Proceedings of the 6th Nordic Signal Processing Symposium - NORSIG 2004
June 9 - 11, 2004
Espoo, Finland

WAVELET-PACKET SUBBAND STRUCTURES IN THE EVOLUTION OF THE JPEG 2000 STANDARD

Markus Reisecker and Andreas Uhl

Department of Scientific Computing, Salzburg University, AUSTRIA
e-mail: uhl@cosy.sbg.ac.at

ABSTRACT

We evaluate the suggested employment of wavelet packet subband structures in the different stages of the JPEG 2000 standardization process and find the solution finally adopted for JPEG 2000 Part II to be satisfactory for amost all type of images.

1. INTRODUCTION

Wavelet packet subband structures [9] generalize the pyramidal decomposition structure as used by most classical wavelet-based coding schemes by iterating the decompositions on the high pass subbands as well. Especially for images with highly textured content (compare for example the FBI fingerprint compression standard [1]) rate-distortion performance of adaptively generated wavelet packet subband structures is superior to classical pyramidal ones.

Best basis algorithms [2] are the computationally most efficient techniques to find wavelet packet subband structures for a given image. This is done by optimizing rate independent information cost functions operating in the transform domain which provides suboptimal rate-distortion performance only. Other techniques involving rate-distortion based optimization have been introduced [5, 6], often at a high computational cost.

Whereas the first generation wavelet coding schemes based on (possibly adaptive) uniform quantization and subsequent entropy coding could be easily generalized to the wavelet packet case, the extension of zerotree-based schemes like SPIHT or similar have been more tricky to generalize, especially concerning the parent-child relationship of coefficients across subbands. However, these problems have been also solved in the last years [4]. EBCOT and JPEG 2000 do not rely on zerotree based coding and can therefore be extended much easier to the wavelet packet case.

In this work we describe and experimentally evaluate the performance of the wavelet packet structures admitted in the different stages of the JPEG 2000 standardization process and try to answer the question whether the scheme finally adopted for JPEG 2000 Part II is sufficient from the rate-distortion viewpoint.

This work has been partially supported by the Austrian Science Fund (project FWF-13732).

2. JPEG 2000

The JPEG 2000 image coding standard [7] uses a scheme originally proposed by Taubman and known as EBCOT ("Embedded Block Coding with Optimized Truncation" [7]). JPEG 2000 is based on the wavelet transform, the major difference between previously proposed wavelet-based image compression algorithms such as EZW or SPIHT [8] is that EBCOT as well as JPEG 2000 operate on independent, non-overlapping blocks of quantized wavelet coefficients which are coded in several bit layers to create an embedded, scalable bitstream (Tier-1 coding). Instead of zerotrees, the JPEG 2000 scheme depends on a per-block quad-tree structure since the strictly independent block coding strategy precludes structures across subbands or even code-blocks. These independent code-blocks are passed down the "coding pipeline" and generate separate bitstreams. Transmitting each bit layer corresponds to a certain distortion level. The partitioning of the available bit budget between the code-blocks and layers ("truncation points") is determined using a sophisticated optimization strategy for optimal rate/distortion performance (Tier-2 coding).

It is well known that JPEG 2000 clearly outperforms JPEG with respect to rate/distortion performance, especially at lower bitrates. Note that for all subsequent experiments JPEG 2000 VM6.0 is used which has been extended to support arbitrary isotropic subband structures.

3. WAVELET PACKET SUBBAND STRUCTURES IN JPEG 2000

During the standardization process of JPEG 2000 Part I, wavelet packet subband structures have been considered but have finally not been included. This technology is now contained in Part II. The software verification model 6.0 (VM6.0) reflects the discussion dur-

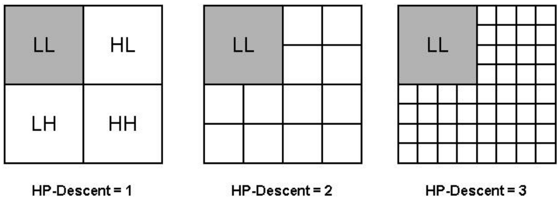

Figure 1: Definition of a resolution level's subband structure with the HP-descent parameter in VM6.0

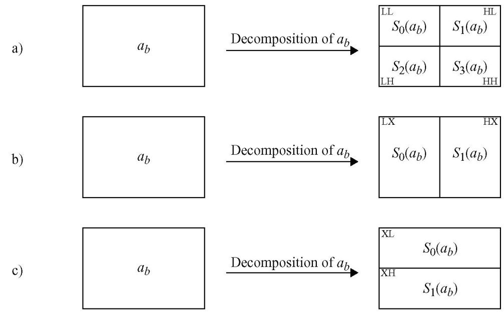

Figure 2: Possible division of a resolution or a subband a_b in JPEG2000 Part II.

ing the standardization process of Part I and consequently posesses some limited possibilities to define the subband structure of the wavelet decomposition. The HP-descent parameter of a resolution level indicates the number of decompositions for that level. Different HP-descent parameters can be set for different resolution levels. The parameters are specified via the `-Fdecomp` command line argument. Figure 1 illustrates the three possible decompositions for one resolution level.

Figures 3, 5, and 7 in the next section show how particular subband structures are approximated in VM6.0. The choices of the HP-descent parameters for the approximation have been selected to be optimal in a way that the area below the rate-distortion curve is maximized (for rate-distortion results of these approximations see the next section as well, Figs. 4, 6, and 8, respectively). Note that the "optimal" decomposition structures have been created by applying the best basis algorithm with different cost functions.

Part II of the JPEG2000 standard provides extended functionality in defining the wavelet decomposition and subband structure of an image to compress. However it is still not possible to define really arbitrary structures, although this is pointed out in the FCD document.

When constructing the several resolution levels, there is now the choice to omit splitting a resolution in one of the two directions. This results in resolutions which — in one direction — have the same dimension as their "mother" resolution. Figure 2 illustrates how the division of a resolution into subbands can now be performed. The subbands $S_0(a_b)$ form the next lower resolution level, while the remaining subband(s) make(s) up the current resolution level. In VM6.0 and JPEG 2000 Part I only case a) is possible.

In the same manner as the resolution formation is performed, the subbands of a decomposition level can be further divided into a maximum depth of three sub-levels, counting the first sub-level as the one resulting from the construction of the current resolution as described above. In each division process again, either horizontal or vertical splitting can be omitted.

The concept of sub-levels was also already implemented in VM6.0 with means of the HP-descent parameter, although only a bi-directional division into four subbands (Figure 2 a)) was possible. JPEG2000 Part I allowes no further division of the high-pass subbands at all.

The most interesting innovation in JPEG2000 Part II is however the fact that arbitrarily selected subbands can be further divided within the range of three sub-levels while others are kept as they are. This enables non-uniformly sized subbands at a single resolution level which is impossible in JPEG2000 Part I and VM6.0.

The figures in the following section also contain an approximation of an optimal subband structure with the extended possibilities of JPEG2000 Part II. The maximum number of three sub-levels per decomposition level however still prevents really arbitrary subband structures.

4. EXPERIMENTAL RESULTS

For the *"Barbara"* image the optimal subband structure has been found using the best basis algorithm with cost function log. of energy. Especially the deep decompositions in the high pass subband of the optimal structure show the limits of the VM6.0 and Part II approximations (Figure 3) where only very coarse approximations are possible.

Interstingly, the coarseness of the approximations are not reflected in the results concerning rate/distortion performance. In Fig. 4 we see that all three wavelet subband structures (the optimal one, and the VM6.0 and Part II approximations, respectively) perform almost identically and outperform the pyramidal structure by 0.8 dB across the entire range of bitrates.

The *"D105"* image is contained in the Brodatz texture collection (see Fig. 5). The optimal subband structure exhibits a different amount of high frequency emphasis in horizontal and vertical direction, which can be approximated only very poorly with the original VM6.0 techniques. The JPEG 2000 Part II approximation is better, but limited to the more low

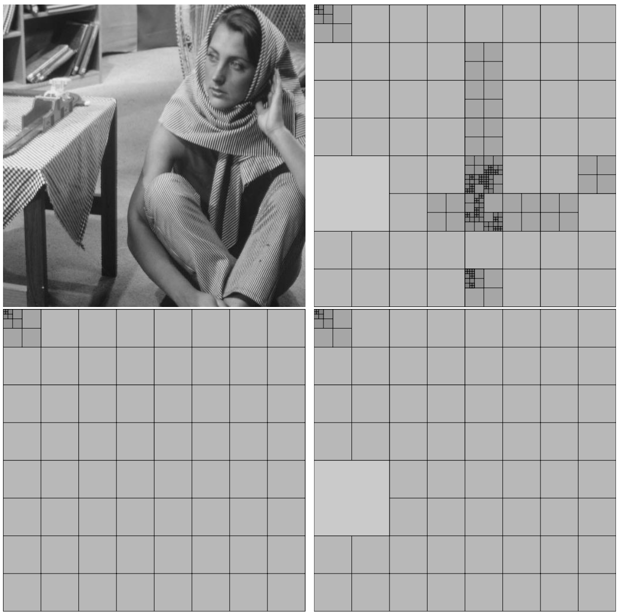

Figure 3: Testimage "Barbara", its optimal subband structure (top right), and the approximations in VM6.0 (argument `-Fdecomp 321`, bottom left) and JPEG2000 Part II (bottom right).

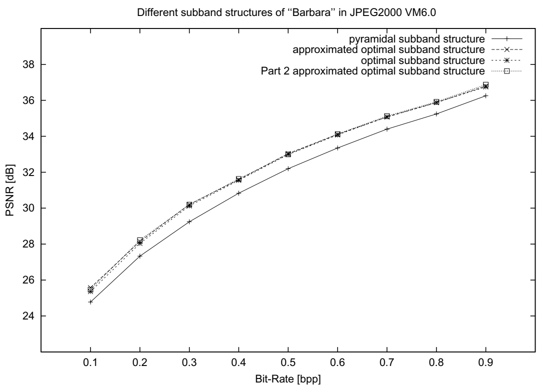

Figure 4: Compression results, Barbara testimage.

frequency subbands.

It turns out that we get a slightly different result with respect to compression behaviour as compared to the Barbara image (Fig. 6). Again, all three wavelet packet subband structures outperform the pyramidal scheme, in this case by more than 2 dB. Whereas the behaviour of the wavelet packet schemes is again identical for higher bitrates (i.e. 0.4 – 0.9 bpp), there is still a performace gap of up to one decibel between the optimal stucture and the approximations, the Part II approximation being the superior one.

The *"Artificial"* image is an artificially generated image with high frequency fluctuations in vertical direction and low frequency behaviour in horizontal direction. The optimal subband structure for this image (Figure 7) has the same small subbands at the

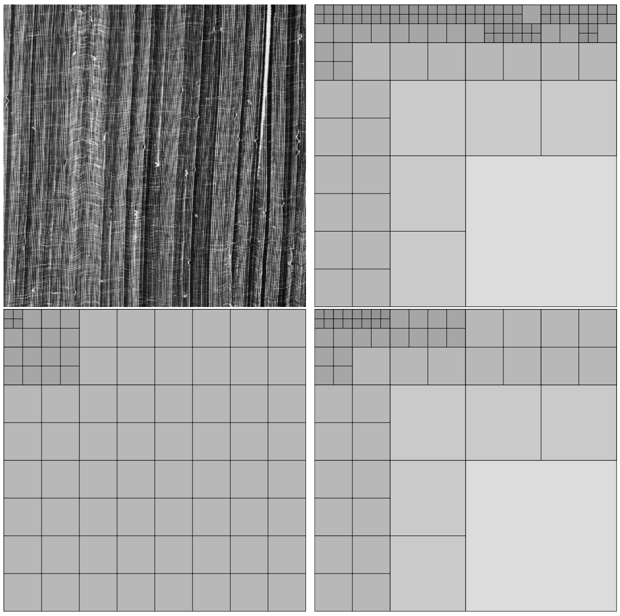

Figure 5: Testimage "D105", its optimal subband structure (cost function: norm l, top right), and the approximations in VM6.0 (argument `-Fdecomp 3221`, bottom left) and JPEG2000 Part II (bottom right).

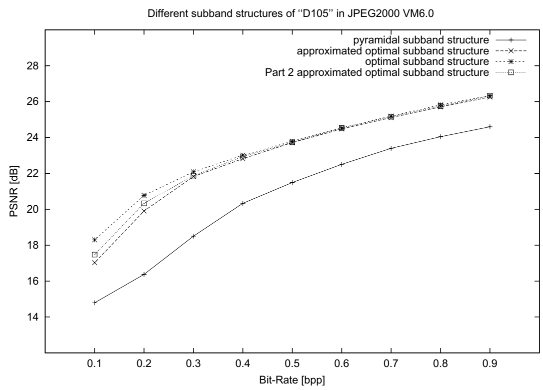

Figure 6: Compression results, D105 testimage.

horizontally-low-pass sections in all decomposition levels. This results in a significantly non-pyramidal subband structure — almost impossible to approximate with the rather pyramidal oriented possibilities of VM6.0. The smallest subbands can only be reconstructed from the third decomposition level on, which prevents a big refinement potential in the highest decomposition levels from being exploited. The Part II approximation is much better, but again restricted to the more low frequency parts.

The compression performance with respect to the "Artificial" image is shown in Fig. 8. The VM6.0 approximation hardly improves the pyramidal scheme at all (only at low bitrates), whereas the Part II approximation is cleary better on the one hand. On the other hand, the Part II scheme is still more than 2

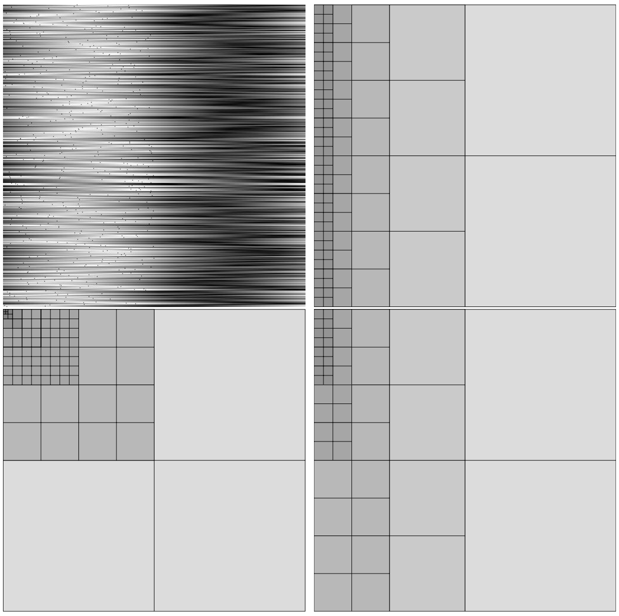

Figure 7: Testimage "Artificial", its optimal subband structure (cost function: log. of energy, top right), and the approximations in VM6.0 (argument `-Fdecomp 12321`, bottom left) and JPEG2000 Part II (bottom right).

dB below the results of the optimal subband structure. Note that this result is not a typical one and is caused by the highly unnatural frequency content of this image. Images of this type can be compressed much better with anisotropic decomposition schemes as allowed as well in Part II [3].

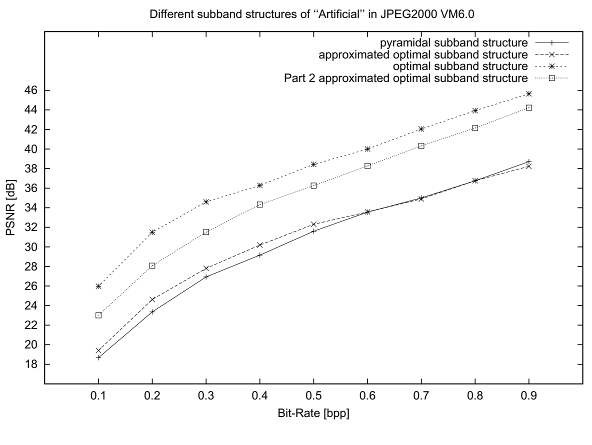

Figure 8: Artificial Testimage.

5. CONCLUSION

It turns out that the possibility to use wavelet packet subband structures in JPEG 2000 Part II clearly improves the suggestions evaluated in JPEG 2000 VM6.0. However, the techniques contained in VM6.0 are already sufficient for most natural images to reach almost optimal coding performance. Only in the case of the highly specific image material the enhanced facilities of JPEG 2000 Part II result in significantly better performance as compared to VM6.0. It seems that the minor restrictions still present in JPEG 2000 Part II are almost entirely irrelevant with respect to rate-distortion performance for almost all types of images.

6. REFERENCES

[1] Jonathan N. Bradley, C. M. Brislawn, and T. Hopper. The FBI wavelet/scalar quantization standard for gray-scale fingerprint image compression. In *SPIE Proceedings, Visual Information Processing II*, volume 1961, pages 293–304, Orlando, FL, USA, April 1993.

[2] R.R. Coifman and M.V. Wickerhauser. Entropy based methods for best basis selection. *IEEE Transactions on Information Theory*, 38(2):719–746, 1992.

[3] R. Kutil. Zerotree image compression using anisotropic wavelet packet transform. In T. Ebrahimi and T. Sikora, editors, *Visual Communications and Image Processing 2003 (VCIP'03)*, volume 5150 of *SPIE Proceedings*, pages 1417–1427, Lugano, Switzerland, July 2003. SPIE.

[4] N. M. Rajpoot, R. G. Wilson, Francois G. Meyer, and R. R. Coifman. A new basis selection paradigm for wavelet packet image coding. In *Proceedings of the IEEE International Conference on Image Processing (ICIP'01)*, pages 816–819, Thessaloniki, Greece, October 2001.

[5] K. Ramchandran and M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Trans. on Image Process.*, 2(2):160–175, 1993.

[6] T. Schell and A. Uhl. New models for generating optimal wavelet-packet-tree-structures. In *Proceedings of the 3rd IEEE Benelux Signal Processing Symposium (SPS 2002)*, pages 225–228, Leuven, Belgium, March 2002. IEEE Benelux Signal Processing Chapter.

[7] D. Taubman and M.W. Marcellin. *JPEG2000 — Image Compression Fundamentals, Standards and Practice.* Kluwer Academic Publishers, 2002.

[8] P.N. Topiwala, editor. *Wavelet Image and Video Compression.* Kluwer Academic Publishers Group, Boston, 1998.

[9] M.V. Wickerhauser. *Adapted wavelet analysis from theory to software.* A.K. Peters, Wellesley, Mass., 1994.