JOINT ADAPTIVE SPACE AND FREQUENCY BASIS SELECTION

John R. Smith

IBM T.J. Watson Research Center
30 Saw Mill River Road
Hawthorne, NY 10532
jrsmith@watson.ibm.com

Shih-Fu Chang

Dept. of Electrical Engineering
Columbia University
New York, N.Y. 10027
sfchang@ctr.columbia.edu

ABSTRACT

We develop a new method for building a representation of an image from a library of basis elements that is facilitated by a joint adaptive space and frequency (JASF) graph. The JASF graph combines partitionable frequency expansion and spatial segmentation of the image, symmetrically. We demonstrate by using a rate-distortion framework for basis selection that the JASF graph improves compression performance over recent wavelet packet and double-tree methods by offering exponentially more bases in which to represent the images.

1. INTRODUCTION

In this paper, we present a method for the jointly-adaptive, space and frequency selection of an image basis. The joint adaptive space and frequency (JASF) graph provides a symmetric decomposition of the image into a library of spatially and frequency localized basis elements. By performing the frequency expansions in a partitionable-form, the JASF graph provides a commutativity in the frequency and spatial operations which allows the basis elements to be more efficiently indexed by a graph.

1.1. Adaptive image decomposition

Recent methods have been developed for adaptively compressing images using space- or frequency-based image decompositions that involve tree structured basis selection methods [1, 3, 2]. The objective is to derive a segmentation or filter bank that is customized to the image. The two extreme approaches decompose the images either by frequency, such as wavelet packets (WP) [1], or spatially, such as quad-tree (QT) segmentation. Hybrid approaches such as the double-tree (DT) incorporate both segmentation and frequency expansion, but do so asymmetrically [2].

The JASF graph, illustrated in Figure 1, treats the space and frequency operations symmetrically in a graph structured cascade [4]. The WP-tree, QT and DT are embedded within the JASF graph. Since the JASF graph provides a more complete decomposition of the image and generates a larger number of alternative bases, the JASF graph improves image compression performance over these methods.

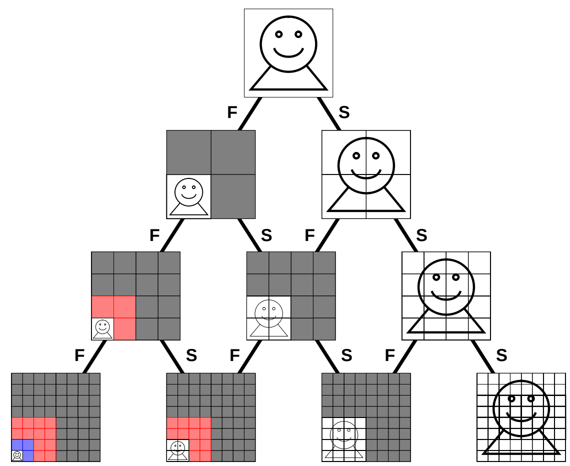


Figure 1: The JASF graph generates a joint space and frequency decomposition of the image.

In general, the tree- and graph-based decompositions generate libraries of basis elements. A *basis element* consists of a set of *basis functions* that is generated and coded as a group; each basis element corresponds to one node in the tree or graph. The objective of the search for the best basis is to select the set of basis elements that have the least total coding cost and provide a "complete" set of basis functions. The completeness requirement guarantees perfect reconstruction in the absence of quantization.

1.2. Outline

We present the JASF graph and the JASF basis generation and selection system. We present the framework for *partitionable* frequency expansions, which are fundamental to the JASF graph. We show that by using partitionable expansions, the generation of the basis elements and the reconstruction of the image may follow a number of equivalent paths. We demonstrate examples of image coding by extending the fast minimum rate-distortion cost basis selection method developed in [3]. We demonstrate that the JASF graph provides approximately 10^{20} more bases than the DT which generates a basis element library the same size as the JASF graph and improves compression by 0.5 to 0.9dB.

2. PARTITIONABLE EXPANSIONS

In order to provide commutativity in the frequency and segmentation operations in the JASF graph, the frequency expansions are performed in a partitionable form. In general, any orthonormal expansion produced by filter banks is not partitionable but may be made partitionable as we explain shortly. Producing the JASF graph expansion of depth M, requires frequency analysis matrices $\mathbf{H}_0$ and $\mathbf{H}_1$ that are at least $(M-1)$-partitionable.

2.1. 1-partitionable expansion

We define the 1-partitionable frequency expansion (1-PFE) as follows:

Definition 1 *A frequency expansion matrix* $\mathbf{H}$ *is* 1-*partitionable if and only if it has only zeros in the upper right and lower left quadrants.*

If the frequency expansion matrix set $\{\mathbf{H}_0, \mathbf{H}_1\}$ generates a 1-PFE then the expansion is comprised of separate expansions over the two half-length signals, which we now illustrate. First, observe that for QMF filter banks the finite-signal frequency expansion matrices $\mathbf{H}_i$, $i \in \{0,1\}$, can be written in the following form [4]:

$$\mathbf{H}_i = \begin{pmatrix} \mathcal{H}_i^a & \mathcal{H}_i^b \\ \mathcal{H}_i^b & \mathcal{H}_i^a \end{pmatrix}. \quad (1)$$

We construct the 1-PFE from $\mathbf{H}_i$ by $\mathcal{H}_i = \mathcal{H}_i^a + \mathcal{H}_i^b$ as follows, note that for a length N signal, this corresponds to circular convolution of period $N/2$:

$$\mathbf{H}_i^{(1)} = \begin{pmatrix} \mathcal{H}_i & \mathbf{0} \\ \mathbf{0} & \mathcal{H}_i \end{pmatrix} = \begin{pmatrix} \mathcal{H}_i^a + \mathcal{H}_i^b & \mathbf{0} \\ \mathbf{0} & \mathcal{H}_i^a + \mathcal{H}_i^b \end{pmatrix}. \quad (2)$$

We now state the following useful results and definition: if the original frequency expansion set $\{\mathbf{H}_0, \mathbf{H}_1\}$ satisfies the condition of perfect reconstruction and orthonormality, that is $\mathbf{H}_0^T\mathbf{H}_0 + \mathbf{H}_1^T\mathbf{H}_1 = \mathbf{I}^N$, then the 1-PFE set $\{\mathbf{H}_0^{(1)}, \mathbf{H}_1^{(1)}\}$ is orthonormal and the set $\{\mathcal{H}_0, \mathcal{H}_1\}$ generates an orthonormal expansion of length $N/2$.

Definition 2 *The 1-partitionable frequency expansion (1-PFE) matrix set* $\{\mathbf{H}_0^{(1)}, \mathbf{H}_1^{(1)}\}$*, as constructed above is* 1*-partitionable and orthonormal if and only if* $\{\mathbf{H}_0, \mathbf{H}_1\}$ *satisfies the perfect reconstruction condition. For proof, see [4].*

The 1-PFE and segmentation are combined into a joint space and frequency expansion using a graph of depth = 2 as follows: in each of the four decomposition and reconstruction paths depicted in Figure 2: start from x, generate v_{ij}'s, and resynthesize x, we have perfect reconstruction of x. That is,

$$v_{ij} = \mathbf{S}_j^{N/2}\mathbf{H}_i^{(1)}x \text{ and } v_{ij} = \mathbf{H}_i^{(1)}\mathbf{S}_j^N x, \text{ and}$$

$$\begin{aligned} x &= \mathbf{G}_0^{(1)}v_{00} + \mathbf{G}_1^{(1)}v_{10} + \mathbf{G}_0^{(1)}v_{01} + \mathbf{G}_1^{(1)}v_{11}, \text{ and} \\ x &= \mathbf{G}_0^{(1)}(v_{00}+v_{01}) + \mathbf{G}_1^{(1)}(v_{10}+v_{11}), \end{aligned} \quad (3)$$

where $\mathbf{G}_i = \mathbf{H}_i^T$, and $\mathbf{S}^N$ is an $N \times N$ segmentation matrix [4].

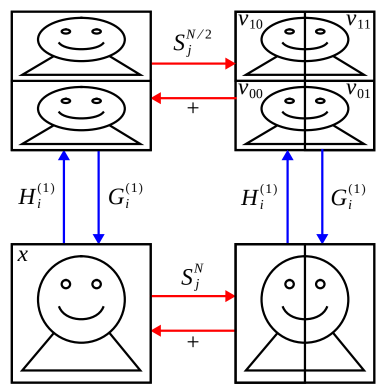


Figure 2: JASF graph of depth = 2: 1-partitionable frequency expansions (1-PFE) $(H_i, i \in \{0,1\})$ and segmentations $(S_j, j \in \{0,1\})$.

2.2. M-partitionable expansion

In order to construct the JASF graph of depth ≥ 2, we generalize the 1-PFE to the M-partitionable frequency expansion (M-PFE) as follows:

Definition 3 *A frequency expansion matrix* $\mathbf{H}_i$ *is* M*-partitionable if and only if the upper left and lower right quadrants are* $(M-1)$*-partitionable.*

The M-PFE analysis transform matrices $\mathbf{H}_i^{(M)}$ for $i \in \{0,1\}$ are described recursively as follows:

$$\mathbf{H}_i^{(M)} = \begin{pmatrix} \mathbf{H}_i^{(M-1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_i^{(M-1)} \end{pmatrix}. \quad (4)$$

As a result, we have that, in general, $\mathbf{H}_i^{(M)}$ are block diagonal with 2^M partitions, as follows:

$$\mathbf{H}_i^{(M)} = \underbrace{\left(\begin{array}{c|c|c} \begin{array}{c|c} \mathcal{H}_i & 0 \\ \hline 0 & \mathcal{H}_i \end{array} & \ddots & \mathbf{0} \\ \hline \mathbf{0} & \ddots & \mathbf{0} \\ \hline \mathbf{0} & \ddots & \begin{array}{c|c} \mathcal{H}_i & 0 \\ \hline 0 & \mathcal{H}_i \end{array} \end{array} \right)}_{2^M \text{ partitions}} . \quad (5)$$

Definition 4 *The M-PFE matrix set* $\{\mathbf{H}_0^{(M)}, \mathbf{H}_1^{(M)}\}$ *is M-partitionable and orthonormal if and only if the matrix set* $\{\mathbf{H}_0, \mathbf{H}_1\}$ *satisfies the perfect reconstruction condition (follows from Eq 4 and proof in [4] for 1-PFE case).*

3. TREE AND GRAPH EXPANSIONS

The tree- and graph-based decompositions differ in the sizes of their basis element libraries and/or the number of bases they provide. The results are summarized in Table 1 for a depth = 6 image decomposition. Examples of bases from the WP, QT, DT and JASF graph image decompositions are illustrated in Figure 3.

	QT	WP	DT	JASF	RSFT
# basis elements	1365	1365	7737	7737	37449
# bases	10^{78}	10^{78}	10^{127}	10^{147}	10^{147}
expansion factor	6	6	21	21	1365

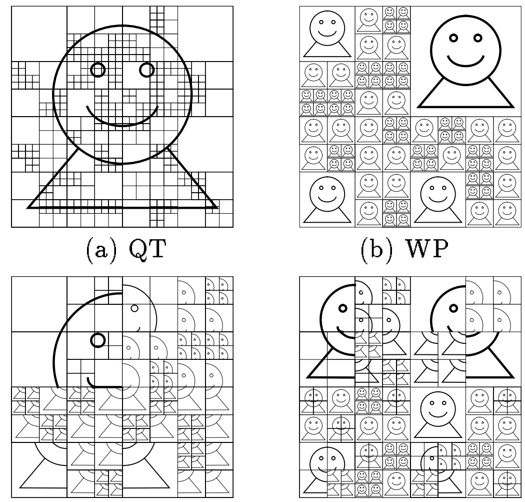
Table 1: Comparison of image decompositions of depth = 6 using QT, WP, DT, JASF and RSFT.

3.1. Single-trees

The WP and spatial QT image decompositions utilize single-trees. Each generates a library of N_s basis elements from which may be chosen B_s bases to represent the image. For a single-tree of depth=D with splitting factor β (for both quad-tree segmentation and four-band subband decomposition of images, $\beta = 4$), N_s is given recursively by $N_s(D) = 1 + \beta N_s(D-1)$ and B_s is given by $B_s(D) = 1 + B_s(D-1)^\beta$, where $N_s(0) = B_s(0) = 0$. For $D = 6$ the single-trees generate $B_s(6) \approx 10^{78}$ bases.

3.2. Double-tree (DT)

The DT generates a separate WP-tree for each spatial node in the spatial QT. The DT increases the number of basis elements and number of bases. The DT generates $N_d(D) = N_s(D) + \beta N_d(D-1)$ basis elements and $B_d(D) = 1 + B_s(D-1)^\beta + B_d(D-1)^\beta$ bases, where $N_d(0) = B_d(0) = 0$. For $D = 6$ the DT generates $B_d(6) \approx 10^{127}$ bases.



(a) QT (b) WP (c) DT (d) JASF graph

Figure 3: Example QT, WP, DT and JASF graph image bases. Each rectangle (node) corresponds to a selected basis element. (a) The QT nodes are image segments. (b) The WP nodes are image subbands. (c) The DT nodes are image-segment subbands. (d) The JASF graph nodes are, equivalently, segment-subbands and subband-segments.

3.3. JASF graph

The JASF graph integrates the spatial and partitionable frequency expansions, symmetrically. The JASF graph generates the same number of basis elements as the DT but significantly more bases, $B_g(D) = 1 + 2B_g(D-1)^\beta + B_g(D-2)^{\beta^2}$ bases, where $B_g(0) = 0$ and $B_g(1) = 1$. For $D = 6$ the JASF graph generates $B_d(6) \approx 10^{147}$ bases.

3.4. RSFT

By symmetrically combining segmentation with *non-partitionable* frequency expansion, a redundant space and frequency tree (RSFT) is generated. The RSFT increases the number of basis elements to $N_r(D) = 1 + 2\beta N_r(D-1)$ and the number of bases to $B_r(D) = 1 + \beta B_r(D-1)^\beta$, where $N_r(0) = 0$, $B_r(0) = 0$ and $B_r(1) = 1$.

When the frequency expansion is inherently *partitionable* (i.e., Haar filter bank has $\mathbf{H}_i$'s which are already block-diagonal, that is $\mathbf{H}_i = \mathbf{H}_i^{(M)}$), the RSFT and JASF graph generate the identical basis elements. However, the RSFT generates multiple copies of each basis element. For example, in Table 1, of the 37,449 basis elements generated by the RSFT, only 7,737 are unique.

Otherwise, when using a non-partitionable frequency expansion in the RSFT, many of the basis elements are nearly redundant. The difference between many of the basis elements stems only from the border extension used in the filtering operations. We have observed that these additional nearly redundant RSFT basis elements provide for little gain in compression performance, while they greatly increase the complexity.

4. JASF BASIS SELECTION

The selection of a basis from the JASF graph involves a three-way decision at each node: (1) choose F (frequency-expansion), (2) choose S (segmentation), or (3) choose neither. An example of a selected basis from the JASF graph is depicted in Figure 4. The basis selection procedure is carried out as follows:

1. Assign a coding cost ($J_i = R_i + \lambda D_i$) to each basis element (node) i in the JASF graph, where R_i, D_i gives the rate-distortion at trade-off λ for basis element i (see [3]).

2. Starting from the root node, and recursively at each F and S child node, choose the least cost path:
$$\min(\sum J_{i,f_k}, \sum J_{i,s_k}, J_i),$$
where $\sum J_{i,f_k}$ is the total cost of the F child path, $\sum J_{i,s_k}$ is the total cost of the S child path, and J_i cost of choosing neither.

3. The final embedded tree gives the basis with the least total cost.

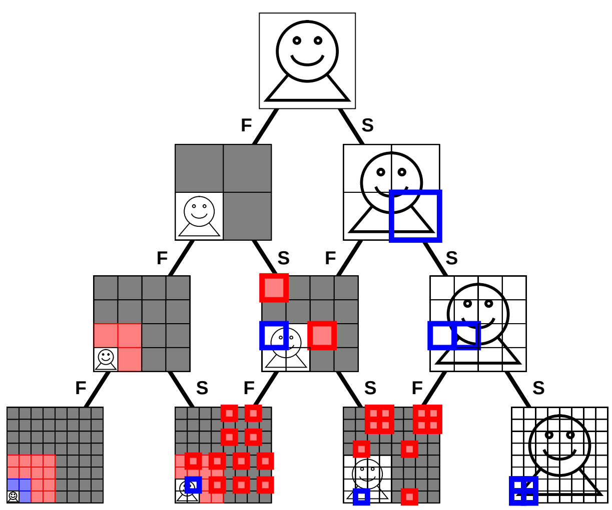


Figure 4: Example basis selected from JASF graph.

5. COMPRESSION EVALUATION

We now examine the compression performance of the basis selection procedure carried out on the spatial QT, WP-tree, DT, JASF graph and RSFT. We used a twelve-tap QMF filter for the frequency expansions. For the JASF graph, the frequency-expansions were carried out in the partitionable-form as discussed in Section 2. The results, given in Table 2, show that the JASF graph improves image compression performance over the spatial QT, WP-tree and DT. Furthermore, the bases selected by the JASF graph are not available in the spatial QT, WP-tree or DT. We see also that the RSFT tree provides no compression improvement over the JASF graph. The addition in the RSFT of the nearly-redundant basis elements does not improve the compression performance.

	spatial QT	WP tree	DT	JASF graph	RSFT
0.25 bpp	N/A	27.9 db	27.9 db	**28.4** db	**28.4** db
0.5 bpp	19.0 db	32.3 db	32.3 db	**32.7** db	**32.7** db
1.0 bpp	25.1 db	36.8 db	36.8 db	**37.5** db	**37.5** db
2.0 bpp	33.1 db	42.9 db	42.9 db	**43.8** db	**43.8** db

Table 2: Compression results on the *Barbara* image.

6. SUMMARY

We developed a method for the jointly adaptive, space and frequency selection of an image basis. The JASF generates a symmetric decomposition of the image by combining partitionable frequency expansion and spatial segmentation. The JASF graph generates a greater number of bases than recent wavelet packet (WP), spatial quad-tree (QT) and double-tree (DT) methods. The bases are selected from the JASF graph by choosing at each node from a frequency expansion, spatial segmentation or neither. We demonstrated that image compression performance using the JASF graph improves over the QT, WP-tree and DT.

7. REFERENCES

[1] R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Trans. Inform. Theory*, 38(2), March 1992.

[2] C. Herley, J. Kovačević, K. Ramchandran, and M. Vetterli. Tilings of the time-frequency plane: Constructions of arbitrary orthogonal bases and fast tiling algorithms. *IEEE Trans. Signal Processing*, December 1993.

[3] K. Ramchandran and M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Trans. Image Processing*, June 1993.

[4] J. R. Smith. *Integrated Spatial and Feature Image Systems: Retrieval, Analysis and Compression.* PhD thesis, Graduate School of Arts and Sciences, Columbia University, New York, NY, 1997.