dirac in frequency are detected with good resolution. (b) Optimal single tree (WP) split. Note how the impulse in time is not detected. (Shown for the Daubechies D2 filter, quantizer step size of 0.1 (fine quantization), $\lambda = 1$, and the tree grown to a maximum depth of 6, i.e. to a leaf size of 16. First order entropy and MSE are used as the rate and distortion measures.)

4.2 Brief Review of Wavelet Packet Algorithm

Optimal bit allocation in a wavelet packet (WP) framework has been solved in [6] 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).

4.3 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 2. 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 2(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 2(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 [6].

5 EXPERIMENTAL RESULTS

The double tree algorithm was used to find the optimal tiling for several test signals, both real and synthetic. For a test signal composed of of an impulse and a sinusoid, Figure 3(a) and (b) show the optimal tiling representations for the double tree and the "best basis" algorithms respectively. Darker tiles denote basis functions with more energy. A scalar quantizer of step size 0.1 is used (to ensure fine quantization) with the tree grown to a smallest split of 16, for $\lambda = 1$. As seen, the double tree split "adapts" well to the input signal, finding a split with good time resolution around the impulse, and good frequency resolution around the sinusoid. The wavelet packet tiling of Figure 3(c), on the other hand, fails to isolate the impulse. This example should highlight the usefulness of adaptive time–frequency tiling, as possible using the double tree algorithm.

REFERENCES

[1] R. Coifman, Y. Meyer, and V. Wickerhauser. Wavelet analysis and signal processing. In M. B. Ruskai *et al.*, editor, *Wavelets and their Applications*, pages 153–178, Boston, MA, 1992. Jones and Bartlett.

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

[3] C. Herley and M. Vetterli. Orthogonal time-varying filter banks and wavelets. In *Proc. IEEE Int. Symp. Circ. and Systems*, Chicago, IL, May 1993. To appear.

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

[5] K. Nayebi, T. P. Barnwell III, and M. J. T. Smith. Time domain filter bank analysis: A new design theory. *IEEE Trans. on Signal Proc.*, IV(6):1414–1429, June 1992.

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

[7] O. Rioul and M. Vetterli. Wavelets and signal processing. *IEEE Signal Proc. Mag.*, pages 14–38, Oct. 1991.

[8] M. Vetterli and C. Herley. Wavelets and filter banks: theory and design. *IEEE Trans. on Signal Proc.*, 40(9):2207–2232, Sept. 1992.