BEST WAVELET PACKET BASES USING RATE-DISTORTION CRITERIA

Kannan Ramchandran *Martin Vetterli*

Department of Electrical Engineering
and Center for Telecommunications Research,
Columbia University,
New York, N.Y. 10026

ABSTRACT

The use of an adaptive tree structure using wavelet packets as a generalized wavelet decomposition for signal compression was recently introduced by Coifman, Meyer, Quake, and Wickerhauser [1]. The idea is to decompose a discrete signal into all wavelet packet bases of a given wavelet kernel, and then to find the "best basis" wavelet packet. Unlike the work of [1], in this paper, we conduct our best basis search in a rate–distortion framework. We formulate a fast algorithm to prune the full tree, signifying the entire library of admissible wavelet packet bases, into that basis subtree which minimizes the total distortion for a given coding bit budget. Arbitrary finite quantizer sets are assumed for each hierarchical level of the basis-family tree. Finally, a DCT "wavelet packet" basis quadtree segmentation is described as an image coding application in a JPEG [5] environment, with good improvement shown over non-adaptive JPEG quantization.

1. INTRODUCTION

Wavelet packets, introduced recently by Coifman, Meyer, Quake, and Wickerhauser (CMQW) [1] as a family of orthonormal (ON) expansions, include the well-known wavelet basis and the Short-Time-Fourier-Transform-like (STFT) basis as its members. They represent the entire family of subband coding tree decompositions, from which the optimal decomposition subtree can be selected to maximize compression by permitting the signal characteristics to be matched "on the fly."

This scheme enables the coder to exhibit, for example, a STFT-like characteristic (regular tree) at one source instance, a wavelet characteristic (logarithmic tree) at another instance, or any intermediate characteristic (arbitrary wavelet packet subtree) at yet other instances to best match the signal's non-stationary statistics. Figure 1 shows the complete depth-(log N) binary subband or wavelet packet (WP) tree, while Figure 2 shows some typical admissible "pruned" WP trees or subband topologies. The popular wavelet and STFT decompositions are mere special cases of permissible WP structures. The ON decomposition, which enables each internal tree node to spawn off branches providing a complete disjoint basis cover for the space spanned by their parent, is vital to the development of the fast pruning algorithm.

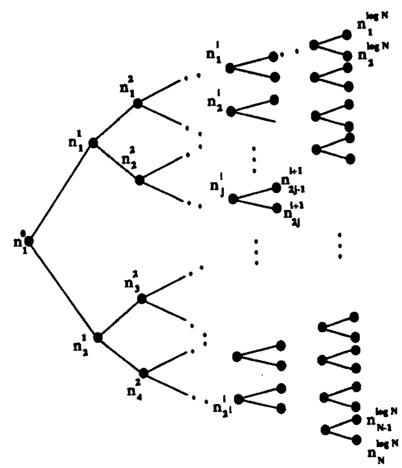

Figure 1: Complete wavelet packet tree of depth log N to code signal block of dimension N. Each node n_j^i contains the basis vector b_j^i with wavelet packet coefficient vector c_j^i. The complete set of all pruned subtrees represents the library of all admissible wavelet packet bases, or equivalently, all subband decomposition topologies.

Related work and contribution of this paper

While the adaptivity and the speed of the best-basis search of [1] are unmistakable, the cost criterion and the coding (quantization) method used there to exploit this speed and flexibility are somewhat ad hoc. In this paper we formulate a fast algorithm, for a given total coding bitrate budget, to pick the optimal WP basis, together with the optimal quantizer choice for that optimal WP subtree, for each of the independent segments or "blocks" that the signal comprises. Optimality is with respect to a *global* distortion criterion that is additive over the signal blocks, e.g. m.s.e or weighted m.s.e. We conduct our best basis hunt in a rate-distortion (R-D) framework, a generalization of the treatment in [1] where a one-sided "entropy" or "m.s.e" criterion is used. Our approach could be viewed, ***in its quadtree application***, as an extension of [3] to provide a fast algorithm covering ***hierarchies*** of admissible quantizers. As a practical

0-7803-0593-0/92 $3.00 1992 IEEE

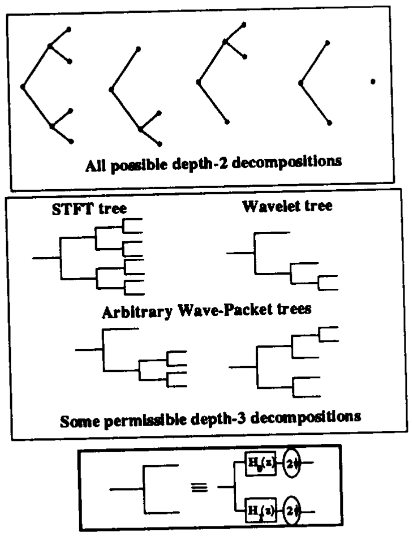

Figure 2: (a) All possible binary wavelet packet decompositions of depth 2. (b) Some typical depth-3 binary wavelet packet subtree decompositions. Note that H_0 and H_1 represent the "low pass" and "high pass" analysis filters.

contribution, a quadtree-based image compression application using a family of DCT "wavelet packet" bases is described. Our application is similar to that of independently done work in [4], involving efficient quadtree segmentation using VQ. We, however, use classified quantizers in a JPEG (DCT-based) [5] coding environment.

Note that the DCT-basis family tree of our application and the standard basis tree of [4] are not strictly WP trees, which are derived recursively using QMF filter banks or using multi-resolution wavelet analysis. The scope of applicability of our algorithm extends to all classes of structures which permit the construction of a hierarchy of basis covers for the input signal space. While this obviously includes structures like quadtrees and ON transformed (e.g. DCT) quadtrees, other powerful structures such as the CMQW multiresolution decomposition wavelet packets, and hierarchical subband coders are also applicable.

2. BASIC IDEA OF THE ALGORITHM

We convert our constrained best basis WP (and quantizer) sequence search into an unconstrained one via the Lagrangian cost functional $J(\lambda) = D + \lambda R$. See [2] for mathematical details. This makes it feasible to deal with each block independently. It can be shown that, at optimality, all nodes of all subtrees of all blocks must operate at "constant quality slope λ". See Figure 3. For a given $\lambda = | \Delta D / \Delta R |$, we populate each node of each tree block *independently* with the Lagrangian cost function associated with the best quantizer for that node. The best quantizer for a particular tree node is that one which "lives" at absolute slope λ on the convex hull of the operational R–D curve for that node, as shown in Figure 3. Then, by applying, *in parallel for each signal block*, the pruning criterion of

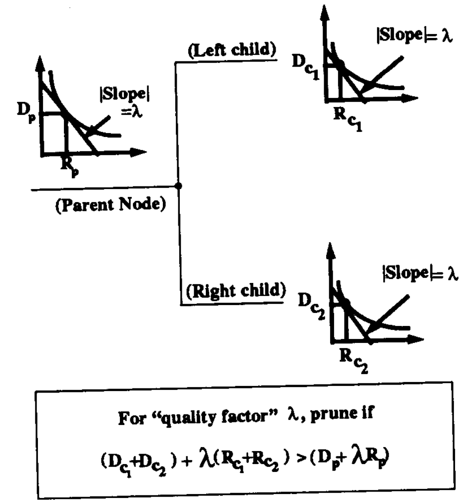

Figure 3: Lagrangian cost pruning criterion for "quality criterion" λ for each parent node of the wavelet packet tree. This condition is used recursively to do fast pruning from the complete tree depth towards the root to find the optimal subtree for a given λ.

Figure 3 recursively on every node, starting from the full-depth tree and proceeding towards the root, we find the sequence of best WP bases and associated best quantizers with which to code the signal. The recursive algorithm exploits Bellman's optimality principle by eliminating quickly a host of suboptimal subtrees from contention for the optimal solution, in a manner reminiscent of the popular Viterbi algorithm. Finally, we show a fast convex iterative search, using Newton's method, over the Lagrangian multiplier λ, to find the optimal λ^* that satisfies the given budget constraint R_{budget}.

3. FORMAL PROBLEM DEFINITION

Without loss of generality, we will consider the problem of a binary WP decomposition tree of a discrete input signal (vector) of size N in $l^2(N)$ See Figure 1. The analysis and synthesis filters of each branch satisfy the standard orthonormality conditions of paraunitary perfect reconstruction filter banks (PRFB's). As is well known, iterating the orthonormal filter templates to the complete tree depth results in an equivalent generalized multiresolution decomposition tree (i.e. WP tree) whose nodes represent a family of orthonormal bases [1]. We assume that there are M signal blocks to be coded independently, each of size N. Let T denote the complete WP tree, for each signal block, of depth $log\ N$ as shown in Fig. 1, t denote any node of T. Let $S \preceq T$ be any pruned subtree of T, i.e. a WP basis subtree of T that shares its root; thus, S corresponds to any admissible WP basis, while $\tilde{S}$ defines the set of terminal nodes or leaves of S. Also, let $\mathbf{q}_\mathbf{a}(t)$ be the set of all admissible quantizers for node $t \in T$, while $\mathbf{Qa}(S)$ represents the vector set of all admissible quantizers for the collection of individual leaf nodes of subtree S. Figure 1 shows the definitions of n_j^i, b_j^i, c_j^i as the jth node, basis, and coefficient vector respectively, at the ith tree-depth or "scale" (for $i = 1, 2, \ldots, \log N$). Thus, b_j^i represents the $\mathbf{R}^{\mathbf{N}}$-basis members associated with node n_j^i, while c_j^i represents the inner product of the signal with the basis vectors in b_j^i. Note also that to simplify notation, $\{t, b_t$ and $c_t\}$ will be invoked where convenient.

Let $D_q(t), R_q(t)$ be the distortion and bitrate, respectively, associated with quantizing WP coefficient vector c_t of node t using quantizer $q \in \mathbf{qa}(t)$, and $D_Q(S), R_Q(S)$ be the distortion and rate, respectively, associated with coding subtree S using quantizer $Q \in \mathbf{Qa}(S)$. In our case, they are both linear tree functionals: i.e. total distortion $= D_Q(S) = \sum_{t \in \tilde{S}} D_q(t)$ and total rate $= R_Q(S) = \sum_{t \in \tilde{S}} R_q(t)$.

The problem to solve, then, is that of finding, given a total budget of R_{budget} to code M independent signal blocks, that sequence of (pruned) subtree best-bases $S_i^* \preceq T$ (for $i = 1, 2, \ldots, M$) together with their associated optimal quantizers $Q_i^* \in \mathbf{Qa}(S_i^*)$ which minimize the global coding distortion. Mathematically, this boils down to determining: $D_{min} = \sum_{i=1}^{M} D_{Q_i^*}(S_i^*)$, where

$$D_{Q_i^*}(S_i^*) = \min_{S_i \preceq T} \left[\min_{Q_i \in \mathbf{Qa}(S_i)} D_{Q_i}(S_i) \right] \qquad (1)$$

such that $R_{total} = \sum_{i=1}^{M} R_{Q_i^*}(S_i^*) \leq R_{budget}$.

4. FAST SOLUTION

Without loss of generality, due to the "parallelization" of the problem, we consider the problem of pruning a single tree (M=1), i.e. determining:

$$D_{Q^*}(S^*) = \min_S \min_Q D_Q(S) \text{ s.t. } R_{Q^*}(S^*) \leq R_{budget} \quad (2)$$

We solve the constrained problem of Eq.(2) by converting it to an unconstrained problem using Lagrange multipliers. Our problem becomes a hierarchical extension, using a fast pruning algorithm, of the "flattened" problem of [3]. After introducing the Lagrangian cost functional $J(\lambda) = J_\lambda(S, Q) = [D_Q(S) + \lambda R_Q(S)]$, the equivalent unconstrained problem, becomes the solution to:

$$J^*(\lambda) = \min_S \min_Q J_\lambda(S, Q) \qquad (3)$$

$$= \min_S \Big(\sum_{t \in \tilde{S}} \min_q [D_q(t) + \lambda R_q(t)]\Big) \qquad (4)$$

Thus, one can equivalently solve the above unconstrained equation for the optimal values of Q, S, and λ. The optimal quantization choice Q (for a fixed subtree S and a fixed operating slope λ) is the inner minimization of Eq.(4), expressed as the sum of the (Lagrangian) costs of the leaf nodes of S. The optimal subtree basis $S \preceq T$ (for the fixed quality slope λ) of the outer minimization of Eq. (4) is found by the fast pruning method described in Section 4.1. Finally, the optimal λ that solves the given budget R_{budget}, to within a convex hull approximation, is found using a convex search method described in Phase II of Section 5.. Refer to [2] for mathematical details.

4.1. Fast pruning to find optimal basis subtree

A Viterbi-like fast dynamic programming technique is feasible due to the ON property of the WP basis family, that enables the signal space spanned by an arbitrary subtree rooted at internal node t of the tree to be identical to the space spanned by the twin subtrees rooted at the two branches emanating from node t. To be specific, let $t = n_j^i$, i.e. t is the jth node of the ith hierarchical level (or depth) of the tree T. Its two children are $t_1 = n_{2j-1}^{i+1}$ and $t_2 = n_{2j}^{i+1}$. See Figure 1. Then, because of the ON property, the subtrees rooted at t_1 and t_2 cover disjoint halves of the signal space spanned by their parent node t.

This allows a direct *quantitative* one-to-one comparison between the $N/2^i$ basis coefficients $\{c_j^i\}$ associated with the basis subset $\{b_j^i\}$ of node t with the $(2 \times (N/(2^{(i+1)})))$ coefficients $\{\{c_{2j-1}^{i+1}\}, \{c_{2j}^{i+1}\}\}$ associated with the basis subsets $\{b_{2j-1}^{i+1}\}$ and $\{b_{2j}^{i+1}\}$ of nodes t_1 and t_2 respectively. The "split/merge" decision will be based on which option leads to a cheaper Lagrangian cost, as spelled out in Figure 3.

Assume known the optimal subtree from node $t = n_j^i$ "onwards" to the full tree-depth $\log N$. The subtrees could be likened to surviving paths in the Viterbi algorithm. Then, by Bellman's optimality principle, we know that all surviving paths passing through node $t = n_j^i$ at depth i must invoke this same optimal "finishing" path. There are two contenders for the "surviving path" at every node of the tree, the parent and its children, with the winner having the lower Lagrangian cost. Using this, we begin at the complete tree-depth $\log N$ and work our way towards the root of the tree, using the above cost criterion at each level i to determine whether to split or merge.

5. COMPLETE ALGORITHM

5.1. Initialization

Step 1: Generate the coefficients $\{c_j^i\}$ for the entire WP family.

Step 2: Gather the given quantizer set dependent R-D points for each node $t \in T$ $\forall q \in \mathbf{qa}(t)$.

Phase I: Optimality For A Given Operating Slope

Phase I of the algorithm is run for a given slope value λ:

Step 3: For the λ of the current iteration, populate all the nodes $t \in T$ with $J_t(\lambda)$, (or equivalently, populate node n_j^i with $J_j^i(\lambda)$), where: $J_t(\lambda) = \min[D_{q_t}(t) + \lambda R_{q_t}(t)]$

Step 4: Initialize $i \leftarrow n$, where $n = \log N$ is the maximum signal block tree-depth. For $t = n_j^n$, if q_t^* is the value of q_t that minimizes $J_t(\lambda)$ initialize:

$\tilde{R}_j^n \leftarrow R_j^n$ (where $R_j^n = R_{q_t^*}(t)$)

$\tilde{D}_j^n \leftarrow D_j^n$ (where $D_j^n = D_{q_t^*}(t)$))

$\tilde{J}_j^n \leftarrow J_j^n$

Step 5: $i \leftarrow i - 1$. If $i < 0$, go to Step 8.

Step 6: $\forall j = 1, 2, \ldots, 2^i$ at the ith tree level:

if $J_j^i(\lambda) < \tilde{J}_{2j-1}^{i+1}(\lambda) + \tilde{J}_{2j}^{i+1}(\lambda)$,

then $\{split(n_j^i) \leftarrow NO; \tilde{R}_j^i = R_j^i;\ \tilde{D}_j^i = D_j^i;\ \tilde{J}_j^i = J_j^i\}$
else $\{\ split(n_j^i) \leftarrow YES;\ \tilde{R}_j^i = R_{2j-1}^{i+1} + R_{2j}^{i+1};$
$\tilde{D}_j^i = D_{2j-1}^{i+1} + D_{2j}^{i+1}; \tilde{J}_j^i = J_{2j-1}^{i+1} + J_{2j}^{i+1}\}$

Step 7: Go to Step 5.

Step 8: Starting from the root t_0, and using, in a linked-list fashion, the node data-structure element *split*(node), selected optimally for all the nodes of T, carve out the optimal subtree $S^*(\lambda)$ and its associated optimal quantizer choice $Q^*(\lambda)$. Also readily available at the data-structure for root node t_0 are $R_{Q^*}(S^*) = \tilde{R}_1^0$ and $D_{Q^*}(S^*) = \tilde{D}_1^0$, the rate and distortion of the optimal subtree $S^*(\lambda)$.

Phase II: Iterating towards the optimal operating point

We now describe the bisection iteration process (convex search) to find the optimal slope λ^*.

Step 1: Pick $\lambda_l \le \lambda_u$ such that

$$\sum_i R_i^*(\lambda_u) \le R_{budget} \le \sum_i R_i^*(\lambda_l)$$

If the inequality above is an equality for either slope value, stop. We have an exact solution. Otherwise, proceed to Step 2.

Step 2: $\lambda_{next} \leftarrow \left| \frac{\sum_i [D_i^*(\lambda_l) - D_i^*(\lambda_u)]}{\sum_i [R_i^*(\lambda_l) - R_i^*(\lambda_u)]} \right| + \epsilon$, where ϵ is a vanishingly small positive number picked to ensure that the lower rate point is picked if λ_{next} is a "singular" slope value.

Step 3: Run the Phase I optimal algorithm for λ_{next}.
$\Longrightarrow$ if $\{\sum_i R_i^*(\lambda_{next}) = \sum_i R_i^*(\lambda_u)\}$, then stop. $\lambda^* = \lambda_u$.
$\Longrightarrow$ else if $(\sum_i R_i^*(\lambda_{next}) > R_{budget})$, $\lambda_l \leftarrow \lambda_{next}$. Go to Step 2.
$\Longrightarrow$ else $\lambda_u \leftarrow \lambda_{next}$. Go to Step 2.

6. QUADTREE APPLICATION RESULTS

As an application, Figure 4 shows a typical result of a quadtree-based image compression example using a family of DCT "wavelet packet" basis matrices of sizes 4x4, 8x8, and 16x16 used independently over the non-overlapping 16x16 subblocks into which the original image is divided. Classified quantizers, with four perceptually consistent classes for each block-size, were chosen to be the admissible quantization set for this application. Figures 4 and 5 show some typical results. As can be seen, for the "Barbara" image, our adaptive scheme outperforms the static JPEG scheme by about 2-3 dB at fixed bitrate, or equivalently about 25-35% compression advantage at fixed SNR, over an entire range of bit rates of interest.

7. CONCLUSION

We have shown, for a given hierarchy of admissible quantizers, an efficient scheme for coding adaptive trees whose individual nodes spawn off descendents forming a disjoint and complete basis cover for the space spanned by their parent nodes. The scheme presented guarantees operation on the convex hull of the operational R-D curve for the admissible hierarchy of quantizers. Applications for this coding technique include the CMQW generalized multiresolution wavelet packet decomposition, iterative subband coders, and quadtree structures. An application to image processing involving quadtrees with a family of DCT bases has been demonstrated in a JPEG-like coding environment with good improvement shown over the static JPEG coding scheme.

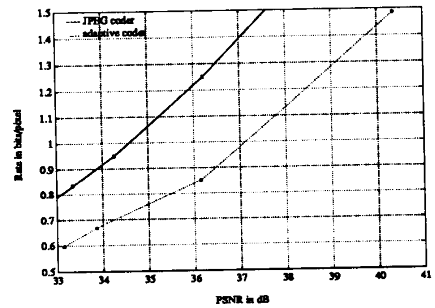

Figure 4: Comparison of adaptive depth-3 DCT basis quadtree coding scheme with non-adaptive JPEG coding scheme for Barbara image.

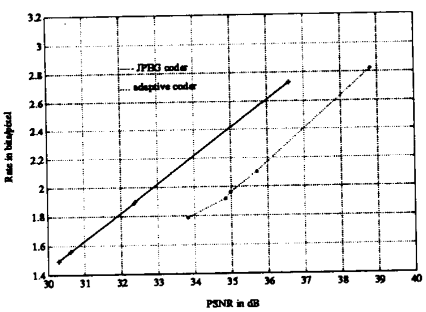

Figure 5: Comparison of adaptive depth-3 block DCT basis quadtree coding scheme with non-adaptive JPEG coding scheme for "mit" sequence frame.

REFERENCES

[1] R. Coifman, Y. Meyer, D. Quake, and V. Wickerhauser, Numerical Algorithms Research Group, Dept. of Mathematics, Yale University, "Signal compression with wave packets," Wavelet Workshop, Marseille, Sept. 1990.

[2] K. Ramchandran and M. Vetterli, "Best Wavelet Packet Bases in a Rate-Distortion Sense," Submitted to IEEE Trans. Image Proc., 1992.

[3] Y. Shoham and A. Gersho, "Efficient Bit Allocation for an Arbitrary Set of Quantizers," IEEE Trans. ASSP, vol. 36, no. 9, pp. 1445-1453, Sept. 1988.

[4] G. J. Sullivan and R. L. Baker, "Efficient Quadtree Coding of Images and Video," Proceedings ICASSP '91, vol. 4, M7.11, pp. 2661-2664, May 1991.

[5] Joint Photographic Experts Group, "JPEG Technical Specification: Revision 8 (DRAFT)," ISO/IEC JTC1/SC2/WG8, CCITT SGVIII, Aug. 1990.