Frequency and Spatially Adaptive Wavelet Packets

John R. Smith and Shih-Fu Chang

Columbia University Center for Telecommunications Research and
Electrical Engineering Department, New York, N.Y. 10027

jrsmith@ctr.columbia.edu and sfchang@ctr.columbia.edu

Abstract

In this paper we consider a method for image compression based on frequency and spatially adaptive wavelet packets. We present a new fast directed acyclic graph (DAG) structured decomposition, with both spatial segmentation and orthogonal frequency branching from each node. Whereas traditional wavelet packet decomposition adapts to a global frequency distribution, this technique finds the best joint spatial segmentation and local frequency basis. The algorithm is derived from the fast double tree algorithm proposed by Herley, et. al., for 1-D signals [1], with an extension to 2-D and modification to include spatial segmentation of frequency nodes. By collecting redundant nodes in this full adaptive tree, we have derived a directed acyclic graph (DAG) structure which contains the same number of nodes as the double tree, but includes new connections between nodes. We present the adaptive wavelet packet DAG algorithm and examine image compression performance on test images.

1. INTRODUCTION

By including spatial segmentations in the search for the most efficient basis for representation, image decomposition via spatially adaptive wavelet packets adapts to image non-stationarity as well as local frequency content. Previously [2][3], wavelet packets have been used to arbitrarily partition the image frequency plane to adapt to signal frequency characteristics. However, the search for the best orthogonal wavelet packet basis considers only a single unsegmented signal space, and thus does not adapt to signal non-stationarity.

When fixed spatial decomposition is performed prior to frequency decomposition, the effect of non-stationarity can often be reduced. DCT, as used in *JPEG*, and other block-based transforms use this approach. Alternatively, the frequency bands themselves may be partitioned [4][5] and coded separately using fixed spatial segmentations. However, these approaches do not find the best joint frequency decomposition and spatial segmentation. We propose that the frequency and spatially adaptive wavelet packet decomposition do just this.

2. SPATIALLY ADAPTIVE TREE STRUCTURE

The adaptive decomposition may be easily organized into a tree structure as indicated in Figure 2(a). Starting with the full image at the root node, each subsequent node is decomposed both spatially (S) and spectrally (F). The spatial decomposition is carried out using a quad-tree split. Since the spatial partitions do not overlap, this segmentation allows the four children nodes to be further analyzed independently. The frequency decomposition is performed using a four channel quadrature mirror filter (QMF) bank. This produces a library of wavelet packet bases which consist of orthonormal functions that are easily constructible from a single filter kernel [2]. An element of the wavelet packet basis may be identified completely by its tree position and original filter kernel. This characteristic has made wavelet packets extremely tractable for image compression.

The extension of wavelet packets to include optimization across spatial segments was proposed in [1]. There, a double tree was formulated such that the best wavelet packet basis is found for a hierarchy of binary segmentations of the signal, then the optimum one is chosen. The overall search identifies the most efficient joint spatial segmentations and corresponding wavelet packet decompositions based on an additive cost function. We extend this double tree, with some modifications, to attain a more extensive best basis search which will yield improved compression performance.

2.1 Rate-Distortion Criterion

In keeping with the goal of image compression, and as in [1], we adopt a rate-distortion (R-D) framework. This two-sided criterion is a more general case of one-sided criterion such as entropy only [2], or distortion only. The operational R-D characteristic is constructed at each node by sweeping through a series quantization step sizes and measuring the first order entropy (R) and mean square error (D). Uniform quantizers are used for all nodes which contain the d.c. frequency component, while a deadzone quantizer centered at zero is used for all other nodes. The search for best basis then combines the search for minimum cost expansion with minimum (R-D) operating costs at nodes. The optimization uses adaptive pruning of the tree based on the Lagrangian cost function, $J(\lambda) = D + \lambda R$. In general, selection of optimal basis and R-D operating points (quantizers) at each node will be a function of the signal characteristics, original filter kernel, and the target rate for compression. A detailed presentation of this technique can be found in [6].
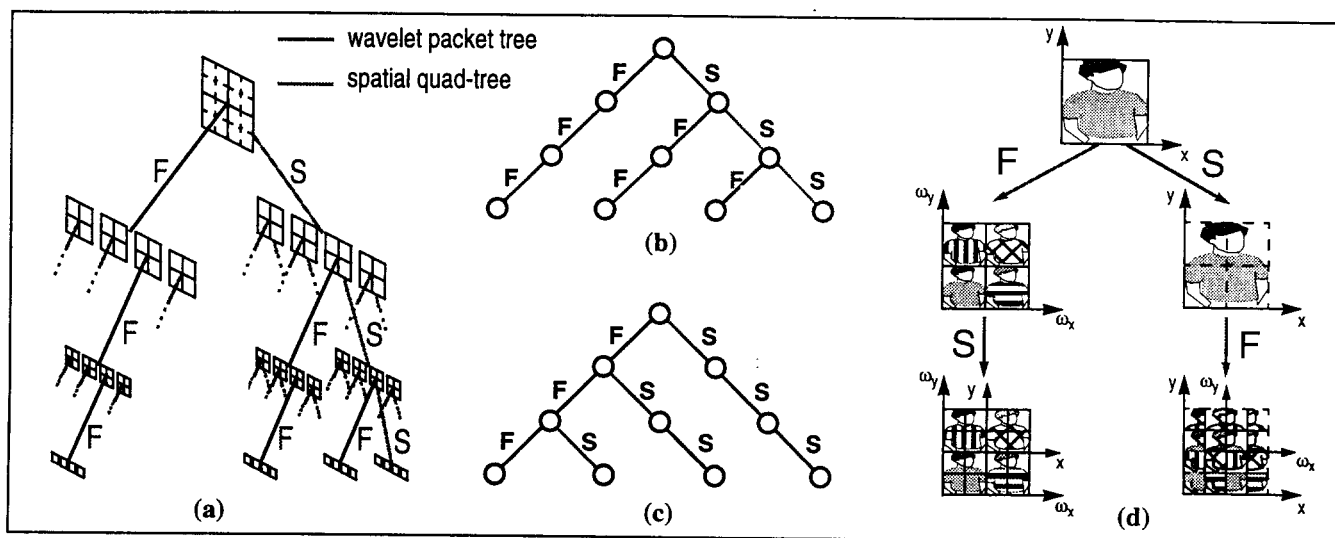
FIGURE 1. (a) Adaptive double tree proposed by Herley, et. al. [1], applied here for images. Notice that a separate wavelet packet tree is generated for each spatial segmentation. (b) Symbolic representation of the double tree in (a), (c) Dual of double tree, with spatial segmentation of frequency nodes. (d) Approximate commutativity of spatial and frequency decompositions; notice that only arrangements of blocks in the bottom two images differ.

2.2 Double Tree

The double tree structure was proposed to find jointly the best wavelet packet basis and signal segmentation [1]. Essentially, this corresponds to growing individual wavelet packet trees for a hierarchy of segmentations of the signal. This is illustrated for images in Figure 1(a). The symbolic representation, which will be used in the remainder of this paper, appears in Figure 1(b). The double tree algorithm was applied to various 1-D signals in [1], and produced improved results over pure wavelet packets. However, although the double tree considers segmentations of the original signal in the selection of the best basis, it does not evaluate the benefits of segmenting the frequency (F) nodes of the wavelet packet trees. In other words, a more complete tree than the double tree can be generated.

2.3 Full Adaptive Tree

Circumstances may dictate that it is more efficient to split a signal first based on frequency, then select which frequency nodes to segment (split spatially), rather than first splitting spatially followed by frequency. The double tree considers the latter case in the optimization, but not the former. To accommodate all combinations of frequency and spatial splits, all nodes in the tree must be split both spatially and by frequency. We call the tree with both splits at each node a full adaptive tree. This tree is illustrated using symbolic notation in Figure 2(a). The full adaptive tree can be contrasted to the double tree, which splits nodes spatially only if no prior frequency splits were performed; compare Figure 1(b) with Figure 2(a). The full adaptive tree combines the double tree, Figure 1(a) with its dual, Figure 1(c).
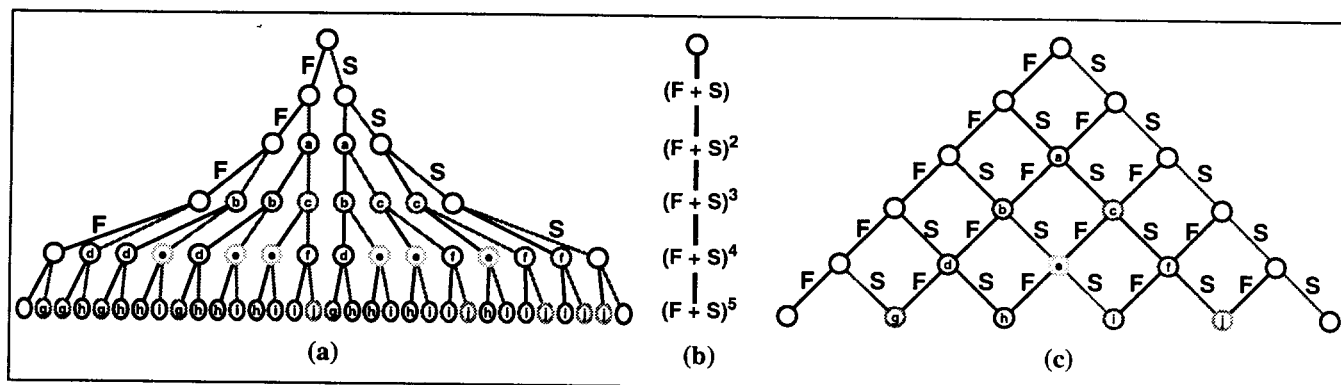

FIGURE 2. (a) Full adaptive tree with frequency and spatial branching from each node. Letters indicate redundant nodes. (b) Equivalent polynomial expansion. (c) Commutativity illustrated in Figure 1(b) allows the conversion of the full adaptive tree into this directed acyclic graph. Note, in these symbolic representations, it appears as if all children of a parent node are further spawned collectively using either frequency or spatial branching. In actual implementation, each child branches independently of the other children. This gives the trees a depth into the page, which is illustrated more completely in Figure 4.

	JPEG	wavelet	spatial quad-tree	wavelet packet	double tree	double tree dual	adaptive DAG
Spatial decomp.	Fixed	None	Adaptive	None	Hierarchical	Adaptive	Adaptive
Freq. decomp.	Fixed	Fixed	None	Adaptive	Adaptive	Hierarchical	Adaptive
# nodes: $D = depth$	$\frac{w}{8} \cdot \frac{h}{8}$	$3 \cdot D + 1$	$\sum_{i=0}^{D} 4^i$		$\sum_{i=0}^{D} (i+1) \cdot 4^i$		
0.1 bpp, PSNR =	22.7 dB	23.8 dB	N/A	25.4 dB	25.4 dB	25.5 dB	**25.5 dB**
0.5 bpp, PSNR =	28.3 dB	29.5 dB	19.1 dB	32.7 dB	32.7 dB	32.9 dB	**33.0 dB**
1.0 bpp, PSNR =	33.1 dB	34.6 dB	26.5 dB	37.1 dB	37.1 dB	37.6 dB	**37.7 dB**
2.0 bpp, PSNR =	38.9 dB	40.7 dB	34.7 dB	43.0 dB	43.0 dB	43.8 dB	**43.8 dB**

TABLE 1. Comparison of several adaptive image compression algorithms using Barbara image.

2.4 Adaptive Directed Acyclic Graph

While the full adaptive tree offers a more complete decomposition of the signal, the number of over-expansions of the signal grows exponentially with tree depth. This is clearly less favorable than the double tree which grows only linearly with tree depth. Clearly with the full adaptive tree, the new combinations come at considerable computational cost! However, by looking more closely at the series of branchings in the full adaptive tree, we can reduce the full adaptive tree into an equivalent graph structure that is the same size as the double tree.

As illustrated in Figure 1(c), the payoff comes from recognizing the commutativity in the series of branchings in the full adaptive tree. For example, for two step hybrid branching, we have FS = SF, and for three step hybrid branching, SFF = FSF = FFS. In other words, if the order of operations (S and F) is reversed, the equivalent nodes are generated. In the two step case, for example, there are two paths leading to each hybrid node. This implies that the double tree and it's dual have identical nodes, but with different connectivity.

When the full adaptive tree is modified accordingly to collect the redundant nodes, a directed acyclic graph (DAG) structure is produced. The corresponding adaptive directed acyclic graph (ADAG) appears in Figure 2(c). Interestingly enough, it has the same nodes as the double tree, but with a rich variety of new connections between nodes. Now the selection of best basis corresponds to selecting the best paths through the ADAG. The Lagrangian method of joint best basis selection and bit allocation used for the double tree, and mentioned earlier, applies to the ADAG as well. Figure 4(a) gives another look at the full adaptive tree and the ADAG for depth = 2, with all children nodes included.

2.5 Compression Results

The double tree algorithm, its dual, and ADAG wavelet packet compression were applied to several test images. The results show a significant increase in performance over non-adaptive methods, as indicated in Table 1. The ADAG and dual double tree show improvement over the double tree. Interestingly, the double tree most often arrives at the same result as wavelet packets. The reason is that for images, it is usually too expensive to sacrifice globally an iteration of wavelet decomposition for a spatial segmentation. Spatial segmentations, however are often chosen liberally in the dual double tree and the ADAG, when the effect is isolated to a sub-tree, where the trade-off is still beneficial.

The tabulated results are based on compression of the Barbara image, using the Johnston filter QMF12a in the wavelet packet filter banks [7]. The overhead cost for coding the tree was also embedded in the optimization. Bits were added to the rate (R) terms of the R-D points at each node as follows: since 16 quantizers were evaluated at each node, 4 bits were added to identify the optimum quantizer. To describe the position of each node, $3 \cdot depth + 4$ additional bits were added to each node, where depth is the distance from the node to the root. Clearly the overhead becomes more significant as the number of nodes in the basis increases. By including the overhead bits in the optimization, the algorithm evaluates if the penalty of selecting nodes deeper in the tree is worthwhile.

3. CONCLUSION

We have presented a new algorithm for image compression using spatially and frequency adaptive wavelet packets. This approach adapts jointly to the image spatial and frequency distributions. We started by extending the double tree algorithm to images. Then the double tree structure was modified to include spatial segmentation of frequency bands, to form a more complete tree. Following from the commutativity of hybrid series of branchings, the full adaptive tree was then reduced into the adaptive wavelet packet directed acyclic graph (ADAG). We tested both the double tree, its
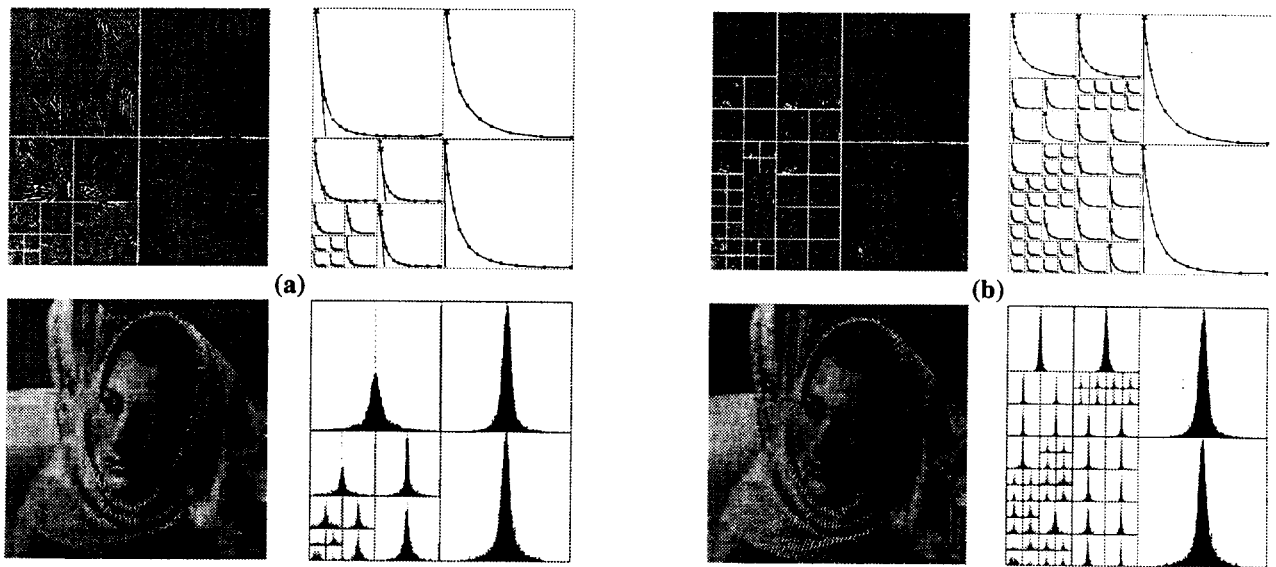
FIGURE 3. Compression of face image of Barbara at 0.25 bpp, (a) wavelet (24.9dB), (b) ADAG compression (27.2 dB). Clockwise, (i) optimal basis, (ii) D-R curves for basis, X's indicate optimal D-R points and slopes (rescaled), (iii) histograms, (iv) reconstructed image. Spatial segmentations indicated with black borders, frequency decompositions with white borders.

dual and the new ADAG and found improved performance using these spatially adaptive approaches. Having identical over-expansions of the image data, the ADAG offers greater potential for optimal basis identification, and thus better performance. Further, the ADAG benefits from the same fast basis selection approach devised for the double tree [1][6].

References

[1] C. Herley, J. Kovacevic, K. Ramchandran and M. Vetterli, "Tilings of the Time-Frequency Plane: Construction of Arbitrary Orthogonal Bases and Fast Tiling Algorithms," *I.E.E.E. Transactions on Signal Processing*, December 1993.

[2] R. Coifman and M. V. Wickerhauser, "Entropy-Based Algorithms for Best Basis Selection," *I.E.E.E. Transactions on Information Theory*, Vol. 38, No. 2, March 1992.

[3] M. L. Wickerhauser, "Picture Compression by best-basis subband coding," Yale University, May 1990.

[4] J. Woods and S.'Neil, "Subband Coding of Images," *I.E.E.E. Transactions on Acoustics, Speech and Signal Processing*, Vol. ASSP-34, No. 5, October 1986.

[5] P. W. Wong and Steven Noyes, "Space-Frequency Localized Image Compression," *I.E.E.E. Transactions on Image Processing*, Vol. 3, No. 3, May 1994.

[6] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *I.E.E.E. Transactions on Image Processing*, 1992.

[7] J. D. Johnston, "A Filter Family for Use in Quadrature Mirror Filter Banks, "*Proc ICASSP*, pp. 291--294, 1980.
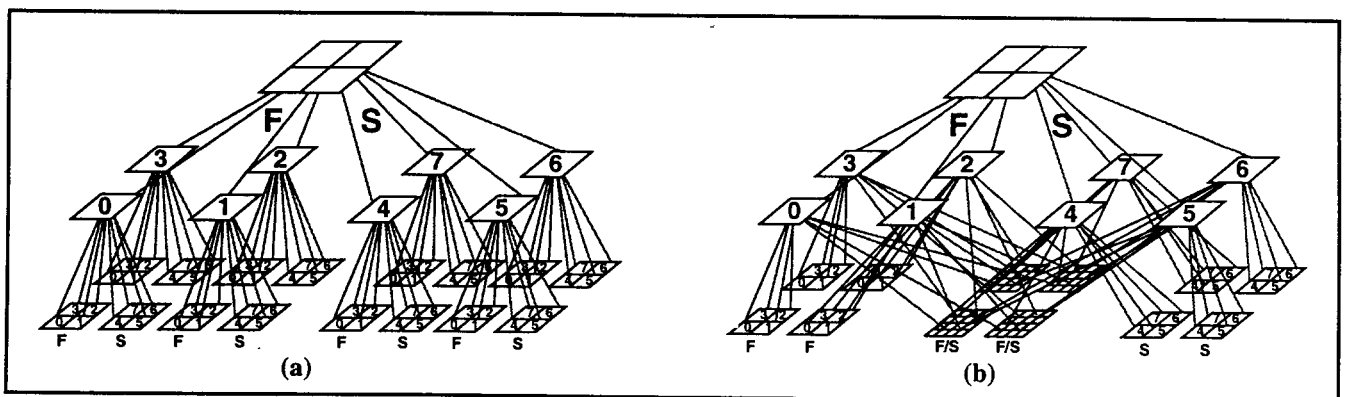
FIGURE 4. (a) Full Tree for depth = 2, showing branching from all children nodes; frequency children are numbered 0 - 3, and spatial children are 4 - 7. After two generations, redundant nodes will have codes that are inverses, i.e., node 2-5 = node 5-2, (b) after collecting redundant nodes, corresponding digraph for depth = 2.

2236