AN EFFICIENT TOP-DOWN APPROACH FOR THE DESIGN OF TREE-STRUCTURED ORTHONORMAL FILTER BANKS*

Rajeev Gandhi and Sanjit K. Mitra

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
rajeev@iplab.ece.ucsb.edu, mitra@ece.ucsb.edu

ABSTRACT

Wavelet and tree-structured filter bank based coding schemes find applications in a number of image coding algorithms. The efficiency of such coding schemes can be enhanced further by adapting the parameters of the coder to the statistics of the signal being compressed. In this paper, we propose an algorithm to determine the coefficients of the filter bank matched to the signal statistics, at every node of a tree-structured filter bank. In the proposed algorithm, we visit each node in a top-down fashion and determine the bit-allocation and the coefficients of the filters used to carry out the decomposition. An estimate of the coding gain provided by the later stages is used to account for the interaction between the filter bank at the current node and the filter banks at the later nodes.

1. INTRODUCTION

Uniform and non-uniform tree-structured filter banks are used in a number of image and audio compression algorithms. The efficiency of the compression algorithm can be improved by adapting the parameters of the tree-structured filter bank to the statistics of the input signal. The problem of designing tree-structured filter banks involves – determining the non-uniform/uniform decomposition, the coefficients of the filter bank at each decomposition and the bit allocation amongst the various nodes. The objective of the design procedure is to chose these parameters such that the mean square error (MSE) between the input and the reconstructed signal is minimized subject to a constraint on the total number of bits used to code the signal. While there exist techniques [1], which determine the bit-allocation and the non-uniform decomposition to obtain the optimal rate-distortion (R/D) performance, the problem of incorporating [10], [11] the optimization of the filter bank coefficients into this R/D criterion seems to be of intractable complexity. To reduce the complexity of the design algorithm, typically the coefficients of the filter banks are obtained by maximizing the coding gain provided by the filter bank. Under a high bit-rate assumption, maximizing the coding gain of the filter bank is equivalent to minimizing the MSE between the input and the reconstructed output sequence. For the rest of the paper, we use the term "node" to refer to a particular subband and "splitting a node" to refer to a two-band decomposition of the subband. For simplicity, we assume that a two-channel filter bank is used to carry out the subband decomposition. The algorithm can be easily generalized when an M-channel filter bank is used instead. We also assume that orthonormal filter banks are used at every node so that minimizing the mean square error (MSE) between the input and the output is equivalent to minimizing the MSE between the original subband coefficients and their quantized counterparts.

Most of the previous work [2]-[4] in the design of tree-structured filter banks have focussed on determining the filter coefficients such that the coding gain provided by the filter bank at that particular node is maximized. The algorithm developed in [10] computes the decomposition and bit-allocation using the R/D optimal approach while the criterion of coding gain maximization is used to determine the filter banks at every node. The techniques for the filter bank design in [2]-[4], [10] suffer from the fact that no attempt is made to take into account the interaction between the filter bank at a parent node and the filter banks at its child nodes. An interaction exists between them, since the filter banks at the child nodes depend upon the psd of the input at these nodes, which in turn depends upon the filter bank used at the parent node.

In this paper, we develop an efficient algorithm to determine the coefficients of the filter banks that maximize the *overall* coding gain of the tree structure. A top-down design approach is proposed to determine the coefficients of the filter bank. However, unlike previous techniques, we use an estimate of the coding gain provided by the filter banks at the child nodes to account for the above mentioned interac-

*This work was supported by a University of California MICRO grant with matching supports from Lucent Technologies, National Semiconductors, Raytheon Missile Systems, Tektronix Corporation and Xerox Corporation.

0-7803-5700-0/99/$10.00©1999 IEEE

tion between various stages. For the sake of completeness, a computationally simple scheme to allocate bits amongst the various nodes and to decide on whether to split a node is also proposed. However, the algorithm to determine the filter bank coefficients can be used in conjunction with any other technique to determine the bit-allocation and the non-uniform decomposition. The main objective of our work is to systematically formulate the design of tree-structured filter banks that includes the interaction between the different stages of the tree structure.

2. DESIGN PROCEDURE

Consider first the problem of determining whether to split the node n_i or not. The approach proposed in [1] could be used to make the node-split decisions, however its complexity increases exponentially with the depth of the tree to be searched. A number of heuristic approaches have been proposed to remedy this problem. Coding gain of the filter bank has been proposed as a criterion [3], [4] to determine whether to split a node or not. If the coding gain of the filter bank at that node is greater than $1 + \epsilon$, for a pre-determined $\epsilon << 1$, then the node is split. The problem with this approach is that – such split decisions are local in nature and there exist situations where it might be useful to split a particular node even when the coding gain provided by the filter bank at that node is just one. It is possible that the later splits will result in an overall coding gain greater than one. In order to illustrate this, consider that the input to a node in the tree has a psd of the form shown in Figure 1. Assume that ideal brick-wall filters can be used to carry

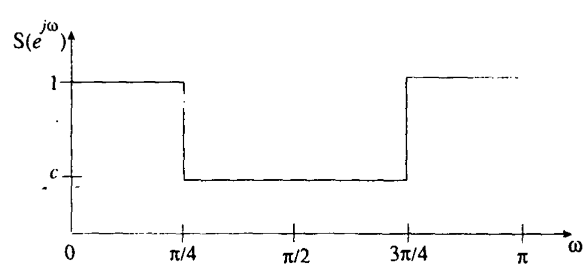

Figure 1: Power Spectral density of the input to one of the nodes in the tree

out the subband decomposition. The maximum coding gain provided by any two-channel orthonormal filter bank, for the input psd of Figure 1, is one. Thus in a conventional top-down approach the node will not be split any further. However, if were to split this node and then split each of the child nodes with ideal brick-wall filters, then an overall coding gain of $0.5(\sqrt{c} + \frac{1}{\sqrt{c}})$ is obtained. This example illustrates the "sub-optimal" nature of the conventional top-down techniques. The reason for this sub-optimality lies in the fact that the approach does not consider the interaction between the different stages of the tree. To overcome this difficulty, we make use of the "spectral flatness" [5] measure for the input psd at a given node. The value of the spectral flatness measure is given by

$$\gamma_x = exp(\int_{-\pi}^{\pi} log_e S_{xx}(e^{j\omega}) \frac{d\omega}{2\pi})/\sigma_x^2, \qquad (1)$$

In the proposed approach, if $1/\gamma_x$ is greater than $1 + \epsilon$, then the node is split. The rationale behind the use of spectral flatness measure is that if $1/\gamma_x$ is equal to one, then the spectrum is flat and there is no gain in splitting the node any further. However, if the value of $1/\gamma_x$ is greater than one, then we can potentially obtain higher coding gains by splitting the node even if the immediate split provides a coding gain of only one (it is possible that splitting one of the child nodes obtained by the split of the node could provide a large coding gain). Using the inverse of spectral flatness measure to decide whether to split a node or not, is equivalent to using the coding gain provided by an infinite channel filter bank at the node. If the coding gain provided by the infinte channel filter bank is not greater than one, then there is no gain in splitting the node any further. The use of spectral flatness measure to determine the node-split decisions assumes that there is no limit on the depth upto which a tree can be grown. However, if the maximum number of levels upto which a tree can be grown is constrained to L_{max}, then some modifications are required to determine the split decisions. Consider the case when the current node is at a depth of n levels from the root node. In this case, we compute the maximum coding gain provided by a $2^{L_{max}-n}$-channel filter bank at the current node. If the maximum coding provided by a $2^{L_{max}-n}$-channel orthonormal filter bank at this node is greater than $1 + \epsilon$, then we split the node. The approach outlined in [8] is used to determine the maximum coding gain that can be provided by any arbitrary channel orthonormal filter bank. As stated earlier the use of theoretical coding gain as a criterion to determine the node-split decision is biased, since splitting a node can only increase the coding gain of the overall tree. The coding gain of the tree can only increase by splitting a particular node. The use of coding gain to make the node-split decisions could be used as a technique to prune out the tree struture before applying the R/D approach to determine the node-split decisions.

After the decision to split n_i has been made, the next step is to determine the coefficients of the filter bank at that node and also the bit-allocation amongst the child nodes. Assume that the input psd at the node is $S_{x_i x_i}(e^{j\omega})$ and the number of available bits per sample are b_i. Let b_{ik} be the number of bits per sample allocated to the kth child node, n_{ik}, for

$k = 1, 2$ then the constraint on the bit allocation is

$$\frac{1}{2}\sum_{k=1}^{2} b_{ik} = b_i. \qquad (2)$$

Moreover, the variance of the signal at n_i is related to the variance of the signals at its child nodes, n_{ik}, through

$$\sigma_{x_i}^2 = \sum_{k=1}^{2} \sigma_{x_{ik}}^2. \qquad (3)$$

This is due to the fact that the analysis filters in an two-channel orthonormal filter bank satisfy the property

$$|H_{i1}(e^{j\omega})|^2 + |H_{i2}(e^{j\omega})|^2 = 1, \qquad (4)$$

where, $H_{ik}(e^{j\omega})$ represents the frequency response of the kth analysis filter. If the subband signals at the child nodes were being quantized then the noise variance at each node would have been given by

$$\sigma_{e_{ik}}^2 = c\, \sigma_{x_{ik}}^2 2^{-2b_{ik}}. \qquad (5)$$

The presense of a tree structure at one or more child nodes, say n_{im}, causes the variance of the quantization noise at n_{im} to be reduced by a factor equal to the coding gain provided by the tree structure at n_{im}. Thus, the total distortion due to quantization of subband signals at node n_i is

$$D_i = \sum_{k=1}^{2} \frac{c\, \sigma_{x_{ik}}^2 2^{-2b_{ik}}}{G_{ik}},$$

where G_{im} is the coding gain provided by the tree at node n_{im}. The coding gain provided by the tree structure at n_{im} cannot be determined until the filter bank and bit-allocation at the parent node n_i is determined. However, an estimate of the coding gain provided by the later stages can be used instead of its true value. The conventional top-down approaches [10], [2]-[4] can be thought of as special cases of this approach with G_{ik} set to unity. However, using G_{ik} equal to one may not necessarily be the best estimate of the coding gain provided by the later stages. In this algorithm, we estimate the value of G_{ik} at each node to be equal to the value of the asymptotic coding gain G_{ik}^{∞} (given by $1/\gamma_{x_{ik}}$) at that node. The justification for using G_{ik}^{∞} as an estimate of the coding gain provided by the tree at node n_{ik} is the fact that the tree starting from this node will be grown till the coding gain provided by the tree is closed to its asymptotic value. If there is a restriction on the number of levels upto which a tree can be grown, then we estimate G_{ik} to be the maximum coding gain provided by a $2^{L_{max}-n}$ channel filter bank at the node, where n is depth of the node from the root node. Let $\hat{G}_{ik}$ be the estimate of the coding gain provided by the tree at the node n_{ik} then D_i can be written as

$$D_i = \sum_{k=1}^{2} \frac{c\, \sigma_{x_{ik}}^2 2^{-2b_{ik}}}{\hat{G}_{ik}},$$

The bit allocation amongst the various child nodes of n_i is done such that D_i is minimized subject to the bit constraint given in Eq. (2). The optimal bit allocation is given by

$$b_{ik} = b_i + \frac{1}{2} log_2 \left(\frac{\sigma_{x_{ik}}^2 / \hat{G}_{ik}}{(\prod_{j=1}^{2} \sigma_{x_{ij}}^2 / \hat{G}_{ij})^{1/2}} \right). \qquad (6)$$

With the bit-allocation given in Eq.(6), the distortion at the node n_i is given by

$$D_i = 2c \left(\prod_{k=1}^{2} \sigma_{x_{ik}}^2 / \hat{G}_{ik} \right)^{1/2} 2^{-2b_i}. \qquad (7)$$

The coefficients of the filter bank at n_i are designed such that the value of D_i given in Eq. (7) is minimized subject to the constraint that the resulting filter bank is orthonormal. This can be done by using standard algorithms for the design of two-channel paraunitary [6] filter banks. Once the coefficients of the filter bank at n_i have been optimized, the psd of the signals at the child nodes can be determined. The number of bits allocated to each child node are given by the bit-allocation result given in Eq. (6). A similar procedure is recursively carried out at all the nodes in the tree.

It is also possible to use an iterative algorithm instead of using the asymptotic coding gain as an estimate of the coding gain provided by the rest of the tree at a node. In the iterative algorithm, we start with some initial filter coefficients at every node. This facilitates computation of the coding gain provided by the rest of the tree at every node (the value of $\hat{G}_{ik}$). However, after each iteration, the coefficients of the filter banks are updated and thus the estimates of G_{ik} used to compute the filter bank coefficients are not equal to the actual G_{ik} provided by the updated filter coefficients. In the next iteration, the updated filter banks are used to estimate the coding gain provided by the rest of the tree at every node. The iterative procedure is carried out till the estimate of the coding gain provided by the rest of the tree at every node is approximately equal to the actual coding gain, after the coefficients of the filter banks at all the nodes are updated (thereby implying that the true value of the coding gain provided by the rest of the tree at each node was approximately equal to the assumed value). While the iterative approach can provide a better estimate of $\hat{G}_{ik}$ than the value of the asymptotic coding gain, especially for filters of small lengths, the disadvantage of this iterative procedure is its enormous computational complexity.

3. ANALYTICAL EXPRESSIONS FOR THE OPTIMAL FILTER BANKS

In this section we determine the optimal filter bank at n_i that minimizes Eq. (7), under the assumption that there is no constraint on the length of the filters and asymptotic coding gain is used as an estimate of the coding gain provided by the rest of the tree at every node. We are interested in finding the optimal filter bank that minimizes the distortion given in Eq. (7). The expression for D_i can be simplified as

$$D_i = cM\ exp(\sum_{k=1}^{M} \int_{-\pi}^{\pi} log_e(S_{x_{ik}x_{ik}}(e^{j\omega}))\frac{d\omega}{2\pi})^{1/M} 2^{-2b_i},$$
$$= cM\ exp(\frac{1}{M}\int_{-\pi}^{\pi} log_e\left(\prod_{k=1}^{M} S_{x_{ik}x_{ik}}(e^{j\omega})\right)\frac{d\omega}{2\pi})2^{-2b_i}$$
$$= cM\ exp(\frac{1}{M}\int_{-\pi}^{\pi} log_e\left(\prod_{k=1}^{M} (\mathbf{B}(e^{j\omega}))_{kk}\right)\frac{d\omega}{2\pi})2^{-2b_i},$$

where $\mathbf{B}(e^{j\omega}) = \mathbf{E}(e^{j\omega})\mathbf{S}_{x_ix_i}(e^{j\omega})\mathbf{E}^*(e^{j\omega})$, $(\mathbf{A})_{kk}$ denotes the diagonal element of the matrix $\mathbf{A}$, $\mathbf{E}(e^{j\omega})$ denotes the polyphase matrix of the analysis filter bank, $*$ is used to indicate the transpose conjugate of a matrix and $\mathbf{S}_{x_ix_i}(e^{j\omega})$ denotes the psd matrix [7], [9] of the signal at the input to the analysis poplyphase matrix at n_i. Minimizing D_i in the above equation is equivalent to finding the polyphase matrix transfer function $\mathbf{E}(e^{j\omega})$ such that

$$C_i = \int_{-\pi}^{\pi} log_e(\prod_{k=1}^{M} (\mathbf{E}(e^{j\omega})\mathbf{S}_{x_ix_i}(e^{j\omega})\mathbf{E}^*(e^{j\omega}))_{kk})\frac{d\omega}{2\pi}. \quad (8)$$

is minimized (since exponential is a monotonocally increasing function). Now

$$C_i = \int_{-\pi}^{\pi} log_e(\prod_{k=1}^{M} (\mathbf{E}(e^{j\omega})\mathbf{S}_{x_ix_i}(e^{j\omega})\mathbf{E}^*(e^{j\omega}))_{kk})\frac{d\omega}{2\pi},$$
$$\geq \int_{-\pi}^{\pi} log_e(det\ (\mathbf{E}(e^{j\omega})\mathbf{S}_{x_ix_i}(e^{j\omega})\mathbf{E}^*(e^{j\omega})))\frac{d\omega}{2\pi},$$
$$= \int_{-\pi}^{\pi} log_e(det(\mathbf{S}_{x_ix_i}(e^{j\omega})))\frac{d\omega}{2\pi}, \quad (9)$$

where, we have made use of the inequality, $\prod_{i=1}^{M} \mathbf{A}_{ii} \geq det(\mathbf{A})$, for a positive definite matrix with equality only if $\mathbf{A}$ is a diagonal matrix. The positive definiteness of the matrix $\mathbf{E}(e^{j\omega})\mathbf{S}_{x_ix_i}(e^{j\omega})\mathbf{E}^*(e^{j\omega})$ follows from the fact that $\mathbf{S}_{x_ix_i}(e^{j\omega})$ is a spectral matrix [9] and is hence positive definite. Thus the quantity C_i in Eq. (8) is minimized when the inequality becomes an equality. This happens when $\mathbf{E}(e^{j\omega_0})$ is chosen to be the eigenvector of the psd matrix $\mathbf{S}_{x_ix_i}(e^{j\omega_0})$ for each value of $\omega_0 \in [0, 2\pi)$. Thus the optimal filter bank at n_i that minimizes the distortion, D_i in Eq. (7), is given by the eigenvalue decomposition of the psd matrix of the input at n_i.

4. RESULTS

The proposed design algorithm was used to design the tree-structured filter banks for an AR(1) process with $\rho = 0.95$ and an AR(2) process with poles at $0.98e^{\pm j9\pi/10}$. The value of the asymptotic coding gain for the AR(1) process was computed to be 10.25 and that for the AR(2) process was determined to be 134. The conventional top-down approach was also implemented, to compare the results with our algorithm. In the conventional top-down design, the product filter approach [10], [12] was used to maximize the coding gain of the filter bank at every node. The maximum depth of the tree-structure was chosen to be 5 and the value of ϵ was set to 0.01. In the proposed algorithm, the asymptotic coding gain was used as an estimate of the coding gain provided by the rest of the tree at every node. This ensured that the complexity of the proposed design procedure is similar to the complexity of the conventional top-down approach. One of the problems with our design technique is the fact that the resulting cost function is highly non-linear. A simple way to deal with non-linear cost function is to initialize the optimization algorithm at a good starting point. We used the output of the conventional top-down approach as the starting point for our algorithm. This results in an increase in the complexity of our algorithm by a factor of two. Figure 2 shows the coding gain of the overall tree-structured filter bank for the AR(1) process for filters of different lengths while Table 1 shows the results for the AR(2) process.

Length of Filter bank	Conventional Approach	Proposed Approach
4	44.5	48.3
8	85.5	98.2
12	101.06	114.7
16	105	116.3

Table 1: Comparison of the overall coding gain achieved by the conventional and proposed top-down approaches

5. CONCLUSIONS

In this paper, we developed an algorithm for the design of tree structured filter banks. Unlike existing techniques, we took into account the interaction between the filter banks at a given node and those at its child nodes. An estimate of the coding gain provided by the tree at a particular node was used to account for this interaction. The value of asymptotic coding gain at every node was used as an estimate of the coding gain provided by the later stages. Simulation results indicate improvements of about 0.2 – 0.6 dB over the conventional top-down techniques. The results could possibly be improved further by using more efficient algorithms to minimize non-linear cost functions.

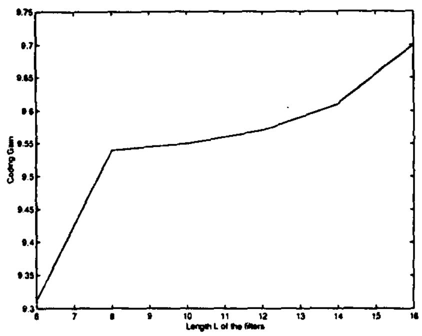

Figure 2: The coding gain of the tree-strucutred filter bank obtain by the proposed approach versus the lengths of the filters used to carry out the subband decomposition

6. REFERENCES

[1] K. Ramchandran and M. Vetterli, "Best wavelet packet basis in a rate-distortion sense," *IEEE Trans. Image Proc.*, pp. 160-175, April 1993.

[2] I. Balasingham, A. Fuldseth and T. A. Ramstad, "On optimal tiling of the spectrum in subband image compression," *Proc. Int. Conf. Image Process.*, vol. 1, pp. 49-52, 1998.

[3] M. V. Tazebay and A. N. Akansu, "A smart time-frequency exciser for spread spectrum communications," *Proc. Int. Conf. Acoust. Speech and Sig. Process.*, vol. 2, pp. 1209-1212, 1995.

[4] M. S. Chung, B. K. Rhim, J. H. Choi and H. S. Kwak, "Wavelet packet based on the top-down method," *Proc. IEEE Conf. Acoust. Speech and Sig. Process.*, vol. 4, pp. 3113-3116, 1997.

[5] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer, 1992.

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

[7] P. P. Vaidyanthan and A. Kirac, " Results on optimal biorthogonal filter banks," *IEEE Trans. Circuits and Sys. II: Analog and Digital Sig. Process.*, vol. 45, pp. 932-947, August 1998.

[8] P. P. Vaidyanthan, "Theory of Optimal Orthonormal Filter Banks," *Proc. IEEE Conf. Acoust. Speech and Sig. Process.*, 1996.

[9] K. C. Aas and C. T. Mullis, "Minimum mean-squared error transform coding and subband coding," *IEEE Trans. Info. Theory*, pp.1179-1192, July 1996.

[10] V. Pavlovic, P. Moulin and K. Ramchandran, "An Integrated Framework for Adaptive Subband Coding," *IEEE Trans. Signal Processing*, Vol. 47, pp. 1024-1038, April 1999.

[11] C. Taswell, "Satisficing search algorithms for selecting near-best bases in adaptive tree-structured wavelet transforms," *IEEE Trans. Signal Processing*, vol. 44, pp. 2423-2438, Oct. 1996.

[12] P. Moulin, M. Anitescu, K. O. Kortanek and F. Porta, "The role of linear semi-infinite programming in signal-adapted QMF bank design," *IEEE Trans. Signal Processing*, vol. 45, pp. 2160-2174, Sept. 1997.