A Complexity Constraint Best-Basis Wavelet Packet Algorithm for Image Compression[†]

Detlev Marpe[*], Hans L. Cycon[*] and Wu Li[‡]

[*] Fachhochschule für Technik und Wirtschaft
Allee der Kosmonauten 20–22
D-10315 Berlin, Germany
Email: dmarpe@fhtw-berlin.de
hcycon@fhtw-berlin.de

[‡] Department of Mathematics and Statistics
Old Dominion University
Norfolk, USA
Email: wuli@math.odu.edu

Abstract

The concept of adapted waveform analysis using a best basis selection out of a predefined library of wavelet packet (WP) bases allows an efficient representation of a signal. This method usually has the disadvantage of high computational complexity. In this paper we introduce an extension of the best-basis method, the complexity constrained best-basis algorithm (CCBB) which allows an adaptive approach, has relatively low complexity and is memory saving. Our CCBB algorithm generates iteratively a scarce library of WP bases by extending a given library according to the energy distribution of the WP representation of a given signal. This iteration process is terminated by using a complexity measure as a control parameter. Our experimental results for image coding applications show that the rate-distortion (RD) performance increases in distinguished and image dependent jumps as the complexity is increased. This enables us to find optima with relatively high RD performance and relatively low complexity for processing still images as well as video sequences.

[†]This work was supported by Deutsche Telekom AG, Technologiezentrum Darmstadt, Germany

1 Introduction

The rate-distortion performance of transform coding techniques for the purpose of image compression seems to have reached a saturation region for non-adaptive (standard) *discrete wavelet transformation* (DWT) based methods, *i.e.*, only small rate-distortion performance win can be expected by further fine tuning filters, quantisers, coders, etc. For some classes of images, however, adaptive methods relying on *discrete wavelet packet transformations* (DWPT) still show some significant coding gain [7]. The price to be paid for this gain is that this kind of adaptive image compression algorithms usually have much higher computational complexity than DWT-based methods. This might be a tolerable drawback for still image applications but it makes these methods less appealing for video image processing.

To overcome this problem we propose an adaptive approach with controllable complexity. It is a modified best-basis method [2] which allows one to restrict the complexity of the original best-basis algorithm by a suitable chosen complexity bound κ. We use κ to control the upper bound of the number of numerical operations used to generate the restricted library of *wavelet packet* (WP) bases admitted for the best-basis search.

This library of WP bases will be created by an iterative complexity constrained best-basis (CCBB) algorithm which decomposes subbands with relatively high energy first. Within this energy constrained library of bases we do a usual best-basis search to find the best wavelet packet transformation in the sense of minimizing a given information cost functional. Since the decomposition library is tailored to the given image, only a small increase in complexity (with respect to the complexity of DWT) allows the algorithm to find a "suboptimal" best basis which has a rate-distortion performance comparable to that of the unrestricted best-basis method.

In this paper, we are concerned only with the performance of the wavelet (packet) transform, that means, we do not consider the optimization of quantisers or coders. For our experiments, however, we used a complete transform/quantiser/coder scheme followed by their inverse operations. This enables us to measure real bitrate/distortion pairs. For details of the experimental setup see Section 3.1.

We present coding results for still images and video sequences. For a class of still images we discover distinguished jumps in the rate-distortion performance curve (cf. section 3.2). This is the class for which best-basis methods usually show some advantage. For video sequences we recover the same kind of jump in the rate-distortion performance curve, when we apply our method to frame differences (cf. section 3.3). Thus by choosing the complexity parameter suitable, we can find optimal trade-off points where we have relatively low complexity and high rate-distortion performance. This enables us to take the advantage of adaptive methods also for processing video sequences in real time.

In contrast to our proposal, Ramchandran and Vetterli [7] introduced a variation of the best-basis scheme that is based on a joint rate-distortion optimization of WP transformation and quantization. This scheme however has much higher computational cost than the original best-basis algorithm of Coif-

man and Wickerhauser [2].

Taswell [8] proposed a top-down algorithm using a non-additive information cost functional to select so-called near-best bases. Our approach is similar to Taswell's algorithm in the sense of getting a near-best basis with relatively low computational cost. However, our algorithm can be explicitly controlled by a complexity parameter. Such control can be very useful in processing video images in real time.

The paper is organized as follows: Section 2.1 and 2.2 provide a brief introduction of wavelet packets and best-basis methods. In Sections 2.3 and 2.4, we introduce the definitions of relative energy and relative complexity which serve as tools to generate the image adapted κ-restricted library of WP bases and to select the best basis in this library with the CCBB algorithm, as formally described in Section 2.5. Section 3.1 contains the description of the experimental setup for all image coding simulations presented in Sections 3.2 and 3.3.

2 Methods

2.1 Wavelet Packet Transformations

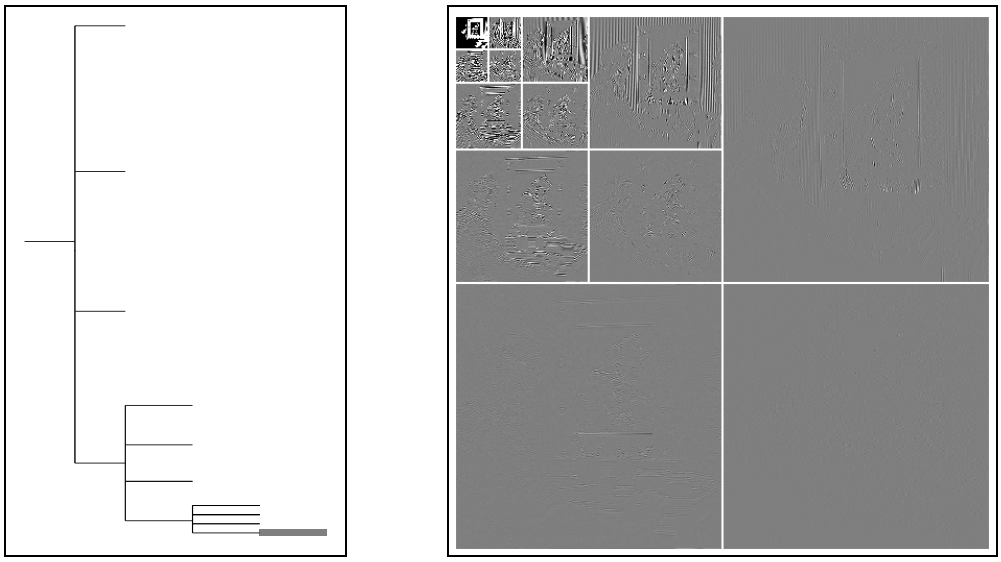
Given an image $\boldsymbol{x} \in \ell^2(\boldsymbol{\mathcal{Z}}^2)$ of size $N = N_1 \cdot N_2$, a decomposition of $\boldsymbol{x}$ into subbands of different frequency localization is performed by repeated application of a pair of *quadrature mirror filters* (QMF) (H, G). If H and G denote a low-pass and high-pass QMF, respectively, acting as convolution-decimation operators on $\ell^2(\boldsymbol{\mathcal{Z}})$, a quadruple $\boldsymbol{F}$ of 2-dimensional separable filters is defined by tensor products of the QMF pair (H, G):

$$\boldsymbol{F} = (F_0, F_1, F_2, F_3) \stackrel{\text{def}}{=} (H \otimes H, H \otimes G, G \otimes H, G \otimes G).$$

By this construction, the first step of a subband decomposition transforms the given image $\boldsymbol{x}$ into a representation having components in 4 subbands $W_{1,0}, W_{1,1}, W_{1,2}$ and $W_{1,3}$. These four bands are the images of $\ell^2(\boldsymbol{\mathcal{Z}}^2)$ under the projection operators related to F_0, F_1, F_2 and F_3, respectively. Iterating the filtering process with the application of $\boldsymbol{F}$ to $\boldsymbol{x}$ yields a quadtree structured decomposition of $\boldsymbol{x}$. The collection of all subband decompositions of $\boldsymbol{x}$ with a maximum number of L levels forms a full, balanced quadtree of depth L. At each node (l, k) of this balanced quadtree, the original image $\boldsymbol{x}$ is represented by a component in the subband $W_{l,k}$ consisting of $4^{-l} \cdot N_1 \cdot N_2$ coefficients $(0 \leq k < 4^l, 0 \leq l \leq L)$.

Given this full quadtree, we are faced with the problem of choosing a specific subband decomposition related to a particular partition of the frequency axis. Note that each subband decomposition corresponds to a basis B in the Hilbert space $\mathcal{H} = \ell^2(\boldsymbol{\mathcal{Z}}^2)$, called *wavelet packet* (WP) basis[1]. For example, the standard wavelet basis which is related to an octave-band tree decomposition is such a particular WP basis (cf. fig. 1).

[1]For mathematical convenience, we use the notation B also for the discrete wavelet packet transformation corresponding to the basis B.



(a) (b)

Figure 1: (a) Octave-band tree of depth $L = 4$ and (b) visualization of corresponding DWT of CLOWN test image (see fig. 7 for original image)

If we assume the QMF pair (H, G) to be orthogonal, then the basis B and the corresponding transformation is also orthogonal. Thus, the wavelet packet analysis provides us with a full depth-L library of orthogonal bases (ONB), denoted by $\boldsymbol{\mathcal{B}}_L$, which is a subset of all ONB of $\mathcal{H}$. The cardinal number A_L of $\boldsymbol{\mathcal{B}}_L$ can be estimated with the help of a recursive relation between the number of subtrees in a full quadtree resulting in the estimation $A_L \geq 2^{4^{L-1}}$.

2.2 Review of the Best-Basis Algorithm

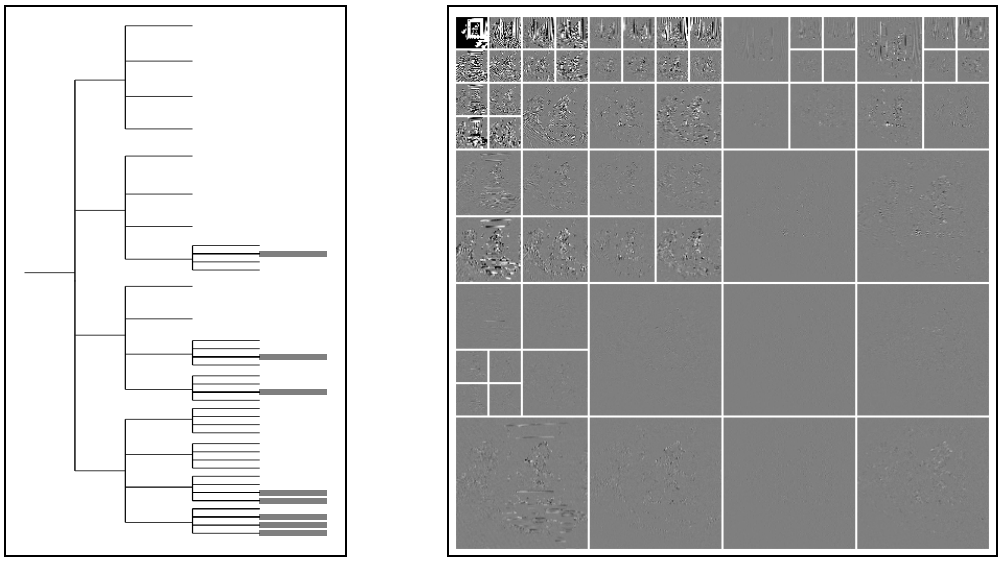
Given a library of WP bases, the search for an efficient representation (in terms of compression) demands for a measure of the information content of a given representation. In order to make this search feasible in a library of e.g. depth $L = 4$ with $A_4 \geq 1.8 \times 10^{19}$ different members we have to restrict to those information measures which exploit the parent-child relationship in the tree, thereby allowing an efficient search method.

An additive functional $M : \ell^2(\boldsymbol{\mathcal{Z}}^2) \rightarrow \boldsymbol{\mathcal{R}}$ is a so-called *information cost functional* with the property, that the information cost of the four children nodes of a quadtree parent node equals the sum of the costs of each of the children. The *best basis* for a given image $\boldsymbol{x}$ relative to any library $\boldsymbol{\mathcal{B}}$ will be a basis B in the library of bases $\boldsymbol{\mathcal{B}}$ for which $M(B\boldsymbol{x})$ is minimal.

Let $B_{l,k}$ the set of WP basis vectors belonging to the subband $W_{l,k}$. Using the parent-child relationship between the subbands

$$W_{l-1,k} = W_{l,4k} \oplus W_{l,4k+1} \oplus W_{l,4k+2} \oplus W_{l,4k+3},$$

which is a consequence of the orthogonality condition introduced in Section 2.1, we can give a formal description of the best-basis algorithm [2]:



(a) (b)

Figure 2: (a) Best-basis WP tree ($L = 4$) and (b) visualization of corresponding DWPT of CLOWN image

Algorithm 1 (Best-Basis Algorithm) *Given an image $\boldsymbol{x}$, the best basis B for $\boldsymbol{x}$ relative to the library $\boldsymbol{\mathcal{B}}_L$ and to the cost functional M is obtained by performing the following steps:*

Step 0: *Expand x in the library $\boldsymbol{\mathcal{B}}_L$ and compute the costs $I_{l,k} = M(B_{l,k}\boldsymbol{x})$ for $0 \le k < 4^l$ and $0 \le l \le L$.*

Step 1: *Choose the basis B related to the subbands at the bottom level $l = L$, i.e., set $B = \{B_{L,0}, \ldots, B_{L,4^L-1}\}$.*

Step 2: *Compare the cost of a parent node to the sum of the costs of 4 children nodes. If the sum of the costs of children is greater than or equal to the cost of the parent, prune the part of the tree below the parent node; otherwise assign the sum of the children's cost to the parent node, i.e., for $k = 0, \ldots, 4^{l-1} - 1$*

if $I_{l-1,k} \le \sum_{j=0}^{3} I_{l,4k+j}$

then *replace* $\{B_{i,j} | \, i = l, \ldots, L;\ j = 4^{i-l+1}k, \ldots, 4^{i-l+1}(k+1) - 1\} \cap B$ *by* $B_{l-1,k}$ *in* B

else *set* $I_{l-1,k} = \sum_{j=0}^{3} I_{l,4k+j}$

Step 3: *Repeat Step 2 for the next lower level until the root is reached, i.e.,*

while $l > 1$

do *set* $l = l - 1$*;*
repeat Step 2.

A useful cost functional for the purpose of transform coding is the pseudo-entropy information cost introduced in [2], *i.e.*, $M_H(u) = -\sum_j |u_j|^2 \log |u_j|^2$, which is an additive analogue to the Shannon-Weaver entropy in the sense that minimizing the former implies minimizing the latter. It has proved to be successful in image compression schemes [5]. This will be the cost functional that we use throughout this paper. (See fig. 2 for a best-basis WP representation of a test image and its corresponding tree.)

2.3 Relative Energy

The energy of an image $\boldsymbol{x}$ given by $E(\boldsymbol{x}) = \sum_{i,j} |\boldsymbol{x}_{i,j}|^2$ is usually very inhomogeneously distributed among the subbands: only a few subbands have high *relative energy*, while others contain almost no significant information of the image. Therefore it is reasonable to control the generation of a library of WP bases in a "top-down" strategy by decomposing the subbands with respect to some energy measure.

The so-called *relative energy* $e_{l,k}$ of $\boldsymbol{x}$ in a given subband $W_{l,k}$ at level l is defined as the relative part of the energy of $\boldsymbol{x}$ in the given child subband $W_{l,k}$ with respect to the energy of $\boldsymbol{x}$ in the corresponding parent subband $W_{l-1,\lfloor \frac{k}{4} \rfloor}$ (where $\lfloor y \rfloor$ denotes the greatest integer less than or equal to y):

$$e_{l,k} \stackrel{\text{def}}{=} \frac{E(B_{l,k}\boldsymbol{x})}{E(B_{l-1,\lfloor \frac{k}{4} \rfloor}\boldsymbol{x})} \qquad (1 \le l \le L;\ 0 \le k < 4^l).$$

It is a measure of energy compaction in the parent-child relationship of subbands. Note that the orthogonality condition implies the additivity of the energy functional $E(\cdot)$, hence the relative energies of 4 children in a subband decomposition sum up to 1.

Usually most of the energy is located in the low-pass branch, *i.e.*, the subbands generated by the DWT will have the highest relative energy. But for some images with a mismatched space-frequency characteristic, there are high-frequency subbands which have high relative energy, too. For this kind of images, a WP basis different from the one generated by the DWT will be a much better representation (as regards to compression applications).

The problem of finding such a representation can be solved in a suboptimal way by constructing an energy constrained library $\boldsymbol{\mathcal{B}}_L^{\varepsilon}(\boldsymbol{x})$ using a (predefined) so-called *energy-threshold* ε to discriminate branches (nodes) with less relative energy than ε. (See [6] for details.)

However, instead of a simple thresholding, it is more efficient to use the measure of relative energy for ordering the nodes of a tree of WP coefficients. We may sort the terminal nodes of a given tree in a list by decreasing order of relative energy, so that at each step of the generation of the library we will decompose the subband related to the topmost node in the list. Given a complexity bound κ, we will stop further decomposition of subbands if the total complexity which is proportional to the sum of the samples in all subbands reaches the upper bound κ. Choosing κ suitably, we can parameterize a family of complexity constrained suboptimal best basis transformations ranging from the fast DWT to the full best basis DWPT.

This idea leads to an algorithm for generating an adapted sublibrary of bases within a given complexity bound. Before we present the details of this so-called *complexity-constrained best-basis algorithm* (cf. Section 2.5), we define an appropriate measure of complexity[2] and derive some related estimates.

2.4 Complexity Measures and Estimates

Suppose that we start our wavelet packet analysis with an image $\boldsymbol{x}$ of size $N = N_1 \cdot N_2$ (*i.e.*, $\boldsymbol{x}$ has "height" N_1 and "width" N_2). Then the maximum level L is restricted by $L \leq \min(\log_2 N_1, \log_2 N_2)$.

For each single step of decomposition at level l, the *computational cost* $\mathcal{C}_l$ for calculating the wavelet packet coefficients is proportional to the number of samples in the corresponding subband, *i.e.*,

$$\mathcal{C}_l \leq C \cdot 4^{-l} N \qquad (0 \leq l \leq L). \tag{1}$$

The constant C in (1) is independent of the level and is determined by the length of the QMF pair (H, G). If, for example, the length of the QMF is M, then C can be estimated by $C \leq 2 \cdot (2M - 1)$ operations (M multiplications and $(M - 1)$ additions for each convolution operation).

For an estimation of the complexity $\mathcal{C}(\mathrm{DWPT})$ of a depth-L DWPT of $\boldsymbol{x}$, we have to count the number p_l of different subbands (nodes) at level l each having computational cost $\mathcal{C}_l$. This adds up to

$$\mathcal{C}(\mathrm{DWPT}) = \sum_{\mathrm{l}=0}^{\mathrm{L}-1} \mathrm{p_l} \cdot \mathcal{C}_\mathrm{l}, \tag{2}$$

where p_l ranges between 1 and 4^l. Thus, any DWPT with one subband per level has the minimum complexity. One such particular DWPT is the DWT. Using (1) and (2), we get

$$\mathcal{C}(\mathrm{DWT}) \sim \frac{4}{3} \cdot (1 - 4^{-L}) \cdot N. \tag{3}$$

The maximum complexity is given by the DWPT corresponding to the full depth-L quadtree, which is identical to the complexity of the construction of the full library $\boldsymbol{\mathcal{B}}_L$. Its computational cost is proportional to the product of L and N, *i.e.*, $\mathcal{C}(\mathrm{full-DWPT}) \sim \mathrm{L} \cdot \mathrm{N}$.

In order to define a complexity measure on the library $\boldsymbol{\mathcal{B}}_L$, which is independent of the signal and the QMF, we introduce the *relative complexity* (which is the relative computational cost) c_l of a subband decomposition at level l with respect to the computational cost of the DWT of depth-L. Combining (1) and (3), we get

$$c_l \stackrel{\mathrm{def}}{=} \frac{\mathcal{C}_l}{\mathcal{C}(\mathrm{DWT})} = \frac{3 \cdot 4^{-l}}{4(1 - 4^{-L})} \qquad (0 \leq l \leq L). \tag{4}$$

The relative complexity c_l gives the amount of complexity in units of $\mathcal{C}(\mathrm{DWT})$ for a subband decomposition at level l. Note that the relative computational

[2] We simply use the notion of complexity with the semantics of computational complexity.

cost of the construction of the whole library $\boldsymbol{\mathcal{B}}_L$ of wavelet packet bases is given by $c_{\max} = \frac{3}{4} \cdot \frac{L}{(1-4^{-L})}$, which e.g. for a maximum decomposition level $L = 4$ means that $\mathcal{C}(\text{full} - \text{DWPT}) \approx 3 \cdot \mathcal{C}(\text{DWT})$.

So far we have only estimated the operations necessary for generating a tree of WP coefficients. Calculating information costs will contribute some additional operations per coefficient, so that the constant C in (1) has to be modified. But this will have no influence on the relative measure c_l, even though the "complexity unit" $\mathcal{C}(\text{DWT})$ will account for an information cost enriched standard DWT.

The additional complexity of the best basis search in Steps 1–3 of Algorithm 1 is proportional to the number of (internal) nodes of the tree corresponding to the library $\boldsymbol{\mathcal{B}}_L$, given by $\sum_{l=1}^{L-1} 4^l$. Adding a constant to the right hand side of (1) adjusts the related complexity estimates for a best-basis wavelet packet transformation.

2.5 The Complexity Constrained Best-Basis Algorithm

The concepts we have developed so far will be merged into the design of a complexity scalable variant of the best-basis algorithm. With the measure of relative computational cost we can parameterize the sublibraries of the full WP library $\boldsymbol{\mathcal{B}}_L$ in discrete steps of c_l $(0 \leq l \leq L)$. Given an image $\boldsymbol{x}$, we will call the corresponding complexity parameter the *complexity bound* κ and the related sublibrary of $\boldsymbol{\mathcal{B}}_L$ the *κ-restricted library* $\boldsymbol{\mathcal{B}}_L^\kappa(\boldsymbol{x})$ of $\boldsymbol{x}$. The range of the complexity bound κ is restricted to the interval $[0, c_{\max}]$ according to the discussion in the last section. Clearly, the construction of a κ-restricted library $\boldsymbol{\mathcal{B}}_L^\kappa(\boldsymbol{x})$ will depend on a "top-down" strategy which in essence is a decomposition rule, as outlined in sec. 2.3. Hence $\boldsymbol{\mathcal{B}}_L^\kappa(\boldsymbol{x})$ depends implicitly on $\boldsymbol{x}$ by means of a measure operating on the expansion of $\boldsymbol{x}$ in $\boldsymbol{\mathcal{B}}_L^\kappa(\boldsymbol{x})$.

A suitable measure for our purposes seems to be the measure of relative energy. With the help of this energy-compaction measure we are able to establish an order on each sublibrary by evaluating the projections of $\boldsymbol{x}$ onto each subband. Ordering the nodes of a tree (related to a sublibrary $\boldsymbol{\mathcal{B}}$) in such a way allows a unique extension of $\boldsymbol{\mathcal{B}}$ by decomposing the terminal node with highest relative energy. This rule of decomposition together with an appropriate chosen complexity bound gives rise to the following algorithm:

Algorithm 2 (Complexity Constrained Best-Basis Algorithm) *Let $\boldsymbol{x}$ be an image and $\kappa \in [0, c_{\max}]$ a complexity bound. Then the best basis B for $\boldsymbol{x}$ relative to the κ-restricted library $\boldsymbol{\mathcal{B}}_L^\kappa(\boldsymbol{x})$ and the information cost functional M is selected as follows:*

Step 0: *Set the initial library $\boldsymbol{\mathcal{B}}_L^\kappa(\boldsymbol{x}) = \{B_{0,0}\}$, initial best basis $B = \{B_{0,0}\}$ and initial value of total complexity $c_{\text{tot}} = 0$. Start with root node $(l, k) = (0, 0)$ and set $I_{0,0} = M(B_{0,0}\boldsymbol{x})$.*

Step 1: *Create a new generation of 4 nodes and replace the parent node by the union of the children nodes in the initial configuration of best basis B, i.e.,*

add $\{B_{l+1,4k}, \ldots, B_{l+1,4k+3}\}$ *to* $\mathcal{B}_L^\kappa(\boldsymbol{x})$;
replace $B_{l,k}$ *by* $\{B_{l+1,4k}, \ldots, B_{l+1,4k+3}\}$ *in* B;
set $I_{l+1,4k+j} = M(B_{l+1,4k+j}\boldsymbol{x})$ *for* $j = 0, \ldots, 3$;
set $e_{l+1,4k+j} = \frac{E(B_{l+1,4k+j}\boldsymbol{x})}{E(B_{l,k}\boldsymbol{x})}$ *for* $j = 0, \ldots, 3$;
set $c_{\mathrm{tot}} = c_{\mathrm{tot}} + c_l$;

Step 2: *Repeat Step 1 for the terminal node with the largest relative energy until complexity budget is exhausted, i.e.,*

while $c_{\mathrm{tot}} < \kappa$

do *determine* (l, k) *with*
$e_{l,k} = \max\{e_{j,k} | \, B_{j,k} \in B, 1 \le j < L, 0 \le k < 4^j\}$;
repeat Step 1 for node (l, k);

Step 3: *Given the initial choice* B *of best basis, the search for best basis in* $\mathcal{B}_L^\kappa(\boldsymbol{x})$ *is performed by recursively traversing the tree in depth-first order from root* $(l, k) = (0, 0)$ *down to the terminal nodes corresponding to the members of initial* B, *i.e.,*

if $B_{l,k} \notin B$

then *repeat Step 3 for the 4 children nodes* $(l+1, 4k), \ldots, (l+1, 4k+3)$;

if $I_{l,k} \le \sum_{j=0}^{3} I_{l+1,4k+j}$

then *replace* $\{B_{i,j} | \, i = l+1, \ldots, L; j = 4^{i-l}k, \ldots, 4^{i-l}(k+1)-1\} \cap B$ *by* $B_{l,k}$ *in* B

else *set* $I_{l,k} = \sum_{j=0}^{3} I_{l+1,4k+j}$

Remarks:

1. The decomposing control structure in Step 2 has access to a "virtual" list of terminal nodes via the initial configuration of best basis B which is updated in each decomposing process of Step 1 by replacing basis vectors of the parent node with the collection of those of the 4 children nodes. Note that for determination of the terminal node with largest relative energy the terminal nodes at deepest level L will not be in competition.

2. Step 3 gives an alternative algorithmic realization of the best-basis search. Here the tree representing the library $\mathcal{B}_L^\kappa(\boldsymbol{x})$ is traversed in depth-first order by recursively performing Step 3 until a leaf node is reached. By resolving the recursion the bottom-up process of selecting an information cost minimizing representation is performed. (cf. steps 1–3 of Algorithm 1, which is a breadth-first oriented description of the best-basis selection.)

3 Results

3.1 Experimental Setup

For our measurements we used a complete transform/quantiser/coder scheme followed by their inverse operations. Since we are primarily concerned with the performance of different WP bases in the transformation part of the generic coding scheme, we do not discuss here the impact of different filters, quantisers or coders. Moreover, we do not vary these components of the coding scheme in our experiments.

In particular, these components are chosen as follows. The quantiser performs a simple scalar quantization of the wavelet packet coefficients with a bandwise adjusted deadzone and uniform step-width. The coder is in fact an entropy coding module for generating variable length code for the quantiser indices and is realized by an adaptive arithmetic coder. Note that the resulting bitstream includes as well the side information of the WP basis description (tree structure).

The chosen wavelet in our experiments corresponds to an orthogonal, "minimum phase" filter with 8 nonvanishing taps and a certain "steepness" around the transition region (cf. [3], table 6.4). This filter was chosen since the corresponding characteristics of the related QMF pair, such as the "approximate" symmetry, short length, and "closeness" to the ideal filter (Shannon wavelet), indicate that it might be a good candidate for best-basis image compression applications. Especially the latter property of a good frequency localization is very important in an iterated wavelet packet decomposition. The "aliasing" part of frequency due to an imperfect frequency localization will spread away from the central transition region in subsequent iterations (see [4] for a discussion of this problem). We implemented this filter with a periodic boundary extension and a maximum decomposition level $L = 4$.

The criteria of distortion we used in our experiments is determined by the ℓ^2-norm of the reconstruction error $D = \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|_2$ between the original image $\boldsymbol{x}$ and a reconstruction $\hat{\boldsymbol{x}}$. We use as a measure of distortion the *peak signal-to-noise-ratio* (PSNR) defined as $10 \log_{10}(255^2/\mathrm{MSE})$, where the *mean square error* (MSE) is given by D^2/N for an image $\boldsymbol{x}$ of size N.

3.2 Still Images

Complexity is not a major concern in still image processing. So that for some classes of images with special space-frequency characteristics it is reasonable to use more complex adaptive methods like best-basis algorithms or matching pursuit methods [1]. Here we use this example to illustrate that the best rate-distortion performance can be achieved by using the complexity-constrained best-basis method with much lower computational cost than the original best-basis method.

Our still image experiments were performed on the grayscale test image CLOWN with 512×512 pixels which shows a critical periodic high-frequency structure (see Fig. 7). For a set of complexity bounds we apply our WP-based CCBB algorithm to the CLOWN image. Figure 3 shows the result of

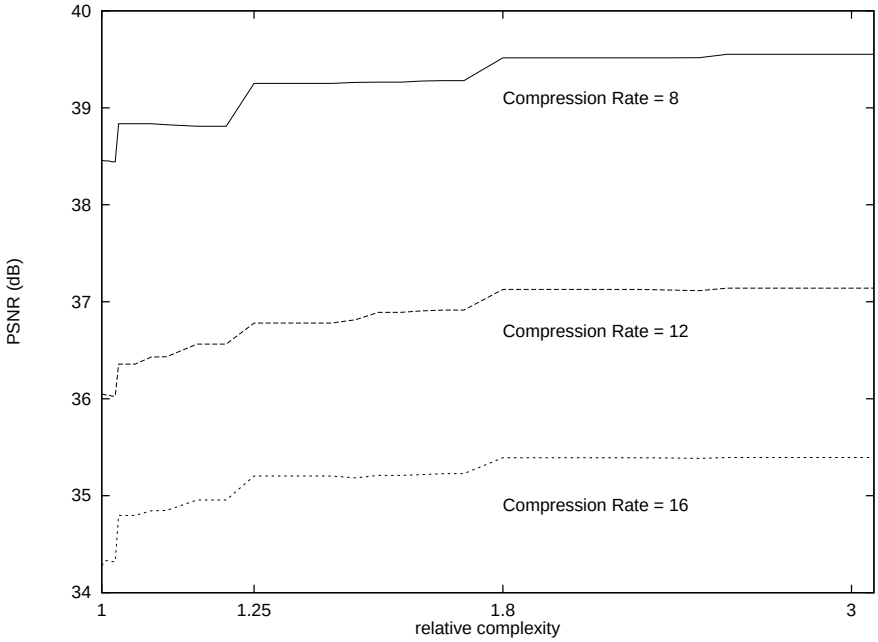


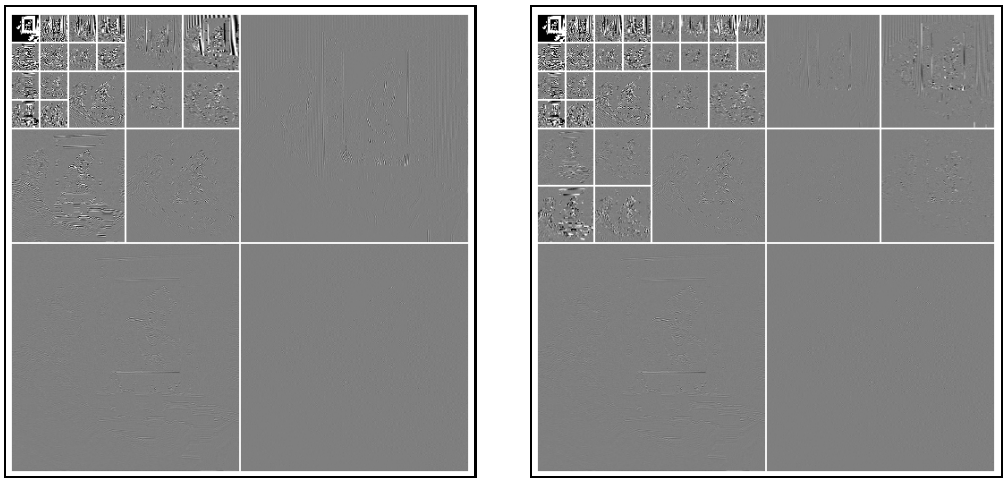
Figure 3: Distortion vs. complexity bound with respect to fixed compression ratios for CLOWN image

this experiment for three different target bitrates, *i.e.*, compression ratios. The plotted graphs display for each fixed compression ratio complexity bound κ (in terms of relative complexity) versus distortion (in terms of PSNR). The complexity scale reaches from the DWT ($\kappa = 1$) to the unrestricted best-basis ($\kappa \approx 3$).

One recognizes several steps in PSNR increase which are due to the iteratively included subbands in the constrained library. There is an obvious first optimum sitting on the edge of the first step with 0.4 dB increase in PSNR but less than 3% increase in complexity. (Note that the scale of complexity is logarithmic.) Fig. 4(a) displays the best basis for this first optimum.

For low compression ratio (1:8) there is a distinguished second step with a gain of 0.4 dB PSNR and an increase of complexity of approximately 25%. The whole coding gain (in PSNR) possible for this compression ratio is 1 dB, so that this second optimum on the edge of the second step contributes 80% of the whole gain with only 25% complexity increase. A comparison of the corresponding best basis shown in fig. 4(b) with the one of the previous step (a) reveals the reason for this jump. The decomposition of the upper right high frequency band in fig. 4(b) contributes a gain of 0.5 dB in PSNR over the first step represented in the subband decomposition (a). Nearly full PSNR gain is achieved at a third step with a complexity increase of 80% (relative to the DWT) which is less than half the maximum complexity (of the full best-basis DWPT). See fig. 2 (b) for a comparison with the unrestricted best basis of the CLOWN image, which corresponds to complexity bound $\kappa = c_{\max} = 3 \cdot \frac{256}{255} \approx 3.01$.

For higher compression ratios we have similar results although the distinguished steps split into smaller steps. This phenomenon is due to the quantiser



(a) (b)

Figure 4: (a) Best basis with complexity bound $\kappa = 1.025$ and (b) best basis with $\kappa = 1.25$ for CLOWN image

which has step sizes that halve with successively deeper tree levels and which has a common step-width at root. Choosing higher compression ratios (*i.e.*, enlarging the step-width of the quantiser) results in dominant significance of subbands on lower levels, since more and more high frequency information (on upper levels) is annihilated.

3.3 Video Sequences

Motion compensated frame differences (MCFD) and frame differences (FD) of video sequences have different statistics than still images. Different classes of sequences show different RD performance when adaptive methods are used. For sequences with a complex motion structure, adaptive methods usually show a superior performance, which is essentially due to the high frequency structure present in the insufficient motion compensated FD or MCFD. Figure 8 displays frame #0 and the first FD frame of the COASTGUARD sequence. This sequence has a frame rate of 25 Hz and a resolution of 352×288 pixels. The FD frame in figure 8 (b) shows the typical high frequency structure as mentioned above.

Our experiment was performed by processing 74 temporally subsampled FD frames out of the first 148 FD frames of the COASTGUARD sequence with our CCBB algorithm. Figure 6 shows that the unrestricted best-basis method achieves approximately 1 dB better quality (in terms of PSNR) on the average than the standard DWT for the same compression ratio. However, this performance win has the penalty of high complexity increase. But as in the still image case, the performance win goes stepwise when complexity is increased, since the algorithm includes stepwise more and more subbands. So we discover "optimum edges" located at the top of the "edge" in the 3D graph of figure 5.

For fixed compression ratio (1:10), we depict in fig. 5 reconstruction quality (in terms of PSNR) versus complexity (in terms of relative complexity) and FD

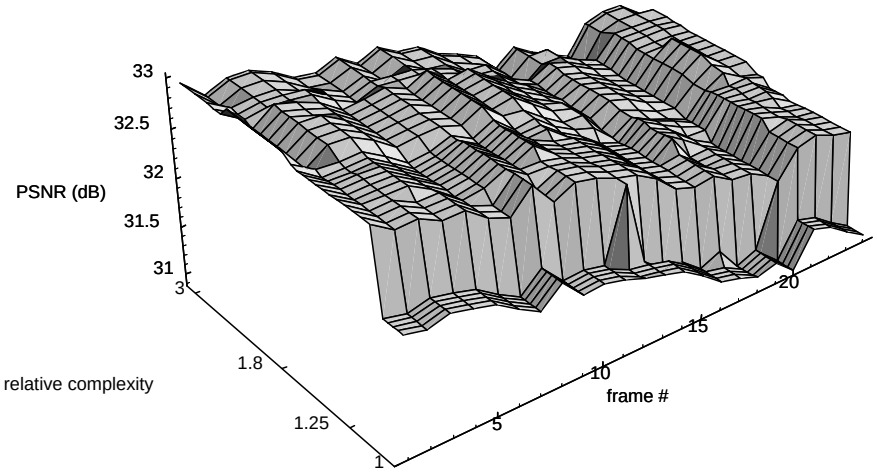


Figure 5: Distortion vs. complexity vs. frame for 25 FD frames of COASTGUARD sequence with a compression ratio 1:10

frame number on a logarithmic complexity scale for the first 25 FD frames. For almost all FD frames the figure shows a distinguished edge of approximately 0.8 dB gain in PSNR with only 25% increase in complexity.

An "integrated" version of these results is shown in fig. 6 where averaged distortion (of 74 FD frames) versus complexity bound is plotted. Due to the integration the jumps around 1.25 relative complexity shown in fig. 5 form a smooth but distinguished area of rapid increase in quality. There is a second area at 1.8 relative complexity with a smaller increase in PSNR. The performance at higher compression ratios is similar but less pronounced.

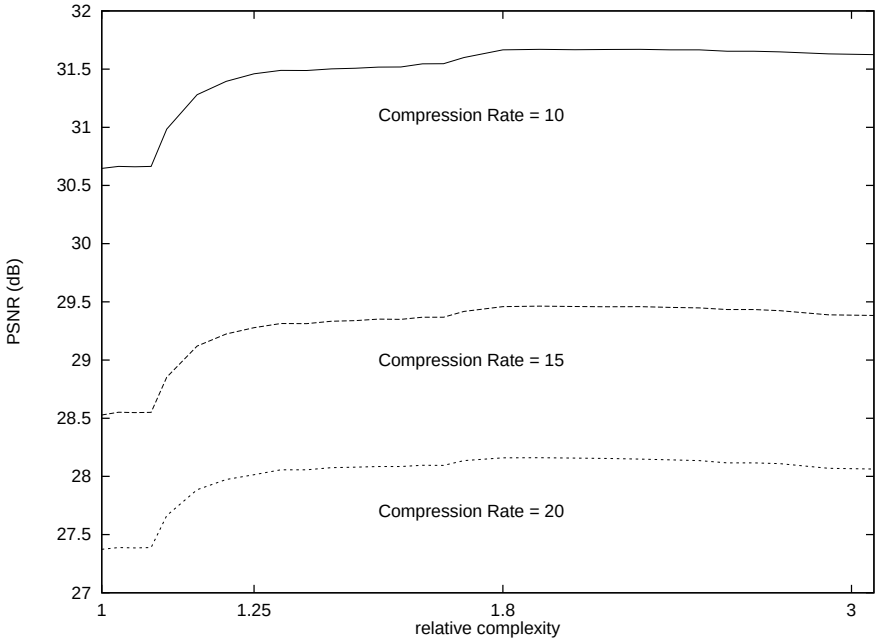


Figure 6: Averaged PSNR vs. relative complexity for COASTGUARD sequence

4 Conclusions

In this paper we introduced an algorithm for image compression which parameterizes the complexity gap between the standard discrete wavelet transformation having low computational complexity and relatively low RD performance, on the one hand, and on the other hand, the best-basis wavelet packet algorithm with relatively high complexity and high RD performance. For some classes of still images and frame differences of video sequences the RD performance shows some distinguished jumps when complexity is increased. Thus we can find suboptimal wavelet packet transformations having relatively low complexity and relatively high RD performance by sitting on the edges of one of these jumps. This may be of special interest for real-time video applications since our complexity constraint best-basis algorithm allows the generation of a fast transformation which outperforms the RD performance of the standard DWT for complex video sequences.

References

[1] BERGEAUD, F. and MALLAT, S., 'Matching pursuit of images', *SPIE Vol.* **2491** (1994), pp. 2–12.

[2] COIFMAN, R. R. and WICKERHAUSER, M. V., 'Entropy-based algorithms for best basis selection', *IEEE Trans. on Information Theory* **38** (2) (1992), pp. 713–718.

[3] DAUBECHIES, I., 'Ten Lectures on Wavelets', No. 61 in CBMS-NSF Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[4] HESS-NIELSEN, N. and WICKERHAUSER, M. V., 'Wavelets and time-frequency analysis', *Proceedings of the IEEE* **84** (4) (1996).

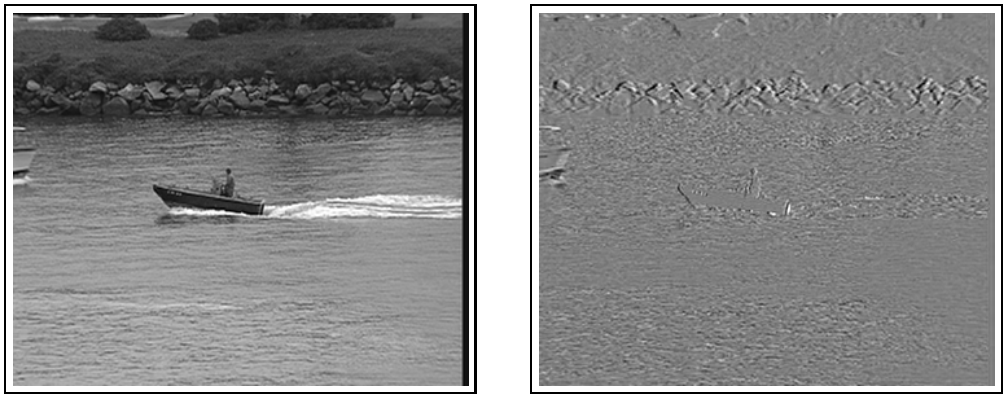
[5] MARPE, D., 'Best bases experiments I and II', Technical Reports, Wavelet Group, Technical University Berlin, 1995.

[6] MARPE, D., CYCON H. L. and LI, W., 'Energy constrained scarce wavelet packet libraries for image compression', Preprint No. 541, Technical University Berlin, 1996.

[7] RAMCHANDRAN, K. and VETTERLI, M., 'Best wavelet packet bases in a rate-distortion sense', *IEEE Trans. on Image Proc.* **2** (2) (1993), pp. 160–175.

[8] TASWELL, C., 'Image compression by parameterized-model coding of wavelet packet near-best bases', Preprint 1996.

Figure 7: CLOWN image with 512×512 pixels and 8 bits/pixel



(a) (b)

Figure 8: (a) Frame #0 and (b) FD frame #1 of COASTGUARD sequence

List of Figures

1 (a) Octave-band tree of depth $L = 4$ and (b) visualization of corresponding DWT of CLOWN test image (see fig. 7 for original image) 4
2 (a) Best-basis WP tree ($L = 4$) and (b) visualization of corresponding DWPT of CLOWN image 5
3 Distortion vs. complexity bound with respect to fixed compression ratios for CLOWN image 11
4 (a) Best basis with complexity bound $\kappa = 1.025$ and (b) best basis with $\kappa = 1.25$ for CLOWN image 12
5 Distortion vs. complexity vs. frame for 25 FD frames of COASTGUARD sequence with a compression ratio 1:10 13
6 Averaged PSNR vs. relative complexity for COASTGUARD sequence 13
7 CLOWN image with 512×512 pixels and 8 bits/pixel 15
8 (a) Frame #0 and (b) FD frame #1 of COASTGUARD sequence . 15