

A Heuristic Algorithm to Find the Best Tree-Structured Image Decomposition

Santosh Kumar Tripathi
 Department of ECE
 Indian Institute of Science
 Bangalore 560 012
 INDIA
 San@protocol.ece.iisc.ernet.in

Dr. Anamitra Makur
 School of EEE
 Nanyang Technological University,
 SINGAPORE
 camakur@ntu.edu.sg

Abstract—In the context of uniform filter bank, width of the band depends upon the order of the subband decomposition. However for non-uniform filter banks there exist infinite possibilities even when the order is fixed. These infinite possibilities can be brought down to a certain number by using tree structured non-uniform decomposition of the spectrum, with a given pair of low pass and high pass filter. We address the problem of finding the decomposition that gives maximum coding gain. This question is extended to two dimensional signal like image.

In this work we suggest a heuristic algorithm that avoids the exhaustive search of max coding gain. We consider separable filters With the assumption of unit norm of each filter in filter pair.

1. INTRODUCTION

A broad class of digital signal compression techniques exploit the *linear dependency* in the input signals in order to achieve compression. This subclass of algorithms treats the input waveform as a *full band signal* and signal spectrum is flattened by linear prediction.

We can think of a better performance when we deal with the signal by dividing its frequency spectrum into M parts and treat each part separately i.e subband coding. If frequency spectrum of the signal is divided into M bands, M is called the order of the subband decomposition. Length of these bands can be uniform or nonuniform.

In case of tree structure decomposition, only a pair of low and high pass filters are used repeatedly to achieve subband decomposition of a given order. One possible way to have three channel decomposition is given in figure(1)

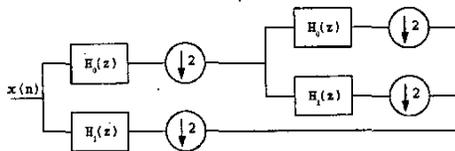


Figure 1. A 3-channel tree-structured filter bank

It is obvious that band length is not same in this case, so it is a non-uniform filter bank, and coding gain expression will be given by equation(1) [1].

$$G_{SBC} = \frac{\sum_{k=0}^{M-1} \frac{\sigma_k^2}{m_k}}{\prod_{k=0}^{M-1} (\sigma_k^2)^{\frac{1}{m_k}}} \quad (1)$$

This idea can be extended to two dimensional filter bank using separable filters. The subband coder structure in two dimensional case is, therefore, similar to one dimensional case. By using the uniform separable filters for the tree structured decomposition of the two dimensional signal, the coding gain expression can directly extended to equation(2)

$$G_{SBC} = \frac{\sum_{k=0}^{M-1} 2^{-2R_{av}} \sigma_x^2}{\sum_{k=0}^{M-1} \frac{2^{-2R_k} \sigma_x^2}{m_k}} \quad (2)$$

Where m_1k and m_2k are the decimation factors in the horizontal and vertical directions.

2. PROBLEM FORMULATION

In the uniform filter bank, width of the band depends upon the order of the subband decomposition. But when we talk about non-uniform filter bank there exist infinite possibilities even when order is fixed. These infinite possibilities can be brought down to a certain number, when we use tree structure decomposition to obtain non-uniform decomposition of the bands, with a given pair of low pass and high pass filter. Among these possibility which type of decomposition will give the maximum coding gain? An example of the block diagram of two dimensional subband decomposition using separable filters is given as shown in figure(2). Z_1 and Z_2 are the Z-transform variable in horizontal and vertical direction respectively.

Wavelet decomposition always chooses the LL band to further decompose. This way of decomposing leads to a special structure. Now the question is, does decomposition of LL band always gives the maximum coding gain? In this work we have shown the answer in the negative. It should be noted that number of bands in n-level

S. No.	image	N=3 O/W	N=6 O/W	N=9 O/W	N=12 O/W
1	lena	17.3762/15.5712	19.7408/19.1555	20.5611/19.9508	21.0983/20.1246
2	baboon	13.8183/12.6737	15.9189/15.4332	16.8469/16.2039	17.4476/16.4095
3	crowd	10.6306/10.3241	12.6956/12.4958	13.3398/12.9316	13.7200/13.0111
4	barbara	14.1651/13.2462	16.4692/16.0407	17.1659/16.6897	17.7459/16.8391
5	boats	15.8494/15.0795	18.7061/18.3636	19.7442/19.1281	20.2729/19.2977

Table 2. Comparison of coding gain values with that of wavelet type decomposition

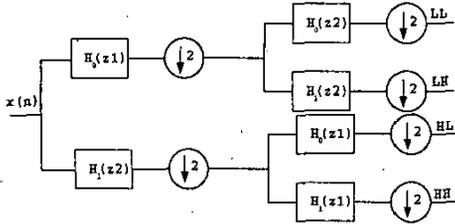


Figure 2. Block diagram of a two dimensional filter bank

S. No.	image	N=3 O/E	N=4 O/E
1	lena	17.3762/17.3762	18.5033/18.5033
2	baboon	13.8183/13.8183	14.8020/14.8020
3	crowd	10.6306/10.6306	11.6768/11.6768
4	barbara	14.1651/14.1651	15.1133/15.1133
5	boats	15.8494/15.8494	17.1892/17.1892

Table 1. comparison with exhaustive search for N=3 and N=4

wavelet type standard decomposition is given by $3n + 1$. Example given in figure(2), shows 1-level wavelet type standard decomposition [2]. Since we are obtaining this by tree structure decomposition, so total number of possibilities for four band decomposition will be fixed and it can be shown in that assuming separability total number of possibility will be given by $N!2^N$, where $N + 1$ stands for order of subband decomposition (N stands for no. of cuts). It is obvious that n -level wavelet type standard decomposition will among one of the possibility and can be shown that is not always the optimum one. Since it is very difficult to find out optimum solution exhaustively as number of possibility will be very large for large value of N . So we are proposing an alternative way by providing an heuristic algorithm to find out maximum coding gain.

3. RELATED WORK

Natural signals like image are not stationary. Therefore, for efficient compression one needs to segment it in the time domain to approximately locally stationary pieces. Thus, for image decomposition, one can segment it in the time domain, or decompose in the frequency domain or both [3]. Some time or/and frequency domain decom-

position methods are briefed below.

Pruning: This may be either in time domain or in frequency domain. First, a complete (binary) tree structure upto the smallest decomposition is built up. At every stage, two leaves are pruned. These two leaves are chosen such that either they have greater cost than their parent node, or pruning them reduces the cost least among all the possibilities at a certain level. The algorithm ends when the tree structure corresponds to the number of desired bands. The globally optimum tree is found under certain conditions by starting at the deepest level of the tree and pruning pairs of branches that have higher cost than their parent branch at each level while working back towards the root. This is the algorithm which is used in [4] and [5].

Single tree algorithm: Another possible approach to finding the best tree in time or frequency domain is the "greedy tree growing" algorithm which starts at the root and divides the signal/spectrum in two if it is profitable to do so (if the cost of the sub signals/subbands generated is less than the cost of the signal/band they come from). It terminates when no more profitable splits remain. This algorithm, however, does not find the globally optimum tree [4] [5].

Double tree algorithm: A shortcoming of the single tree algorithm is that it, say, selects the best frequency decomposition tree and retains that tree for the whole signal. If the signal is non-stationary, the algorithm will choose the basis that works well in some average way for the whole signal, but can not adapt the frequency decomposition tree for different time segments of the signal. Another algorithm that attacks this problem of segmenting the signal both in the time and in the frequency domain is the double tree algorithm [6]. This algorithm first calculates the best single tree for the complete signal and stores the cost. Then it finds the best single trees for the two halves segmented in the time domain and stores the cost. again it finds out the best single trees for the signal segmented in time into quarters, and so on. These costs can be written in the binary tree structure and that can be pruned by using the same single tree algorithm as before [7].

4. HEURISTIC SEARCH ALGORITHM

Our aim is to find out what will be the optimum decomposition of an image when order of the subband coding N is given. First of all we will find out the power spectral density of the image. At each level except level one, we examine only four possibilities. Among these, one decomposition will be decided by the cost function to decompose further and to proceed to the next level.

First of all we will do horizontal and vertical decomposition of an image by given nearly ideal filter pair, and find corresponding *psd's* of each sections of each decomposition. Then find out the cost defined in the equation(4) and choose that decomposition which has more cost as that will be indication of achieving the required decomposition in which all the cells have flatness equation(3) one or nearly one.

Let the desired number of cuts or depth be N (number of subbands is $N + 1$). One step of the algorithm is described now. Let the present depth be L (present number of subbands or cells = $L + 1$).

1. The best decomposition at level $L - 1$ is known. Consider the cell in this decomposition that contains the (0,0) or dc coefficient. Cut this cell both horizontally and vertically to obtain two possibilities 1 and 2. Now, in the level $L - 1$ best decomposition, find the cell (excluding the (0,0) cell) that has the minimum flatness to power of reciprocal of multiplication of decimation factors. Cut this cell both horizontally and vertically to obtain two other possibilities 3 and 4. Flatness of a cell or subband is defined as,

$$\frac{\exp(\frac{1}{2\pi} \int_0^{2\pi} \ln(S_{xx}(\omega)) d\omega)}{\frac{1}{2\pi} \int_0^{2\pi} S_{xx}(\omega) d\omega} \quad (3)$$

2. At level L we now consider only these four possibilities among the $2^L L!$ possibilities. For each possibility, define the weight

$$W_j = \text{geometric mean}(W_{L,j,1}, W_{L,j,2}, \dots, W_{L,j,L+1}) \quad (4)$$

where index j varies from 1 to 4, and

$$W_{L,j,k} = \text{flatness}(L, j, k) \frac{1}{m_{1k} m_{2k}} \quad (5)$$

where index L indicates the present depth, index j provides the numbering to the four possibilities, index k indicates numbering for the cell varying from 1 to $L + 1$, and m_{1k}, m_{2k} are the decimation factors in the horizontal and vertical direction for the k^{th} subband.

3. Choose the largest W_j among these four as the best decomposition to be considered in the next stage.

Explanation: When we have N desired cuts ($N + 1$ cells or bands), we would like to have a final decomposition in

which all the cells should be flat or nearly flat. In other words, no region of low energy should be masked in a cell. The algorithm starts from one cell and reaches to N cuts i.e ($N + 1$ bands). At every stage the algorithm decides which cell to be cut. Since we want to find out the non-flat regions and expose the smoother regions of low energy which are masked by the regions of high energy, intuitively it feels that a lesser flatness measure of a cell would indicate that there may be regions of low energy which should be exposed (by a cut) to increase the coding gain. Therefore, when the weight (cost function) defined above for a decomposition is highest at some level, it exposes the smoother regions of low energy the most, and this decomposition is chosen for passing to the next level. Taking the flatness to the power the reciprocal of the product of the decimation factors will ensure that we end up with the decomposition of a low flatness cell occupying larger area as it may mask low energy regions. i.e we want to choose a smoother region of a larger cell. Even at the first $L = 1$ stage also we will pass the decomposition which will have more weight. After getting the best decomposition passed to the next level, the question arises as to which cell we have to choose for a further cut? Here we would like to emphasize on the point that images are generally low pass signals. Therefore, among all the cells, the cell which contains the (0,0) coefficient will possess maximum power and minimum flatness in most of the cases. So we choose that cell which contains the (0,0) coefficient. Further, we choose the cell among the rest which will have minimum flatness. And in both of the cells we will have horizontal and vertical cut to obtain four decompositions to analyse further. At the final stage we will choose the decomposition of maximum coding gain, among the four.

5. RESULTS

In the table 1 we have compared the results obtained by the algorithm and result obtained by the exhaustive search. O/E refers to obtained/exhaustive search result. We can see that we are getting the same result, for $N = 3$ and $N = 4$. For larger values of N number of possibilities are so large that it will take a big space to store the results of exhaustive search, which is not possible due to the memory constraint.

In the table 2 we can see the gain over wavelet type decomposition, although we are not using the wavelet filters but nearly ideal DFT filters. O/W refers to obtained/wavelet result. Subband of the order of 4,7,10 and 13 corresponds to wavelet type decomposition of level 1,2,3 and 4.

In the figure(3) we can see the performance of the algorithm described above over the wavelet type decomposition. The upper flat line shows the maximum coding gain that can be obtained by infinite order predictive coding, and infinite order subband decomposition. Of

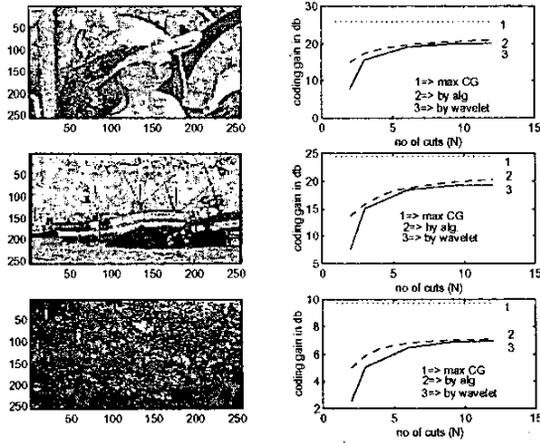


Figure 3. Plots for leena, boats and a high variation image

the remaining two curves, the top one in each case is the result obtained by the algorithm while the bottom one is the wavelet result.

In the figure(4), we can see that decomposing the

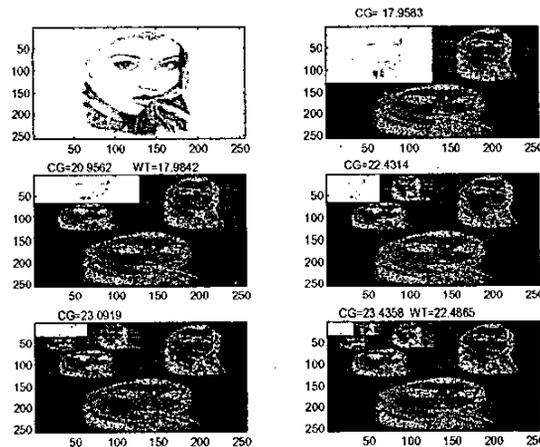


Figure 4. Trend of optimum decomposition with increasing N from 2 to 6

cell corresponding to (0,0) component of *psd*, as done in wavelet, is not always optimum and we can find out better coding gain by decomposing the other cells as well. Typical wavelet type decomposition(WT) can give us modest coding gain but that will not be necessarily optimum one. In the figure WT refers to coding gain corresponding to wavelet type decomposition, and CG is the coding gain for the given decomposition.

REFERENCES

- [1] A. Soman and P. P. Vaidyanathan, "Coding Gain in Paraunitary Analysis and Synthesis Systems," *IEEE Trans. on Signal Processing*, SP-41(5):1824-1835, May 1993.
- [2] R. M. Rao and A. S. Bopardikar, "Wavelet Transforms: Introduction to Theory and Applications," Addison-Wesley, 1998.
- [3] Cormac Herley, Zixiang Xiong, Kannan Ramchandran, and Michael T. Orchard, "Joint Space Frequency Segmentation using Balanced Wavelet Packet Trees for Least-Cost Image Representation," *IEEE Trans. on Image Processing*, vol. 6, no. 9, September 1997, pp. 1213-1230.
- [4] R. Coifman and M. Wickerhauser, "Entropy-based Algorithms for Best Basis Selection," *IEEE Trans. Information Theory*, vol. 38, March 1992, pp. 713-718.
- [5] K. Ramchandran and M. Vetterli, "Best Wavelet Packet Bases in a Rate-distortion Sense," *IEEE Trans. Image Processing*, vol. 2, 1993, pp. 160-175.
- [6] C. Herley and M. Vetterli, "Orthogonal Time-varying Filter Banks and Wavelet Packets," *IEEE Trans. signal Processing*, vol. 42, Oct. 1994, pp. 2650-2663.
- [7] Philip A. Chau, Tom Looabaugh and Robert Gray, "Optimal Pruning with Applications to Tree Structured Source Coding and Modeling," *IEEE Trans. on Information Theory*, vol. 35, no. 2, March 1989, pp. 299-315.