BEST WAVELET PACKET BASES IN A JPEG2000 RATE-DISTORTION SENSE: THE IMPACT OF HEADER DATA

Thomas Stütz, Bernhard Mühlbacher, and Andreas Uhl

Department of Computer Sciences, University of Salzburg, Austria
Email: {tstuetz, bmuehl, uhl}@cosy.sbg.ac.at

ABSTRACT

This paper discusses optimal wavelet packet basis selection within JPEG2000. Two algorithms of Lagrangian rate distortion optimal wavelet packet basis selection for JPEG2000 are presented. The first and more conservative approach considers the JPEG2000 packet body data in the rate distortion optimization only, while the other technique additionally integrates packet header data. The algorithms are evaluated on the FVC2004 fingerprint databases and other textured image data. Results demonstrate that inclusion of header data information into rate distortion optimization leads to superior compression results. For the first time the maximum performance gains of custom isotropic wavelet packets in JPEG2000 can be assessed.

Index Terms— Image compression, JPEG2000, wavelet packet bases, rate distortion optimization

1. INTRODUCTION

Wavelet packet bases (WPBs) [1] offer to adapt the wavelet transform to the source signal (image) characteristics and thus potentially improve the compression performance. WPBs are an alternative to the classical pyramidal wavelet decomposition and allow to further decompose all subbands and not just the LL subband, which leads to an enormous number of possible WPBs. The application of an adapted wavelet packet basis (WPB) for image compression purposes has been subject to investigation since the introduction of the first feasible selection technique called "best basis algorithm" [1]. A brute-force search for the best WPB is computationally infeasible; for 2-D signals and wavelet decomposition depth 5 there are 5.6×10^{78} possible WPBs. In figure 1 examples of WPBs at depth 5 for selected images (see fig. 2) are shown (isotropic decompositions are considered in this work, i.e., a subband is always decomposed horizontally and vertically).

The approach of [1] employs a rate-independent but suboptimal basis selection scheme, which is based on various additive cost functions which only estimate the actual coding cost. An extension to this approach employing non-additive cost functions has been developed soon after [2]. Genetic algorithms have been used [3] to assess the degree of optimality and to further optimize the subband structures found by these algorithms in earlier work.
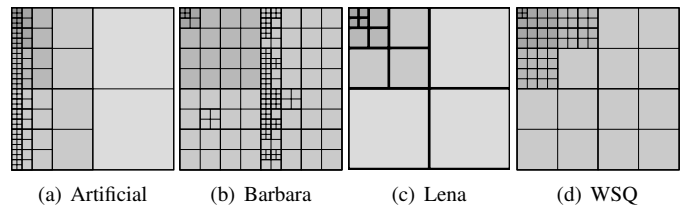

Fig. 1. Best WPBs for specific images and the WSQ-WPB

The employment of rate-distortion optimization criteria for WPB selection has been first demonstrated for classical wavelet-based compression schemes [4]. For certain compression schemes, a certain source image, and a specific target bitrate, the optimal WPB can be computed in feasible time. For zero-tree-based compression algorithms, a Markov chain-based cost function estimating the cost of zero tree coding has been employed to find well suited WPBs [5]. In recent work [6], a wavelet block-based compression scheme has been introduced incorporating the principle of [4] for WPB selection. Subsequent works [7, 8] propose fast and efficient basis selection methods with a lower computational complexity connected with a little loss of rate-distortion performance in comparison with the original work.

The main application field of WPBs in image compression are textured data, with most contributions devoted to fingerprint images. Fingerprint images exhibit characteristic high energy in certain high frequency bands resulting from the ridge-valley pattern and other structures. To account for this property, the WSQ standard for lossy fingerprint compression as adopted by the FBI a specific wavelet packet subband structure which emphasizes the important high frequency bands. Inspired by this algorithm, a few WP-based fingerprint compression schemes have been developed (e.g. [9, 10, 11]).

JPEG2000 Part 2 allows the employment of custom WPBs [12, 13], but WPBs for JPEG2000 have not been subject to extensive investigations so far. In [14], the variants of representing WPBs as discussed during the development of the JPEG2000 Part 2 standard have been assessed with respect to compression performance. For image confidentiality, it has been proposed to use secret wavelet packet bases as a means for compression integrated JPEG2000 encryption [15] (where the impact on compression performance needs to be controlled). Interestingly, at least to the best of the authors' knowledge, the work of [4] has

(a) Artificial   (b) Barbara   (c) Lena   (d) Fingerprint

Fig. 2. Selected test images

not been discussed for JPEG2000 so far.

In this work we show that efficient, best WPB selection is possible in JPEG2000 by an extension of the the approach of [4]. We define (and develop an algorithm for) the Lagrangian cost of a subband in JPEG2000, which enables the determination of the best WPB in a rate-distortion sense. Thus for the first time, the maximal performance gains achievable by an optimal selection of WPBs in JPEG2000 can be assessed. The influence of packet header data on rate-distortion optimal WPB (RDO-WPB) selection is analyzed and evaluated in-depth. Our focus in this work is especially on fingerprint images, for which a custom WPB has been proposed and other highly textured image data. Additionally, the computational complexity of the best WPB selection algorithms for JPEG2000 is discussed as compared to the classical pyramidal decompostion, as mandatory in JPEG2000 Part 1. Section 2 gives an overview of JPEG2000, section 3 discusses two algorithms of Lagrangian rate distortion optimal wavelet packet basis selection within JPEG2000 and section 4 discusses the computational complexity of the two algorithms. Section 5 presents experimental results on fingerprint databases and other textured data.

2. OVERVIEW OF JPEG2000

JPEG2000 employs a wavelet transform and uses the EBCOT-algorithm (embedded block coding with optimized truncation) to encode the wavelet coefficients. The wavelet coefficients of a subband are grouped in rectangular blocks (codeblocks), which are coded independently to separate bitstreams. JPEG2000 Part 2 [16] allows arbitrary even anisotropic WPBs, i.e., a subband is allowed to be only vertically or horizontally decomposed. However, the standard [16, p.54] severely restricts the set of permissible WPBs, every high-frequency subband may only be decomposed two more times (vertically, horizontally or both). In figure 1, the WSQ-WPB is in the set of permissible WPBs, while the best bases for the Artificial and the Barbara image are not. Thus a super-set of the permissible isotropic WPB of JPEG2000 Part 2 is considered in this work. A JPEG2000 file (codestream) consists of a main header followed by several packets. Each packet increases the decoded image quality (i.e. belongs to a certain quality layer). The number of quality layers can be freely chosen. A packet consists of a packet header and a packet body. The packet body is solely comprised of bitstreams (coded codeblock data). The packet header contains information necessary to interpret and decode packet body data. The following data is written in the packet header per each codeblock of the packet: leading zero bitplanes, the length of codeblock contribution, the number of coding passes and the inclusion information.

2.1. Rate-Distortion Optimization in JPEG2000

The embedded bitstream of a single codeblock has several potential truncation points, i.e., each codeblock has a separate RD function. The goal of an encoder is to arrange the bitstream data of all codeblocks in an RD optimal manner, i.e. to find the truncation points which minimize the distortion for a given rate. The most common algorithm for JPEG2000 is PCRD-Optimization (post-compression-rate-distortion). A truncation point of the codeblock B_i is denoted by n_i, all truncation points by $\vec{n}$. The embedded bitstream of the codeblock B_i can be truncated to a rate $R_i^{n_i}$ (for a given truncation point n_i). The rate constraint is then

$$R(\vec{n}) = \sum_{i=1}^{\#cbs\ of\ image} R_i^{n_i} \leq R_{max} \qquad (1)$$

The distortion of each codeblock B_i for a truncation point n_i is given by $D_i^{n_i}$. Given an additive distortion measure, the distortion D of the compressed image is derived by:

$$D(\vec{n}) = \sum_{i=1}^{\#cbs\ of\ image} D_i^{n_i} \qquad (2)$$

An optimal solution (minimizing D) of truncation points $\vec{n}^*$ for this constrained problem can be found by solving the corresponding unconstrained problem (Lagrangian RDO):

$$\vec{n}^* = \operatorname{argmin}_{\vec{n}}[D(\vec{n}) + \lambda R(\vec{n})] \qquad (3)$$

Considering D a function of R, a solution is obtained by setting $D'(R) = -\lambda$, which yields $D_i'(R) = -\lambda$.
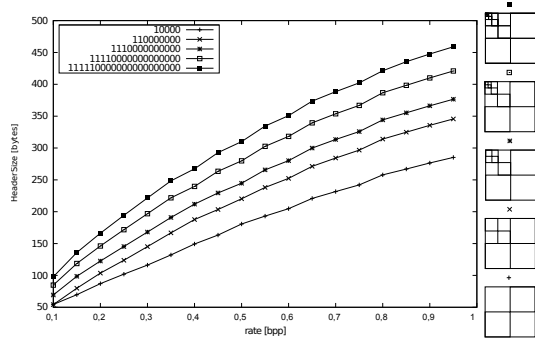
3. BEST WAVELET PACKET BASIS (WPB)

For compression, the best WPB is the one that minimizes the size of the compressed image at a given level of distortion (best WPB in a rate-distortion sense). Thus finding the best solution in a rate-distortion sense depends on the underlying coding mechanisms and might be computationally complex and complex to integrate in a compression framework such as JPEG2000.
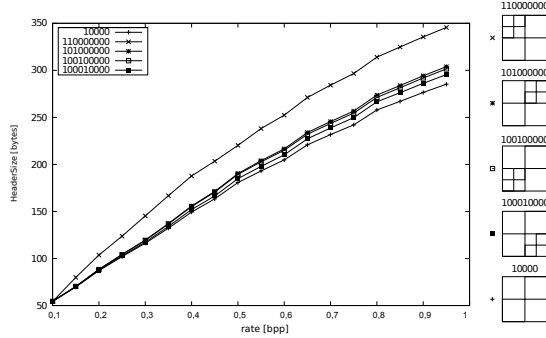
Testing every possible WPB soon becomes infeasible, as the number of possible WPBs is growing tremendously with the decomposition depth d. The following recursion [17] calculates Q_d, the number of possible WPBs at depth j:

$$Q_d = Q_{d-1}^4 + 1 \qquad (4)$$

where $Q_0 = 1$. At depth two we have 17 possible WPBs, at depth three 83522, at depth four 4.9×10^{19}, at depth five 5.6×10^{78}, at depth six 9.9×10^{314}, and at depth seven 9.6×10^{1259}.

(a) Splitting the LL subband

(b) Comparison of the costs of different subbands

Fig. 3. Packet header data for different decomposition depths and subbands

There is a more efficient algorithm for the determination of the "best" WPB (best in the restricted sense of the cost function only): the best basis algorithm (BBA) [1]. The BBA first makes a full wavelet packet decomposition at maximum decomposition depth and starts from the leaves, i.e. the subbands at the deepest decomposition depth. The BBA merges the children subbands of a parent subband if the sum of the costs of its children is higher than the parent's cost.

3.1. Best WPBs in a Rate-Distortion Sense

The best solution for an actual compression framework (JPEG2000) in an RD sense is obtained if the coding costs are not estimated, but actually determined. The cost of a subband sb is calculated by the Lagrangian cost function which is defined as:

$$J(\lambda)_{sb} = D_{sb} + \lambda R_{sb} \qquad (5)$$

Children subbands are not merged if the following split condition holds:

$$J(\lambda)_{parent} \geq \sum_{child=1}^{\#children} J(\lambda)_{child} \qquad (6)$$

In order to obtain a solution for a target bitrate, an efficient bisection search on the parameter λ can be conducted [4].

3.1.1. Lagrangian cost function of a subband for JPEG2000

The essential part of integrating the algorithm for finding the best WPB in a rate-distortion sense into JPEG2000 is to appropriately determine the Lagrangian cost of a subband. A subband consists of several codeblocks, each with a bitstream with its own rate-distortion statistics, i.e. truncation points and associated distortions. These data describe a rate-distortion function with a certain slope at each truncation point and a corresponding Lagrangian cost. The Lagrangian cost of a subband is defined as the sum of the Lagrangian costs of its codeblocks:

$$J(\lambda)_{sb} = \sum_{cb=1}^{\#cbs\ of\ subband} J(\lambda)_{cb} = \sum_{cb=1}^{\#cbs\ of\ subband} D_{cb} + \lambda R_{cb}$$

$$(7)$$

The actual algorithm to determine the Lagrangian cost of a subband is given in pseudo-code (see algorithm 1). This optimiza-

---

Algorithm 1 Lagrangian cost function of a subband

```
Param: λ
costs = 0
for (b = 0; b <all code-blocks of the subband; b + +) do
    for (slopeIdx  =  0; slopeIdx  <all  slopes  of
    codeblock[b]; slopeIdx + +) do
        slope = block[b].slopes[slopeIdx].slope
        if slope < λ then
            currentSlope = slope
            break;
        end if
    end for
    costs += getDistortion(slopeIdx) + lambda * getRate(slopeIdx)
end for
return costs
```

---

tion minimizes the overall packet body size, and also minimizes the overall file size if the cost of coding the headers is not influenced by the selection of the WPB (we refer to this algorithm as RDO-WPB).

3.1.2. Considering the Packet Header in the Lagrangian Cost of a Subband

In figure 3(a) the packet header cost for the LL subband is analyzed in detail, the cost of the packet headers is plotted for increasing decomposition depths and varying rate. The LL suband's packet header cost is compared to the other subbands in figure 3(b). The packet header cost scales well with the overall target bitrate. In the RDO-WPB algorithm the packet header data cost is considered constant and independent of the decomposition. I.e., the packet header cost of a subband and a further decomposed subband (the sum of the packet header costs of its children) are assumed equal. Occasionally, this simplification has to be paid by sub-optimal compression performance, as can be seen in figure 4, where RDO-WPB is clearly outperformed by the algorithm RDOH-WPB, which takes packet header data into account. These results result reflect the general trend (see fig. 8 for average results). Considering only the packet body
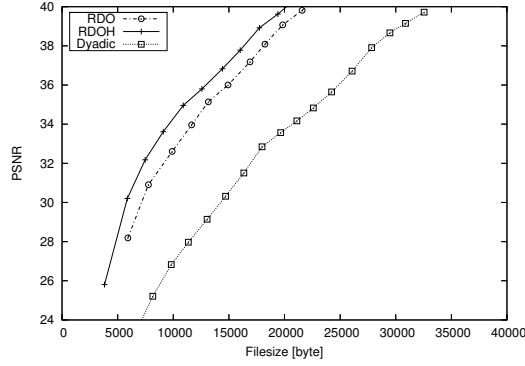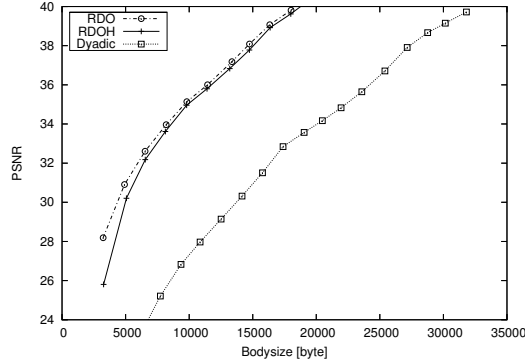
Fig. 4. Artificial image: file size

Fig. 5. Artificial image: body size

size, RDO-WPB is optimal (see fig. 5). Most important for the difference in the packet header cost for subbands and further decomposed subbands is the triple: size of the image (x, y), codeblock size (cb_x, cb_y), and wavelet decomposition depth of the subband under investigation d. If $cb_x > x/2^d$ or $cb_y > y/2^d$, the subband is decomposed into subbands smaller than the codeblock size and additional entries for the new codeblocks have to be added to the packet header, and the packet header length is increased. E.g., for 512x512 images, and $d = 5$, further decomposing the subbands at depth 3 becomes more expensive in terms of packet header cost.

Therefore, the packet header cost has to be integrated in the Lagrangian cost of a subband. More concisely, we take the overall packet header costs at a full decomposition for a specific depth d and for the target rate r (λ is determined with a bisection algorithm to match r [4]) and divide it by the numbers of subbands. The result is the average size of the packet header data $R_{d,r}^H$ of a subband at depth d for a target bit rate of r, which also reflects the rate-dependency of the packet header cost. The Lagrangian cost of a subband sb at depth d is computed as follows:

$$J(\lambda)_{sb} = D_{sb} + \lambda(R_{sb} + R_{d,r}^H). \tag{8}$$

We refer to this extended algorithm by RDOH-WPB.

$cb_x cb_y$	64^2	32^2	16^2	8^2	4^2
No DWT	0.6s	0.7s	0.9s	1.5s	3.4s

Table 1. Runtime performance in seconds depending on code-block size for 512x512 images and no DWT

d	0	1	2	3	4	5	6
$cb_x cb_y$	64^2	64^2	64^2	64^2	32^2	16^2	8^2
FD	0.6s	0.63s	0.7s	0.75s	0.9s	1.8s	15s
WPB	0.6s	1.00s	1.1s	1.30s	1.5s	2.8s	18s

Table 2. Runtime performance in seconds depending on decomposition depth d of full decomposition (FD) and RDO[H]-WPB (WPB) for 512x512 images

4. COMPLEXITY

The asymptotic complexity of rate distortion optimal WPB selection for a maximum decomposition depth d and an N-element signal is of $\mathcal{O}(d \times N)$ and also in $\mathcal{O}(N\log N)$, because d is bounded by $\log N$, which is the maximal decomposition depth.

In order to assess the concrete complexity of RDO[H]-WPB (both have the same order of complexity), we consider the computationally complex parts of the JPEG2000 compression pipeline: DWT (and quantization), coding of codeblocks (and RDO), as well as file I/O (including final bitstream formation). Depending on the implementation, the compression settings and source data the shares vary; for JPEG2000 Part 1, overall 0.8 seconds are needed for a 512x512 image with JJ2000 default settings (all evaluations are performed on an Intel Core2 6700@2.66GHz and the software described in sect. 5). For the RDO[H]-WPB with a maximum decomposition depth d a full wavelet decomposition for every depth l ($1 \leq l \leq d$) has to be performed. The coefficients of a full decomposition at depth l can be used to compute the coefficients of the next depth $l + 1$. In terms of a DWT at depth 1, $\mathcal{W}$, the cost of all decompositions is at least $d \times \mathcal{W}$. However, in practice the runtime complexity is tremendously increased for decomposition depths greater 5 (see table 2). Coding and RDO has to be done at every depth l and for no decomposition as well, which adds up to $(d+1) \times \mathcal{B}$, where $\mathcal{B}$ is the cost of coding all coefficients. As long as the subbands are larger than the codeblocks, their coding and RDO cost remains approximately constant for all depths. If the subbands become smaller than the codeblocks, the runtime performance decreases, however, this effect is implementation-specific (see table 1 for JJ2000's behavior).

The overall cost for RDO[H]-WPB at depth d, $\mathcal{R}(d)$, in terms of a compression at depth 1, $\mathcal{C}$, is approximately $\mathcal{R}(d) \approx d \times \mathcal{C} - (d - 1) \times \mathcal{D}$, where $\mathcal{D}$ represents fixed time, e.g., for the actual JJ2000 implementation: Java start up time, image IO, and bitstream IO (approx. 0.3s).

In conclusion, for a reasonable wavelet decomposition depth of 4 our RDO[H]-WPB implementation only takes less than twice the default JJ2000 compression time, which is in-line with
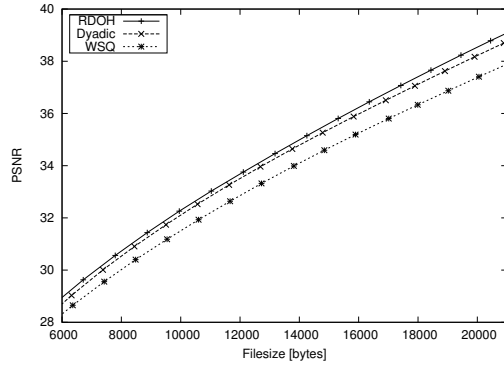
Fig. 6. FVC2004: Dyadic vs RDOH-WPB vs WSQ

our theoretical analysis, which predicts $4 \times 0.63s - 3 \times 0.3s = 1.62s \approx 1.5s$.

5. EXPERIMENTAL RESULTS

The results have been produced with a custom implementation, which is based on the JJ2000 reference implementation. The correctness of our implementation of RDO[H]-WPB has been experimentally verified on the entire FVC2004 database (32000 images) for depth 2 by testing every possible WPB. For single test images verification has been conducted for depth 3 as well. The default settings of the JJ2000 implementation have been employed, e.g., a decomposition depth of 5, 9/7 CDF filter, and 64x64 codeblocks.

The presented results focus on fingerprint data (FVC2004 database), and highly textured data (Brodatz database and the test image "Barbara" and "Artificial", see fig. 2). For natural images, e.g., the "Lena" image, RDO[H]-WPB can not achieve significant performance improvements (tested on 1000 natural images). Also, we tested several well-known cost functions, such as the L1-norm, the L2-norm, the log energy metric, and the entropy information cost [1]; our evaluations revealed that these do not work reliably for best WPB selection in JPEG2000.

5.1. Compression Performance Evaluation for Fingerprint Data

Additionally to our best WPB selection algorithms for JPEG2000 we evaluated the compression performance of the WSQ-WPB on the FVC2004 database. In figure 6 the dyadic-WPB, the WSQ-WPB and the WPB calculated by RDOH-WPB are compared (RDO-WPB is only slightly worse). An improvement with RDOH-WPB can be seen, while the WSQ-WPB significantly reduces the compression performance. Especially at higher decomposition depths the header data needs to be taken into account (see fig. 8).

5.2. Compression Performance Evaluation for Highly-Textured Data

The results for the "Artificial" and the "Barbara" are plotted in figures 4 and 7. The performance gains for RDO[H]-WPB for
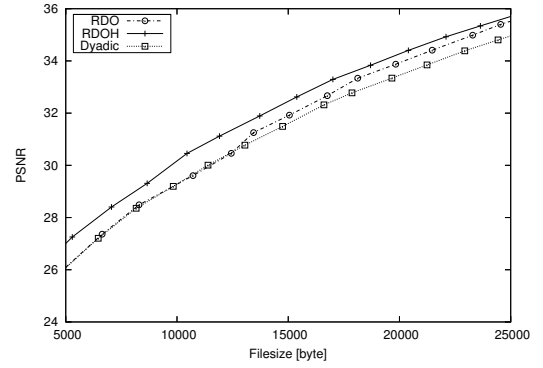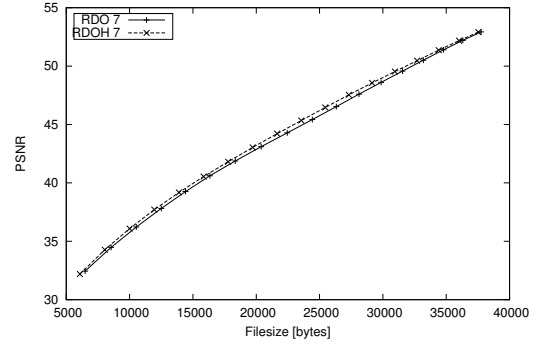

Fig. 7. Barbara image: file size

Fig. 8. The impact of header data for the DB1 of FVC2004 for wavelet decomposition level 7

the "Artificial" image are enormous, and considerable for the "Barbara" image as well. For both RDOH-WPB outperforms RDO-WPB. On the Brodatz database both algorithms perform equally well, especially on a subset consisting of 20% of the Brodatz images: RDO[H]-WPB achieves impressive compression performance gains (see fig. 9). Thus for a considerable share of textured data, best basis selection can be recommended.
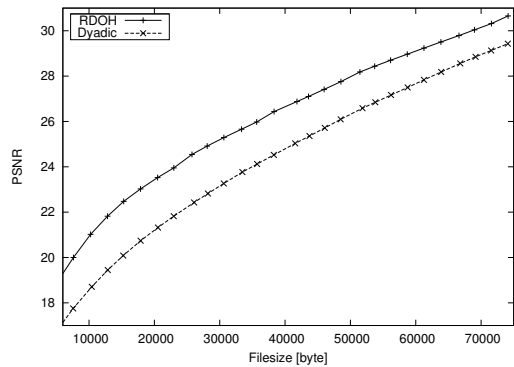

Fig. 9. A subset of 20% of the Brodatz database

6. CONCLUSION

Rate distortion optimal (RDO) wavelet packet basis (WPB) selection for JPEG2000 has been presented. Two algorithms are proposed, namely RDO-WPB, which optimizes only the JPEG2000 packet body data and RDOH-WPB, which also takes the packet header data into account. For the first time the actual compression performance gains achievable by custom isotropic WPB and the JPEG2000 coding framework can be assessed concisely. The complexity of RDO[H]-WPB selection is discussed and evaluated. In terms of compression performance our results show that for highly textured data, RDOH-WPB performs significantly better than RDO-WPB. Therefore, taking the packet headers into account for JPEG2000 WPB optimization is favorable. Furthermore, RDO-WPB and RDOH-WPB improve the compression performance for fingerprint data, while the WSQ-WPB, which has been explicitly proposed for fingerprint images, significantly reduces the rate-distortion compression performance in JPEG2000.

7. REFERENCES

[1] R.R. Coifman and M.V. Wickerhauser, "Entropy based methods for best basis selection," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 719–746, 1992.

[2] C. Taswell, "Satisficing search algorithms for selecting near-best bases in adaptive tree-structured wavelet transforms," *IEEE Transactions on Signal Processing*, vol. 44, no. 10, pp. 2423–2438, Oct. 1996.

[3] T. Schell and A. Uhl, "Optimization and assessment of wavelet packet decompositions with evolutionary computation," *EURASIP Journal on Applied Signal Processing*, vol. 2003, no. 8, pp. 806–813, 2003.

[4] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. on Image Process.*, vol. 2, no. 2, pp. 160–175, 1993.

[5] N. M. Rajpoot, R. G. Wilson, Francois G. Meyer, and R. R. Coifman, "Adaptive wavelet packet basis selection for zerotree image coding," *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1460–1472, 2003.

[6] Yongming Yang and Chao Xu, "A wavelet packet based block-partitioning image coding algorithm with rate-distortion optimization," in *Proceedings of the IEEE International Conference on Image Processing (ICIP'05)*. 2005, vol. III, pp. 201–204, IEEE.

[7] Yongming Yang and Chao Xu, "Fast and efficient basis selection methods for embedded wavelet packet image coding," in *Proceedings of Third International Conference of Image Analysis and Recognition, ICIAR'06*. Sept. 2006, vol. 4141 of *Lecture Notes in Computer Science*, pp. 480–492, Springer-Verlag.

[8] Yongming Yang and Chao Xu, "Fast wavelet packet basis selection for block-partitioning image coding," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS 2006)*, Sept. 2006, p. 4 pages.

[9] S. Kasaei, M. Deriche, and B. Boashash, "A novel fingerprint image compression technique using wavelet packets and pyramid lattice vector quantization," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1365–1378, 2002.

[10] P. Saeedian and B. Shirazi, "A novel fingerprint image compression technique using adaptive subband image coding," in *Proceedings of the 2004 Picture Coding Symposium PCS'04*, 2004.

[11] G.A. Khuwaja, "Best parameter based compression of fingerprints with wavelet packets," *International Journal of Computer Applications in Technology*, vol. 19, no. 1, pp. 51–62, 2004.

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

[13] ISO/IEC 15444-2, "Information technology – JPEG2000 image coding system, Part 2: Extensions," May 2004.

[14] M. Reisecker and A. Uhl, "Wavelet-packet subband structures in the evolution of the JPEG2000 standard," in *CD-ROM Proceedings of the 6th IEEE Nordic Signal Processing Symposium (NORSIG 2004)*, Espoo, Finland, June 2004, IEEE Norway Section.

[15] Dominik Engel and Andreas Uhl, "Secret wavelet packet decompositions for JPEG2000 lightweight encryption," in *Proceedings of 31st International Conference on Acoustics, Speech, and Signal Processing, ICASSP '06*, Toulouse, France, May 2006, vol. V, pp. 465–468, IEEE.

[16] ITU-T T.801, "Information technology – JPEG2000 image coding system, Part 2: Extensions," Feb. 2002.

[17] D. Xu and M. N. Do, "Anisotropic 2-D wavelet packets and rectangular tiling: Theory and algorithms," *in Proc. of SPIE Conf. on Wavelets X*, vol. 19, pp. 619–630, 2003.