Image Compression by Parameterized-Model Coding of Wavelet Packet Near-Best Bases

Carl Taswell

Scientific Computing and Computational Mathematics
Stanford University, Stanford, CA 94305-2140
Internet e-mail: taswell@sccm.stanford.edu

ABSTRACT

Top-down tree search algorithms with non-additive information cost comparisons as decision criteria have recently been proposed by Taswell[9,10] for the selection of near-best bases in wavelet packet transforms. Advantages of top-down non-additive near-best bases include faster computation speed, smaller memory requirement, and extensibility to biorthogonal wavelets in addition to orthogonal wavelets. A new compression scheme called parameterized-model coding was also proposed and demonstrated for one-dimensional signals.[10] These methods are extended here to two-dimensional signals and applied to the compression of images. Significant improvement in compression while maintaining comparable distortion is demonstrated for parameterized-model coding relative to quantized-scalar coding. In general, the lossy compression scheme is applicable for low bit rate coding of the M largest packets of wavelet packet decompositions with wavelet packet basis libraries and the M atoms of matching pursuit decompositions with time-frequency atom dictionaries.

Keywords: Image compression, parameterized-model coding, matching pursuit decomposition, wavelet packet decomposition, near-best basis, non-additive information cost, top-down tree search.

1 INTRODUCTION

Matching pursuit decompositions of images with time-frequency atom dictionaries[7,1] yield finite weighted linear combinations $f(x,y) \approx \sum_{i=1}^{M} a_i w_{\gamma_i}(x,y)$ that approximate the image functions $f(x,y)$ with weighting or amplitude coefficients a multiplying the atomic waveforms $w_\gamma(x,y)$. These atoms are labeled by the multi-index γ identifying the atom within a large collection $\{w_\gamma(x,y)|\gamma \in \Gamma\}$ of such atoms. Both amplitudes a_i and indices γ_i constitute necessary information representing the image. This decomposition is usually obtained with the atoms ordered by decreasing absolute value of the amplitudes. Moreover, if the image is discretized with a total of N pixels, then the matching pursuit decomposition is computed with $M \ll N$ atoms. Thus, a complete basis decomposition with N atomic waveforms in the transform domain is never computed. However, when such a basis with N waveforms is computed for the wavelet transform, coding of the waveforms with the M largest amplitudes has also been shown to be a method important for image compression.[5] For both cases, the M waveforms of a matching pursuit decomposition or the M largest waveforms of a complete basis decomposition, how should the amplitude and index information be further compressed and coded?

This question is investigated here in the context of wavelet packet transforms.[3] For wavelet packet dictionaries, the time-frequency atoms are wavelet packets for which the amplitudes a are the wavelet packet transform coefficients and the waveforms $w_\gamma(x, y)$ are the wavelet packet basis functions labeled by the multi-index γ specifying scale, orientation, and position information. In this paper, the compression and coding of wavelet packet bases is studied for the M largest packets obtained from near-best bases selected by a top-down tree search algorithm with non-additive information cost comparisons as the decision criterion.[9,10] Advantages of top-down non-additive near-best bases include faster computation speed, smaller memory requirement, and extensibility to biorthogonal wavelets in addition to orthogonal wavelets. In a previous study of these wavelet packet near-best bases,[10] a new compression scheme called parameterized-model coding was proposed and demonstrated for one-dimensional (1-D) signals. These methods are extended here to two-dimensional (2-D) signals, applied to the compression of images, and proposed as a possible answer to the question posed above. Feasibility of the parameterized-model coding scheme is demonstrated for compression of both orthogonal and biorthogonal wavelet packet decompositions of images. In particular, significantly improved rates of compression are demonstrated for parameterized-model coding relative to quantized-scalar coding with comparable rates of distortion.

2 METHODS

2.1 Wavelet Packet Decompositions

We consider here various data structures for representing information relevant to 2-D wavelet packet decompositions. A 2-D discrete packet transform is considered to be any multiresolution transform (such as a wavelet packet transform or local trigonometric transform) that yields a table of transform coefficients which can be organized as a balanced quaternary tree. The table is called a discrete packet table $\mathbf{P}^{\text{table}}$ with levels l and blocks b of the table corresponding to levels l and branches b of the tree. For the sake of mnemonics, the term branch is often used here instead of the more customary term node. However, the conventional term node is also used synonymously in this paper. Thus, the root node at level 0 and the terminal nodes at level L are considered to be the top and bottom of of the full balanced tree corresponding to the finest and coarsest resolutions of the data.

There are 4^l blocks on each level and thus $K = (4^{(L+1)} - 1)/3$ blocks in the entire table. Within each block b on level l, there are $4^{-l}N$ cells c where $N = N_1 N_2$ is the number of elements in the original data matrix $\mathbf{X} \in \mathbb{R}^{N_1 \times N_2}$. Thus each coefficient in the packet table $\mathbf{P}$ can be specified as the 4-vector $[a, l, b, c]$ where a is the packet's amplitude and l, b, and c are its level, block, and cell indices. These level, block, and cell indices also correspond to scale, orientation, and position indices. The position index can be a scalar if the cells in each block are labelled $0, \ldots, 4^{-l}N - 1$. Otherwise the position index must be a multi-index specifying both row and column indices for the cell (which is a pixel in the case of an image).

Thus, if the data $\mathbf{X} \in \mathbb{R}^{N_1 \times N_2}$ is an image with $N = N_1 N_2$ pixels, then the wavelet packet transform to a depth of L levels yields a packet table matrix $\mathbf{P}^{\text{table}} \in \mathbb{R}^{N_1 \times (L+1)N_2}$ with a total of $(L + 1)N$ coefficients. A particular basis within this redundant representation can be specified with the basis selection tree $\mathbf{S} \in \chi^K$ where each of the K variables $\chi = \{0, 1\}$ is an indicator variable for the selection of the k^{th} block/branch of the table/tree. The redundant table $\mathbf{P}^{\text{table}} \in \mathbb{R}^{N_1 \times (L+1)N_2}$ can then be converted to the non-redundant basis $\mathbf{P}^{\text{basis}} \in \mathbb{R}^{N_1 \times N_2}$. In WavBox 4,[11] the function *dpt2dpb* (discrete packet table to discrete packet basis) performs this restructuring of the data via the mapping $\mathbf{P}^{\text{basis}} = \text{dpt2dpb}(\mathbf{P}^{\text{table}}, \mathbf{S})$.

To compare various decompositions, it is convenient to convert discrete packet tables $\mathbf{P}^{\text{table}}$ or bases $\mathbf{P}^{\text{basis}}$ to discrete packet lists $\mathbf{P}^{\text{list}}$ representing the selected decompositions. In WavBox 4, the functions *dpt2dpl* and *dpb2dpl* perform this restructuring of the data via the mappings $\mathbf{P}^{\text{list}} = \text{dpt2dpl}(\mathbf{P}^{\text{table}}, \mathbf{S})$ and $\mathbf{P}^{\text{list}} = \text{dpb2dpl}(\mathbf{P}^{\text{basis}}, \mathbf{S})$ where again the functions are named as the abbreviations for their input and output data structures analogous to the naming convention for *dpt2dpb*. Each list contains M packets specified as row 4-vectors $[a_i, l_i, b_i, c_i]$ with rows $i = 1, ..., M$ ordered so that $|a_1| \geq \cdots \geq |a_M|$. To study a complete basis decomposition, we must examine the entire list where $M = N$. However, we may also study subsets of the list where $M < N$, for example, where we choose $M = \mathcal{N}^2_{.99} < N$ (see Section 2.2 for explanation of the data compression number $\mathcal{N}^p_f$).

Thus, there are four data structures presented here: $\mathbf{P}^{\text{table}} \in \mathbb{R}^{N_1 \times (L+1) N_2}$, $\mathbf{P}^{\text{basis}} \in \mathbb{R}^{N_1 \times N_2}$, $\mathbf{P}^{\text{list}} \in \mathbb{R}^{M \times 4}$, and $\mathbf{S} \in \chi^K$. Since packet tables $\mathbf{P}^{\text{table}}$ and selection trees $\mathbf{S}$ are implemented respectively as matrices and vectors, table blocks and corresponding tree branches indexed by (l, b) are respectively submatrices and scalars. They are denoted $\mathbf{P}^{\text{table}}_{lb} \equiv \mathbf{P}^{\text{table}}_{i_{lb}, j_{lb}}$ and $S_{lb} \equiv S_{k_{lb}}$ where for $l \in \{0, 1, \ldots, L\}$ and $b \in \{0, 1, \ldots, 4^l - 1\}$, the row and column vector indices i_{lb}, j_{lb} are for level l block b in a table matrix, and the scalar index k_{lb} is for level l branch b in a tree vector. The same holds true analogously for packet bases $\mathbf{P}^{\text{basis}}_{lb} \equiv \mathbf{P}^{\text{basis}}_{i_{lb}, j_{lb}}$ with the proviso that not all levels l and blocks b of $\mathbf{P}^{\text{table}}$ are stored in $\mathbf{P}^{\text{basis}}$ since $\mathbf{P}^{\text{basis}}$ is not redundant by its definition as a basis. In fact, $\mathbf{P}^{\text{basis}}$ contains only those blocks b on levels l for which $S_{lb} = 1$. Correct manipulation of coefficients stored in $\mathbf{P}^{\text{basis}}$ requires using the level-block indexing information encoded as logical values in $\mathbf{S}$. Finally, with regard to packet lists $\mathbf{P}^{\text{list}}$, the i^{th} packet together with its index information is denoted $\mathbf{P}^{\text{list}}_i \equiv [a_i, l_i, b_i, c_i]$.

2.2 Non-Additive Information Costs

We consider data matrices $\mathbf{X} \in \mathbb{R}^{N_1 \times N_2}$ for one parent block and $\mathbf{Y}_1, \ldots, \mathbf{Y}_4 \in \mathbb{R}^{N_1/2 \times N_2/2}$ for four children blocks in a wavelet packet table represented by a quaternary tree. We wish to compare their information costs by some measure used as a decision criterion when searching the quaternary tree. For decompositions of 2-D images, these costs are computed elementwise for each submatrix block. Thus, the blocks can be described equivalently as vectors obtained by reindexing the elements of the corresponding submatrices so that we compare the parent $\mathbf{x} \in \mathbb{R}^N$ with the children $\mathbf{y}_1, \ldots, \mathbf{y}_4 \in \mathbb{R}^{N/4}$. Additive costs were originally intended for use with the best bases of Coifman and Wickerhauser.[3] Non-additive costs were proposed for use with the near-best bases of Taswell.[9,10] Here we review some of the relevant definitions as needed for this paper.

Definition: A cost functional $\mathcal{C}^{\text{add}}$ from vectors $\mathbf{x} \in \mathbb{R}^N$ to $\mathbb{R}$ is called an *additive information cost function* if $\mathcal{C}^{\text{add}}(0) = 0$ and $\mathcal{C}^{\text{add}}(\mathbf{x}) = \sum_i \mathcal{C}^{\text{add}}(x_i)$.

Definition: A cost functional $\mathcal{C}^{\text{non}}$ from vectors $\mathbf{x} \in \mathbb{R}^N$ to $\mathbb{R}$ is called a *non-additive information cost function* if it serves as a decision criterion for a basis selection algorithm and it is not an additive information cost function $\mathcal{C}^{\text{add}}$.

Definition: The inequality $\mathcal{C}^{\text{non}}(\mathbf{x}) \leq \mathcal{C}^{\text{non}}(\mathbf{y}_1, \ldots, \mathbf{y}_4)$ between vectors $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y}_1, \ldots, \mathbf{y}_4 \in \mathbb{R}^{N/4}$ is called a *non-additive information cost comparision* if $\mathcal{C}^{\text{non}}(\mathbf{y}_1, \ldots, \mathbf{y}_4) \equiv \mathcal{C}^{\text{non}}(\oplus_{i=1}^4 \mathbf{y}_i) \neq \sum_{i=1}^4 \mathcal{C}^{\text{non}}(\mathbf{y}_i)$.

We can construct several examples of $\mathcal{C}^{\text{non}}$ starting from the sorted vector $[x_{(k)}]$ where $x_{(1)} = |x_{i_1}| \geq \cdots \geq x_{(N)} = |x_{i_N}|$ so that $x_{(k)} = |x_{i_k}|$ is the k^{th} largest absolute value element of the vector $[x_i]$. We continue by constructing the decreasingly sorted, powered, and cumulatively summed vector $[v_k(\mathbf{x}, p)]$, and renormalized vector $[u_k(\mathbf{x}, p)]$ where

$$u_k(\mathbf{x}, p) = \frac{v_k(\mathbf{x}, p)}{v_N(\mathbf{x}, p)} \quad \text{with} \quad v_k(\mathbf{x}, p) = \sum_{i=1}^{k} x^p_{(i)}$$

which then makes it convenient finally to define the examples of $\mathcal{C}^{\text{non}}$. (Note that $0 \leq u_k(\mathbf{x}, p) \leq 1$ because of the normalization.) Thus, with $[x_{(k)}]$ and $[u_k(\mathbf{x}, p)]$ obtained from $[x_i]$, define the non-additive information cost functions

$$\begin{aligned} \mathcal{C}^{\text{non}}_1(\mathbf{x}) &= \mathcal{N}^p_f(\mathbf{x}) = \arg\min_k |u_k(\mathbf{x}, p) - f| \\ \mathcal{C}^{\text{non}}_2(\mathbf{x}) &= \mathcal{A}^p(\mathbf{x}) = N - \sum_k u_k(\mathbf{x}, p) \end{aligned}$$

which are respectively the data compression number and data compression area.[9,10]

2.3 Near-Best Bases

For mathematical convenience, we interpret the matrices $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{N_1 \times N_2}$ as the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ with $N = N_1 N_2$ and consider the orthonormal transformation matrix $\mathbf{B} \in \mathbb{R}^{N \times N}$. Then $\mathbf{y} = \mathbf{Bx}$ and $\mathcal{C}(\mathbf{y})$ are respectively the coefficient vector and information cost scalar for $\mathbf{x}$ in the coordinate system represented by the basis $\mathbf{B}$. We wish to find a basis $\mathbf{B}$ for which $\mathcal{C}(\mathbf{y})$ is minimal, subject possibly to some constraint on the search for the basis $\mathbf{B}$. Coifman and Wickerhauser defined the best basis for additive costs.[3] Subsequently, Taswell defined the near-best basis for non-additive costs,[9] and then more generally for both non-additive and additive costs.[10] Here we review definitions for both best and near-best bases.

Definition: The *best basis* relative to $\mathcal{C}^{\text{add}}$ for a vector $\mathbf{x}$ in a library $\mathcal{B}$ of bases is that $\mathbf{B}$ for which $\mathcal{C}^{\text{add}}(\mathbf{Bx})$ is minimal.

Definition: The *near-best basis* relative to $\mathcal{C}$ (either $\mathcal{C}^{\text{non}}$ or $\mathcal{C}^{\text{add}}$) for a vector $\mathbf{x}$ in a library $\mathcal{B}$ of bases is that $\mathbf{B} \in \mathcal{B}^* \subset \mathcal{B}$ for which $\mathcal{C}(\mathbf{Bx})$ is minimal subject to the constraints of the search within the subset $\mathcal{B}^*$ defined by the search type.

The near-best basis with additive or non-additive costs permits either a bottom-up or top-down search through the table to find the basis selection tree $\mathbf{S}$. Searches subject to other patterns of constraint are possible as well. The various search methods are denoted $\mathcal{S}$ in general with $\mathcal{S} = \mathcal{U}$ and $\mathcal{S} = \mathcal{D}$ indicating bottom-up and top-down in particular. In WavBox 4, the function *dpt2bst* performs this mapping from discrete packet table to basis selection tree as $\mathbf{S} = \text{dpt2bst}(\mathbf{P}, \mathcal{S}, \mathcal{C})$. Also in WavBox 4, the function *wpt* for the wavelet packet transform is one of several discrete packet transform functions, and *dcnum* for the data compression number is the mathematical function $\mathcal{N}_f^p(\mathbf{x})$ described above in Section 2.2 with default values of $p = 2$ and $f = 0.99$ for the parameters. Using these functions, then the sequence of function calls

$\mathbf{P}^{\text{table}} = \text{wpt}(\mathbf{X})$
$\mathbf{S} = \text{dpt2bst}(\mathbf{P}^{\text{table}}, \mathcal{S}, \mathcal{C})$
$\mathbf{P}^{\text{list}} = \text{dpt2dpl}(\mathbf{P}^{\text{table}}, \mathbf{S})$
$M = \text{dcnum}(\mathbf{P}^{\text{list}}(1:N, 1))$
$\mathbf{P}^{\text{list}} = \mathbf{P}^{\text{list}}(1:M, 1:4)$

yields a wavelet packet decomposition returned as a packet list truncated to the M largest absolute value packet coefficients constituting 99% of the energy of the transform (and of the original data if the transform mapping is orthonormal). While valid for near-best basis decompositions in general (for further details and discussion, consult previous papers[9,10]), this approach fails to exploit the advantages which can be potentially gained in particular for top-down tree searches with $\mathcal{S} = \mathcal{D}$.

2.4 Top-Down Tree Searches

We wish to design an appropriate algorithm specialized for top-down tree searches with $\mathcal{S} = \mathcal{D}$ operating via $\mathbf{P}^{\text{basis}}$ instead of the more general algorithm operating via $\mathbf{P}^{\text{table}}$ as described in Section 2.3. Naming this function *wpddb* for Wavelet Packet Decomposition by top-Down Basis search, then the pseudocode segment

$[\mathbf{P}^{\text{basis}}, \mathbf{S}] = \text{wpddb}(\mathbf{X})$
$\mathbf{P}^{\text{list}} = \text{dpb2dpl}(\mathbf{P}^{\text{basis}}, \mathbf{S})$

$M = \text{dcnum}(\mathbf{P}^{\text{list}}(1:N,1))$
$\mathbf{P}^{\text{list}} = \mathbf{P}^{\text{list}}(1:M,1:4)$

incorporating *wpddb* replaces the more general one incorporating *wpt* and *dpt2bst* from Section 2.3. It is also possible to combine the two functions *dpb2dpl* and *dcnum* so that the M-packet truncated list is returned directly from the combined function. The current method of first returning the complete N-packet list from the function *dpb2dpl*, returning M from the function *dcnum*, and then truncating the N-packet list to an M-packet list requires significantly more memory. This memory requirement is not necessary and can be eliminated with use of the combined function if packets $\mathbf{P}^{\text{list}}_{M+1}, \ldots, \mathbf{P}^{\text{list}}_{N}$ are never used in subsequent processing.

Because the algorithm runs unidirectionally downward in the tree, it can be performed essentially "in place", thus significantly reducing memory storage requirements from approximately $O((L+1)N)$ for $\mathbf{P}^{\text{table}}$ in *wpt* and *dpt2bst* to $O(2N)$ for $\mathbf{P}^{\text{basis}}$ and a temporary copy in *wpddb*. Furthermore, because the algorithm does not necessarily require that the entire table and tree be generated and searched, it can be performed with significant savings in machine operations and computing time. This reduction in computational cost corresponds to a number $\hat{L}$ representing the number of levels of the transform that need to be computed. This number is estimated by summing over all levels the fraction of computed blocks to total blocks on each level. Computed blocks include all ancestral blocks from the root block to the parental blocks above the selected blocks, the selected blocks themselves, and the four children blocks below each selected block (unless the selected block is already at the maximum level L). This estimate yields $\hat{L}$ as a rational (not necessarily integer) number that ranges between 1 and L. Thus, the "in place" algorithm reduces computational costs from approximately $O(LN)$ for $\mathbf{P}^{\text{table}}$ in *wpt* and *dpt2bst* to $O(\hat{L}N)$ for $\mathbf{P}^{\text{basis}}$ in *wpddb*. The amount by which $\hat{L} \leq L$ is dependent on the data and the image class. However, the reduction in memory storage requirements from $O((L+1)N)$ to $O(2N)$ is independent of image class. Nevertheless, for both issues of memory storage and computational cost, the savings for 2-D images (relative to 1-D signals) can be significant even for small values of N and small differences between $\hat{L}$ and L.

2.5 Compression and Distortion

For experiments investigating lossy compression of images, we wish to minimize the distortion D resulting between the reconstruction $\hat{\mathbf{X}}$ and the original $\mathbf{X}$ following compression and coding of the wavelet packet decomposition $\mathbf{P}^{\text{list}}$. Compression can be achieved by truncating the N packets in the list to the $M < N$ largest absolute-value packets and then quantizing and coding the remaining M packets analogous to the method used for wavelet transforms[5] instead of wavelet packet transforms. Standard methods of coding data include scalar and vector quantization.[6] The quantization and coding of the M packets remaining after truncation of the list applies only to the amplitudes a and not to the level-l, block-b, and cell-c indices which must be coded without loss of information. Since the (l,b,c)-index information is retained for each packet retained in this compression scheme, it is possible to consider other coding schemes for the packet amplitudes a.

Taswell[10] proposed a new method incorporating parameter estimation by statistical regression modeling with $Y_i = f(X_i, \boldsymbol{\theta}) + \epsilon_i$ for some function f, parameter $\boldsymbol{\theta}$, and noise ϵ_i. In particular, Taswell[10] estimated the parameters of the curve traced by a plot of the dependent variable Y_i taken to be the sums v_i defined in Section 2.2 as a function of the independent variable X_i taken to be the index i. This concave curve ascends smoothly to a horizontal aymptote. It can be modeled by a low order polynomial with coefficients estimated by (possibly weighted) linear regression. Alternatively, it can be approximated by a variety of well-known nonlinear models with a small number of parameters including those that can be expressed as generalized linear models and as non-linear dose-response models.[8]

Here we consider modeling the smooth curve traced by a plot of the normalized sums u_i (also as defined in Section 2.2 except with the u_i obtained from normalization of v_i by v_M instead of v_N) as a function of the normalized index $(i-1)/M$. Scaling the problem with these normalizations improves the numerical conditioning of the estimation problem. For this parameterization of the variables then, the coding scheme for the amplitude coefficients of the packet list $\mathbf{P}^{\text{list}}$ consists of the steps

1. retain signs s_i of the packet amplitudes a_i,
2. compute the powered cumulative sums v_i and retain the normalization sum v_M,
3. normalize the sums v_i by v_M to obtain the u_i,
4. regress on curve of $Y_i \equiv u_i$ versus $X_i \equiv (i-1)/M$ estimating $\boldsymbol{\theta}$ in $Y_i = f(X_i, \boldsymbol{\theta}) + \epsilon_i$,
5. quantize the estimate $\hat{\boldsymbol{\theta}}$ to desired level of precision,

and the decoding scheme consists of the steps

1. set $X_i \equiv (i-1)/M$ and estimate $\hat{Y}_i = f(X_i, \hat{\boldsymbol{\theta}})$,
2. set $\hat{u}_i \equiv \hat{Y}_i$ and multiply by retained sum v_M to estimate $\hat{v}_i$,
3. finite difference the $\hat{v}_i$ and multiply by retained signs s_i to estimate $\hat{a}_i$.

The choice of regression model has not been specified in the coding and decoding scheme listed above. Actual implementation of the method requires specifying the functional model f for the shape and fit of the regression curve. Furthermore, the model can also be enhanced by specifying the nature of any batch processing done. For example, regression can be performed on groups of packet amplitudes selected by level l ("level grouping"), or by level l and block b ("block grouping"), or only by packet list order i index without regard to level l and block b indices ("order grouping"). In the latter case, the number M_g of packets in the g^{th} of G groups can be variable or fixed, whereas in the other cases, M_g is determined by the selection criteria. Of course, in all cases, those packets selected for the batch processing of each group do remain in correct decreasing absolute value order so that the coding scheme described above can be applied.

In this paper, we focus our initial pilot study on comparing various batch processing methods used in the statistical regression model for the packet amplitude coefficients. Thus we ignore entropy coding of indices and calculation of compression rates based on bit counts. Instead, we consider two stages of compression: Stage 1 with a fixed compression rate of M packets truncated from the list of N packets obtained from the complete basis decomposition, and Stage 2 with a variable compression rate R measured by summing the number J_g of regression model coefficients estimated for the M_g packets in the g^{th} group parameterized by $\boldsymbol{\theta}_g \in \mathbb{R}^{J_g}$. Then for the original $\mathbf{X}$ and reconstructed $\hat{\mathbf{X}}$ images, we compare relative ℓ^p-norms of reconstruction errors computed as measures of distortion $D_p = \|\mathbf{X} - \hat{\mathbf{X}}\|_p / \|\mathbf{X}\|_p$ for $p = \{1, 2, \infty\}$.

In particular, we perform Stage 1 compression only, simulate scalar quantization of the packet amplitudes by rounding them, call this procedure "quantized scalar coding" (QSC) of the packet amplitudes, and report a compression rate of $R = M$. Alternatively, we perform both Stage 1 and Stage 2 compression, simulate scalar quantization of the model parameter coefficients by rounding them, call this procedure "parameterized model coding" (PMC) of the packet amplitudes, and report a compression rate of $R = \sum_{g=1}^{G} J_g$. In both procedures, we ignore the lossless compression obtained by entropy coding the code indices associated with scalar quantization such as those used in Huffman or arithmetic coding. Furthermore, we also ignore the lossless compression associated with entropy coding the packet (l, b, c) indices making the assumption that it's performed independently of coding the packet a coefficients and thus the same for all the different models. As a consequence, the comparisons here are meant to be valid for the different models relative to each other within this study and not relative to some standard requiring compression with actual entropy coding of all the information (both packet amplitudes and packet indices).

3 RESULTS

Experiments were performed on the test image "Elaine" with 640×480 pixels displayed in the upper left of Figure 1. This image with a total of $N = 307,200$ pixels was scaled to zero-mean unit-variance and then wavelet packet transformed to a near-best basis selected by a top-down search $\mathcal{S} = \mathcal{D}$ with the non-additive cost function $\mathcal{C}^{\text{non}} = \mathcal{A}^1$ for the data compression area. The transform and search were performed to level $L = 5$ with Daubechies' orthogonal least asymmetric length-8 filters[4] (DOLA 8) and the $\mathbf{P}^{\text{basis}}$ of $N = 307,200$ packets was converted and truncated to a $\mathbf{P}^{\text{list}}$ of $M = \mathcal{N}^2_{0.99} = 17813$ packets. This procedure was repeated with the Bradley-Brislawn biorthogonal symmetric analysis length-7 synthesis length-9 filters[2] (BBBS 7,9), yielding a truncated $\mathbf{P}^{\text{list}}$ of $M = 18568$ packets. The same circular-periodized convolution version was used for both orthogonal and biorthogonal sets of filters which were chosen for their similar length 8 versus (7,9).

For orthogonal and biorthogonal decompositions, the computed values of $\hat{L} = 2.95$ and $\hat{L} = 3.05$ respectively were both significantly less than $L = 5$ thus demonstrating one of the important advantages of the top-down tree search over the bottom-up tree search.[10] After obtaining both decompositions, the M packets from each were coded by PMC and QSC as explained in Section 2.5. For PMC, the model used for fitting the data was a polynomial of variable order found by minimizing the regression error subject to the constraint of between 6 to 12 total coefficients in the model for each group of packets. For PMC with order grouping, the maximum number of packets per group was set to 1200 insuring that the compression ratio from packet coefficients to model coefficients for each group could be no greater than 1200/6 or 200/1. For both PMC and QSC, all retained coefficients were rounded to 11 bits per mantissa and 5 bits per exponent for a total of 2 bytes per coefficient.

Table 1 presents the compression rate R and distortion rates D_p described in Section 2.5 for the image Elaine when coded by the various methods. Distortion rates for PMC with order grouping were comparable to distortion rates for QSC despite a significant decrease in the rate R from 17813 to 120 with a ratio of ≈ 148 for the orthogonal case and from 18568 to 142 with a ratio of ≈ 131 for the biorthogonal case. Since all coefficients were rounded (simulating quantization) to 2 bytes, simple ratios of numbers of coefficients permit valid comparisons of the different versions of PMC relative to each other but not to QSC so that the above ratios are misleading. For valid comparison between PMC and QSC, the information retained in the sign bits for PMC must also be considered. Therefore, total bits for PMC with order grouping was $17813 \cdot 1 + 120 \cdot 16 = 19733$ and $18568 \cdot 1 + 142 \cdot 16 = 20840$ for the orthogonal and biorthogonal cases respectively; and total bits for QSC was $17813 \cdot 16 = 285008$ and $18568 \cdot 16 = 297088$ respectively. Thus, the improvement ratio for PMC relative to QSC was ≈ 14.

Figure 1 displays the original and some of the reconstructed images for the orthogonal case. Since the original image was scanned at 100 dpi adequate for redisplay on video monitors with 72 dpi, the various images appear granular when laser printed at 300 dpi. This granularity on printing Figure 1 confounds the subjective visual judgement of lossy compression artifacts. However, when viewed on a video display monitor, the artifacts appeared as patches or blotches that were either inappropriately lightened or darkened. These artifacts were least noticeable for PMC with order grouping. This subjective result was consistent with the objective result of minimal distortion rates D_p for PMC with order grouping relative to the other methods of grouping.

4 DISCUSSION

Parameterized-model coding[10] (PMC) was originally introduced as a compression scheme for 1-D signals. It has been further elaborated here and demonstrated for wavelet packet near-best basis transforms of 2-D signals. In particular, it has been shown to be a feasible method for the compression of images with minimal visual artifacts on reconstruction. As measured by the simple estimates calculated in Section 3, a significant improvement in compression was obtained for PMC relative to quantized-scalar coding (QSC) with a ratio of ≈ 14 for both orthogonal and biorthogonal test cases. As explained in Section 2.5, compression rates and improvement ratios estimated here reflect information coded from the wavelet packet amplitudes a and not the wavelet packet indices (l, b, c). Therefore, the comparisons reported here between QSC and the various versions of PMC must be considered valid only relative to each other within the context of this initial pilot study and not relative to some external standard requiring actual bit counts for information coded from both amplitudes and indices. Future research will focus on developing efficient methods for lossless coding of indices in conjunction with lossy coding of amplitudes. These future experiments will determine the range of compression rates (in terms of bits per pixel or other standard measures) within which parameterized-model coding is competitive with or superior to existing methods. However, it is likely that this range will be the very low bit rates most effectively used for coding a small number of packets with the largest amplitudes thereby minimizing the overhead requirements for coding the indices of the packets. This condition applies to the cases described in Section 1 of coding the M packets of a matching pursuit decomposition or the M largest packets of a complete basis decomposition.

5 REFERENCES

[1] Francois Bergeaud and Stephane Mallat. Matching pursuit of images. In Moeness G. Amin, editor, *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pages 330–333, Philadelphia, PA, October 1994. IEEE Press 94TH8007.

[2] Jonathan N. Bradley and Christopher M. Brislawn. Vector quantization of discrete wavelet transform coefficients. Technical Report LA-UR-92-3771, Los Alamos National Laboratory, Los Alamos, NM, 1992.

[3] Ronald R. Coifman and Mladen Victor Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2):713–718, March 1992.

[4] Ingrid Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41:909–996, 1988.

[5] Ronald A. DeVore, Björn Jawerth, and Bradley J. Lucier. Image compression through wavelet transform coding. *IEEE Transactions on Information Theory*, 38(2):719–746, March 1992.

[6] Allen Gersho and Robert M. Gray. *Vector Quantization and Signal Compression*. The Kluwer International Series in Engineering and Computer Science: Communications and Information Theory. Kluwer Academic Publishers, Boston, MA, 1992.

[7] Stephane G. Mallat and Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, December 1993.

[8] David A. Ratkowsky. *Handbook of Nonlinear Regression Models*, volume 107 of *Statistics: Textbooks and Monographs*. Marcel Dekker, Inc., New York, NY, 1990.

[9] Carl Taswell. Near-best basis selection algorithms with non-additive information cost functions. In Moeness G. Amin, editor, *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pages 13–16, Philadelphia, PA, October 1994. IEEE Press 94TH8007.

[10] Carl Taswell. Top-down and bottom-up tree search algorithms for selecting bases in wavelet packet transforms. In Anestis Antoniadis and Georges Oppenheim, editors, *Wavelets and Statistics*, Lecture Notes in Statistics. Springer Verlag, 1995. Proceedings of the Villard de Lans Conference November 1994.

[11] Carl Taswell. WavBox 4: A software toolbox for wavelet transforms and adaptive wavelet packet decompositions. In Anestis Antoniadis and Georges Oppenheim, editors, *Wavelets and Statistics*, Lecture Notes in Statistics. Springer Verlag, 1995. Proceedings of the Villard de Lans Conference November 1994.

Table 1: Compression Rate R and Distortion Rates D_p for Various Codings of Elaine Image

Coding	Orthogonal (DOLA 8)				Biorthogonal (BBBS 7,9)			
(Grouping)	R	D_1	D_2	D_∞	R	D_1	D_2	D_∞
PMC (none)	12	0.8760	2.4747	33.3320	11	1.3341	3.1428	50.3867
PMC (levels)	47	0.0563	0.1172	5.0625	51	0.0913	0.1543	7.2031
PMC (blocks)	236	0.0549	0.0675	1.1445	351	0.0541	0.0635	1.3203
PMC (order)	120	0.0476	0.0531	0.2188	142	0.0492	0.0555	0.2891
QSC (none)	17813	0.0451	0.0506	0.2227	18568	0.0471	0.0532	0.2969

Figure 1: Original and reconstructed versions of Elaine image following orthogonal near-best basis wavelet packet transform and lossy compression with parameterized-model coding: upper left – original; upper right – PMC with levels grouping; lower left – PMC with blocks grouping; lower right – PMC with order grouping.