UNIVERSITÄT SALZBURG

Best Wavelet Packet Bases in a JPEG2000 Rate-Distortion Sense

Thomas Stütz Bernhard Mühlbacher Andreas Uhl

Technical Report 2010-11 December 2010

Department of Computer Sciences

Jakob-Haringer-Straße 2
5020 Salzburg
Austria
www.cosy.sbg.ac.at

Technical Report Series

Best Wavelet Packet Bases in a JPEG2000 Rate-Distortion Sense

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. Three algorithms for rate distortion optimal wavelet packet basis selection in JPEG2000 are presented. The first approach considers the JPEG2000 packet body data in the rate distortion optimization only, while the other techniques additionally integrate packet header data. The algorithms are evaluated on a wide range of highly 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.**

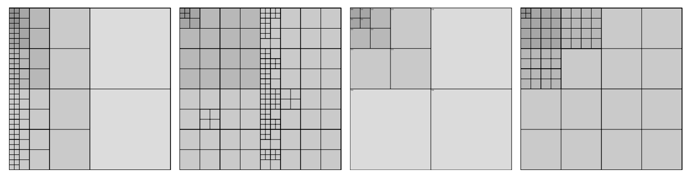
***Index Terms*—Image compression, JPEG2000, wavelet packet bases, rate distortion optimization**

I. 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 dyadic wavelet decomposition (also referred to as pyramidal) 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 even for moderate maximum decomposition depths; 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 sub-optimal 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 the best basis algorithms proposed in earlier work.

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.



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

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

The main application field of WPBs in image compression are textured data, with many 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, optimal wavelet packet basis selection in a rate distortion sense [4] has not been discussed for JPEG2000 so far.

In this work we show that efficient, best WPB selection is

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

Fig. 2. Selected test images

possible in JPEG2000 by an extension of 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 header data on rate-distortion optimal WPB (RDO-WPB) selection is analyzed and evaluated in-depth. Our focus in this work is on highly textured image data, especially on fingerprint images, for which a custom WPB has been proposed. Additionally, the computational complexity of the best WPB selection algorithms for JPEG2000 is discussed as compared to the classical dyadic decompostion, as mandatory in JPEG2000 Part 1.

There are substantial extensions to own previous work [16], [17] as well: The development, implementation and evaluation of a concise header cost determination algorithm, the implementation and evaluation of a lossless coding mode for our rate-distortion optimal wavelet packet coder, and the improvement of the evaluation framework, which now analyzes the compression performance with state-of-the-art quality metrics. The novel packet header cost determination allows to precisely assess the actual cost of the packet header portion of a subband, thus enabling perfect rate-distortion optimization without the imprecisions of the header cost estimation. Thus our JPEG2000-based WPB coder achieves always better or equal results compared to the underlying JPEG2000 Part 1 coder. Furthermore the source code of our coder and evaluations will be made publicly available, making our results easily reproducible [18].

Section II gives an overview of JPEG2000, section III discusses algorithms of Lagrangian rate distortion optimal wavelet packet basis selection within JPEG2000 and section IV discusses the computational complexity of the algorithms. Section V presents experimental results on fingerprint databases and other textured data.

II. 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 [19] allows arbitrary WPBs. The standard [19, p.54] restricts the set of permissible WPBs, every high-frequency subband may only be decomposed two more

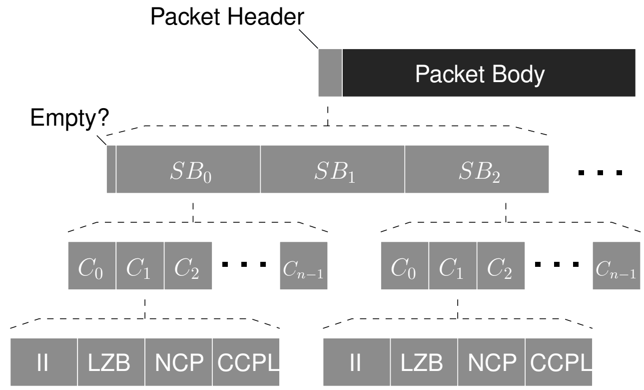


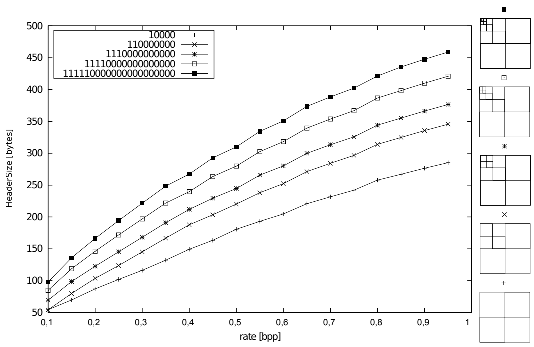
Fig. 3. Packet header formation

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 superset of the permissible isotropic WPB of JPEG2000 Part 2 is considered in this work.

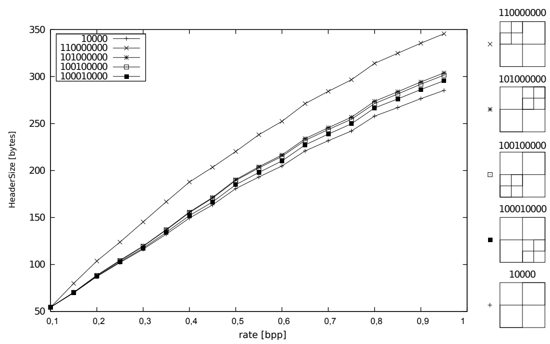
A JPEG2000 file (codestream) consists of a main header followed by several packets. Each packet increases the decoded image quality. Each packet belongs to a certain quality layer and resolution. The number of quality layers can be freely chosen (for the scope of this work we set the number of quality layers to one). 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 for each codeblock of the subbands of the packet's resolution: leading zero bitplanes, the length of codeblock contribution, the number of coding passes and the inclusion information. The packet header formation is illustrated in figure 3. II denotes inclusion information, i.e., whether the code block contributes to the packet. LZB denotes leading zero bitplanes of the coefficients of a codeblock. NCP denotes the number of contributing coding passes (EBCOT employs three coding passes for a single bitplane of the coefficients of a codeblock), and CCPL denotes the length of the coded code block contribution in the packet body.

A. 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).



(a) Splitting the LL subband



(b) Comparison of the costs of different subbands

Fig. 4. Packet header data for different decomposition depths and subbands

The rate constraint is then

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

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:

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

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):

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

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

III. Best Wavelet Packet Basis

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 [20] calculates Q_d, the number of possible WPBs at depth j:

$$Q_d = Q_{d-1}^4 + 1 \tag{4}$$

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}.

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.

A. 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:

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

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

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

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

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:

$$J(\lambda)_{sb} = \sum_{cb=1}^{\#cbs\ of\ subband} J(\lambda)_{cb} = \sum_{cb=1}^{\#cbs\ of\ subband} D_{cb,\lambda} + \lambda R_{cb,\lambda} \tag{7}$$

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

This optimization 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).

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 $slope < \lambda$ **then**
      $currentSlope = slope$
      break;
    end if
  end for
  costs += getDistortion(slopeIdx) + lambda * getRate(slopeIdx)
end for
return costs

2) Considering the Packet Header in the Lagrangian Cost of a Subband: In figure 4(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 4(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. This simplification has to be paid by sub-optimal compression performance, as can be seen in figure 5, where RDO-WPB is clearly outperformed by the algorithm RDOH-WPB, which takes header data into account. Considering only the packet body size, RDO-WPB is optimal (see fig. 6). The performance gains for RDO[H]-WPB for the "Artificial" image are enormous, and considerable for the "Barbara" image as well (see figure 7). For both RDOH-WPB outperforms RDO-WPB. The packet header data is coded per subband, each subband maintains its own coding states, e.g., in the form of associated tag trees. The packet header cost for subbands and further decomposed subbands further depends on the triple: size of the image (x, y), code-block 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 the number code blocks.

The header cost has to be integrated in the Lagrangian cost of a subband in order to reflect the increased coding cost for a further decomposition of a subband. In [16] it is proposed to estimate the actual header cost of a subband by considering the overall 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 header data $R^H_{d,r}$ of a subband at depth d for a target bit rate of r. The Lagrangian cost of a subband sb at depth d is computed as

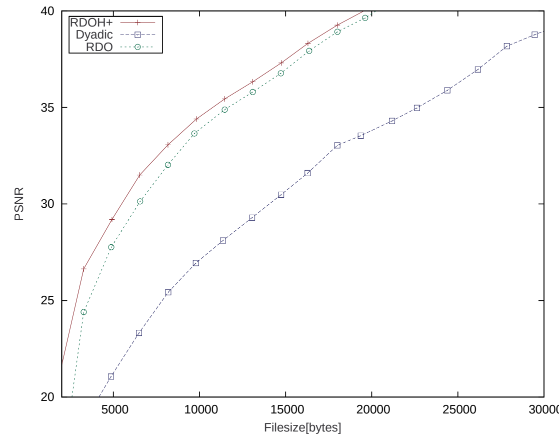


Fig. 5. Artificial image: file size

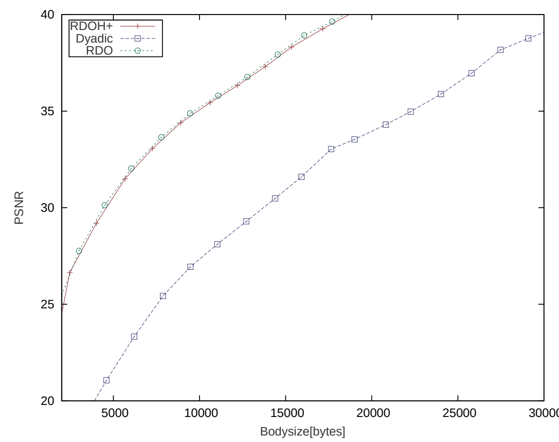


Fig. 6. Artificial image: body size

follows in [16]:

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

We refer to this extended algorithm by RDOH-WPB.

However, this is only an estimation of the header cost of a subband, which nonetheless is shown to lead to significant performance advantages compared to the optimization without the consideration of the header cost (RDO-WPB). Naturally

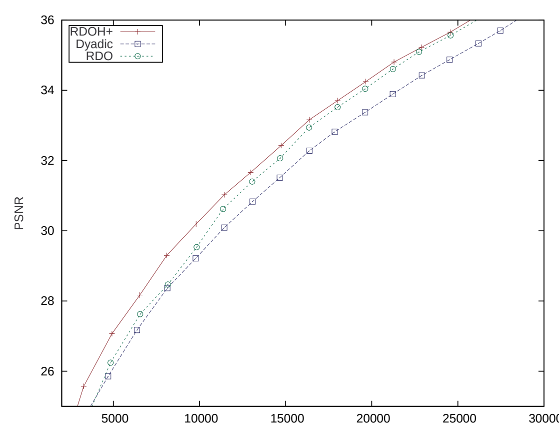


Fig. 7. Barbara image: file size

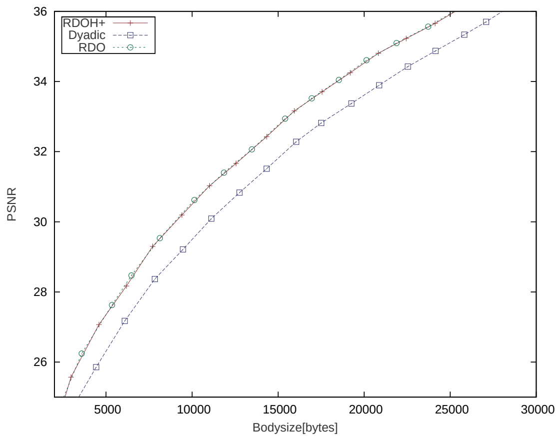


Fig. 8. Barbara image: file size

the best solution would be achieved by the determination of the actual coding cost of the header data of a subband, instead of its estimation. For that end, the coding of the packet header portion of each subband has to be simulated in order to determine its length, i.e., the rate of the Lagrangian cost. More precisely, for each subband and lambda the optimal truncation points for the coded code block data are determined and afterwards the packet header portion of the subband is determined by simulating the coding in dependency of the packet body RDO, i.e., the obtained truncation points. Thus we propose to employ the following Lagrangian cost of a subband sb for a Lagrangian parameter λ:

$$J(\lambda)_{sb} = D_{sb,\lambda} + \lambda(R_{sb,\lambda} + R^H_{sb,\lambda}). \tag{9}$$

We refer to this extended algorithm by RDOH+WPB and because the estimate of the packet header cost is replaced by the actual packet header cost, this algorithm achieves the best possible performance.

3) Bitrate Adjustment: A Bisection Search for the Lagrangian λ: The actual bitrate of the best basis algorithms in a JPEG2000 rate-distortion sense is determined by the parameter λ. Thus in order to achieve a certain target bitrate the appropriate value of λ has to be determined. Our implementation employs a classic bisection method [4], [21]. Starting from a lower (λ_{min}) and an upper bound (λ_{max}) for λ, i.e., equation 10 has to be satisfied, a bisection search is performed, as outlined in algorithm 2.

$$R(\lambda_{min}) \leq R_{budget} \leq R(\lambda_{max}) \tag{10}$$

The initial values of λ_{min} and λ_{max} can be set to the minimum and maximum slope of the rate distortion functions of all codeblocks; suitable initial values have been found experimentally to be 0 and 10.

B. Ad-hoc cost functions

Alternatively to the optimal wavelet packet basis in a rate-distortion sense with the actual coding bitrate as cost function it has often been proposed to employ simpler cost functions for best basis selection (although these may not result in best

Algorithm 2 Bisection search for λ

```
Param: R_budget
λ = λ_min + (λ_max − λ_min)/2
R_result = 0
maxSearchSteps = 32
numSearchSteps = 0
ε = 10^−6
while numSearchSteps < maxSearchSteps do
   numSearchSteps++
   R_current = getRateForLambda (λ)
   if R_current == R_budget then
      λ_result = λ
      R_result = R_budget
      break;
   end if
   if R_current < R_budget then
      if R_result < R_current then
         λ_result = λ
         R_result = R_current
      end if
      if |λ_max − λ| < ε then
         break;
      end if
      λ_max = λ
      λ = λ_min + (λ_max − λ_min)/2
   else if R_current > R_budget then
      if |λ_min − λ| < ε then
         break;
      end if
      λ_min = λ
      λ = λ_min + (λ_max − λ_min)/2
   end if
end while
return λ_result
```

bases in a rate-distortion sense). In this section we will present common cost functions. Let c_i represent the value of the coefficients of a subband. The cost functions are calculated as follows:

- L1-norm metric: $\sum_i |c_i|$
- L2-norm metric: $\sum_i c_i^2$
- LogE - log energy metric: $\sum_i ln(c_i^2)$.
- EIC - entropy information cost or Shannon metric [1]): $\sum_i c_i^2 * ln(c_i^2)$

IV. 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 similar complexity, as only a constant for the header cost is added in Lagrangian cost of a subband), 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

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

TABLE I
RUNTIME PERFORMANCE IN SECONDS DEPENDING ON CODEBLOCK 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 II
RUNTIME PERFORMANCE IN SECONDS DEPENDING ON DECOMPOSITION DEPTH d OF FULL DECOMPOSITION (FD) AND RDO[H]-WPB (WPB) FOR 512X512 IMAGES

default settings (all evaluations are performed on an Intel Core2 6700@2.66GHz and the software described in sect. V). 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 II). 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 I 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 our theoretical analysis, which predicts $4 \times 0.63s - 3 \times 0.3s = 1.62s \approx 1.5s$.

The complexity of RDOH+ is increased by the coding of the packet header portions of a subband, which becomes significant for higher decomposition depths as the number of subbands n_{sb} grows exponentially with the decomposition depth d of a full wavelet packet decomposition ($n_{sb} = 4^d$). At a decompostion depth of 5 a single RDOH+ compression takes about double the time of the RDOH version with the packet header estimate. Thus for higher decomposition depths estimating the packet header cost is the method of choice.

V. 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., 9-7 irreversible filter for lossy compression and the 5-3 reversible filer for lossless, and 64x64 codeblocks. A maximum decomposition depth of 5 has been employed (if not explicitly stated otherwise) and only one quality layer is employed. Additionally to the well-known PSNR we present results for state-of-the-art quality metrics [22], such as the VIF, MSSIM, and SSIM. The matlab package metrix_mux has been employed for quality evaluation.

We present results for highly textured data (Brodatz database) and for fingerprint data (FVC2004 database). For natural images, e.g., the "Lena" image, RDO[H]-WPB can not achieve significant performance improvements (tested on 1000 natural images), as the best bases are very similar to the dyadic wavelet decompostion, i.e., the best WPB basis is the standard WPB basis of JPEG2000 Part 1. 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.

A. Lossless compression

JPEG2000 also offers a lossless compression pipeline. In this context RDO WPB selection is actually a rate optimal WPB selection. In tables III,IV,V, and VI the results for the test images, Artificial, Barbara, and Lena are given. In the first column the achieved size of the dyadic wavelet decomposition is given, the next column gives the reduction of the compressed size compared to the dyadic decomposition for RDO and the next column for RDOH+ (the results for RDOH are in general in between and omitted). For a decompostion depth of 5 compression efficiency is already improved for both algorithms RDO and RDOH+, which yield the same result. Further increasing the decomposition depth reveals the difference between RDO and RDOH+, the compression efficiency of RDO is decreased with increasing depth (as trend also followed by the dyadic decomposition) while RDOH+ can still remain its compression efficiency and even improve it. The improvement of compression efficiency is most pronounced for the Artificial and least significant for the natural Lena image.

Considering texture data, especially a well-performing subset of the Brodatz database, we see that highly textured data is well-suited for wavelet packet compression, almost 7KB are saved compared to the dyadic decomposition (see tables VII,VIII, IX, and X). For lower decomposition depths RDO-WPB performs quite well, only at higher decomposition depths RDOH+ can achieve performance improvements.

For fingerprint data performance gains can also be achieved, but they are less pronounced (see table XI for an overview).

Image	Dyadic	Reduc. RDO	F.reduc. RDOH+
artificial512	78108	19847	0
barbara	156758	209	0
lena	153511	0	0
Avg.	129459.00	6685.33	0.00

TABLE III
SELECTED TEST IMAGES WITH MAX. DECOMP. 4

Image	Dyadic	Reduc. RDO	F.reduc. RDOH+
artificial512	77953	20744	0
barbara	156784	235	0
lena	153529	18	0
Avg.	129422.00	6999.00	0.00

TABLE IV
SELECTED TEST IMAGES WITH MAX. DECOMP. 5

B. Lossy Compression

The RD performance of our algorithms has been evaluated with the objective quality metrics PSNR, VIF, MSSIM, and SSIM, which all support the conclusion from the PSNR analysis. Additional to the PSNR results, we give the result of the most sophisticated quality metric VIF, which shows better correlation with human subjective quality perception than the PSNR [22].

At higher decomposition depths RDOH+ is superior to RDO, i.e., the consideration of header data leads to performance improvements. For the Artificial image concise PSNR and VIF results are summarized in the tables XII and XIII. Enormous PSNR improvements are achieved for RDO and RDOH+ (over 7dB), the RDOH+ algorithm works reliable for all quality ranges and outperforms RDO significantly, especially in lower quality range. The VIF results report a similar objective quality behavior. For the Barbara image improvements of about 1dB are achieved with RDOH+, which outperforms RDO (see table XIV). The VIF shows a similar behavior, although smaller PSNR differences are no longer distinguishable with two digits precision. Thus at higher maximum decompostion depths the consideration of the header data is recommended.

On the Brodatz database RDO wavelet packet bases perform well, especially on a subset consisting of 20% of the Brodatz images: RDO[H+]-WPB achieves impressive compression per-

Image	Dyadic	Reduc. RDO	F.reduc. RDOH+
artificial512	77931	20561	196
barbara	156804	255	0
lena	153547	36	0
Avg.	129427.33	6950.66	65.33

TABLE V
SELECTED TEST IMAGES WITH MAX. DECOMP. 6

Image	Dyadic	Reduc. RDO	F.reduc. RDOH+
artificial512	77937	20560	203
barbara	156820	271	0
lena	153562	51	0
Avg.	129439.66	6960.66	67.66

TABLE VI
SELECTED TEST IMAGES WITH MAX. DECOMP. 7

Image	Dyadic	Reduc. RDO	F.reduc. RDOH+
D101	295537	7218	0
D102	288975	13819	0
D103	345204	2839	0
D105	346258	6871	0
D106	356648	6912	0
D107	324075	6904	0
D109	333892	19790	0
D16	390590	14555	32
D21	362414	12108	6
D49	251011	14049	0
D53	319139	4702	6
D6	293547	5250	0
D64	275410	237	0
D67	317191	8607	0
D68	286131	1259	0
D76	327516	160	1
D77	365771	5247	0
D78	347081	2648	0
D79	334866	2105	4
D82	346783	42	36
D83	340027	2694	30
Avg.	326098.38	6572.19	5.48

TABLE VII
SELECTED IMAGES FROM THE BRODATZ DATABASE WITH MAX. DECOMP. 4

Image	Dyadic	Reduc. RDO	F.reduc. RDOH+
D101	295436	7117	0
D102	288898	13742	0
D103	345236	2871	0
D105	346221	6897	14
D106	356627	6891	41
D107	324110	6939	0
D109	333924	19822	0
D16	390561	14524	73
D21	362293	12675	321
D49	250951	14784	2
D53	319108	4987	114
D6	293454	5900	91
D64	275417	236	8
D67	317239	8655	0
D68	286124	1258	0
D76	327510	152	8
D77	365733	5391	46
D78	347068	2628	24
D79	334867	2093	16
D82	346751	41	0
D83	340033	2669	61
Avg.	326074.33	6679.62	39.00

TABLE VIII
SELECTED IMAGES FROM THE BRODATZ DATABASE WITH MAX. DECOMP. 5

Image	Dyadic	Reduc. RDO	F.reduc. RDOH+
D101	295440	7121	0
D102	288904	13748	0
D103	345252	2887	0
D105	346225	6898	17
D106	356641	6894	52
D107	324135	6964	0
D109	333945	19843	0
D16	390574	14470	140
D21	362281	12263	885
D49	250956	14725	86
D53	319116	4764	348
D6	293437	5808	183
D64	275440	249	18
D67	317252	8668	0
D68	286137	1271	0
D76	327522	153	19
D77	365739	5370	73
D78	347078	2628	0
D79	334871	2093	20
D82	346772	62	0
D83	340030	2678	54
Avg.	326083.19	6645.57	90.24

TABLE IX
SELECTED IMAGES FROM THE BRODATZ DATABASE WITH MAX. DECOMP. 6

Image	Dyadic	Reduc. RDO	F.reduc. RDOH+
D101	295454	7135	0
D102	288913	13757	0
D103	345271	2906	0
D105	346234	6898	26
D106	356654	6907	52
D107	324153	6982	0
D109	333959	19857	0
D16	390582	14457	161
D21	362292	12193	966
D49	250968	14670	153
D53	319127	4739	384
D6	293450	5798	206
D64	275456	265	18
D67	317268	8684	0
D68	286151	1285	0
D76	327536	167	19
D77	365754	5385	73
D78	347092	2642	34
D79	334883	2105	20
D82	346786	76	0
D83	340046	2683	65
Avg.	326096.62	6647.19	103.67

TABLE X
SELECTED IMAGES FROM THE BRODATZ DATABASE WITH MAX. DECOMP. 7

formance gains (see figures 9 and 10). RDO[H+] performs also good on the entire Brodatz database (see figures 11 and 12). Thus for textured data, best basis selection in a JPEG2000 RDO sense can be recommended. The header data has a negligible influence at a decomposition depth of 5 (visually indistinguishable) and thus is omitted in the figures.

On fingerprint data RDO WPB selection also achieves a significantly improved performance; interestingly, the performance gains are more significant in the higher quality region

Max. decomp.	Dyadic	Reduc. RDO	F.reduc. RDOH+
4	59743.93	296.70	0.00
5	59691.11	288.11	0.00
6	59690.60	288.66	1.86
7	59703.36	299.36	3.94

TABLE XI
AN OVERVIEW FOR THE FVC2004 DB1 (B) DATABASE FOR A MAX. DECOMP. 4 TO 7

Dyadic PSNR	RDO PSNR Inc.	RDOH+ PSNR F.inc.
16.07	-2.21	6.39
18.64	5.76	2.24
21.06	6.70	1.43
23.33	6.81	1.37
25.43	6.60	1.03
26.94	6.71	0.75
28.11	6.78	0.56
29.29	6.51	0.53
30.48	6.28	0.54
31.60	6.33	0.39
33.04	5.88	0.35
33.54	6.10	0.40
34.30	6.27	0.39
34.98	6.62	0.37
35.89	6.84	0.34
36.96	6.78	0.27
38.17	6.43	0.26
38.77	6.66	0.25

TABLE XII
ARTIFICIAL, MAX. DECOMP. 7

starting a PSNR of 34dB (see figure 13). The VIF evaluation supports the good performance for fingerprint data (see figure 14). Additionally to our best WPB selection algorithms for JPEG2000 we evaluated the compression performance of the WSQ-WPB on the FVC2004 database, which led to even worse results than the standard dyadic decomposition.

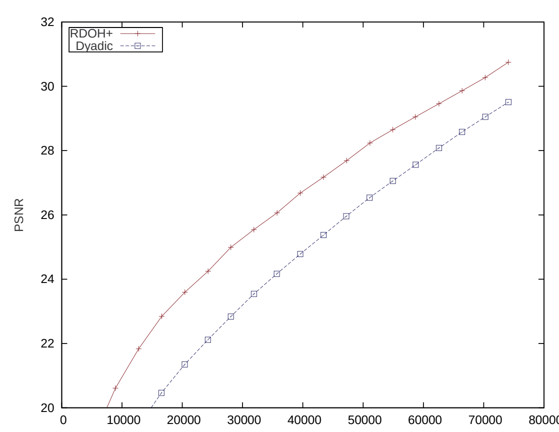


Fig. 9. A 20% subset of the Brodatz database: PSNR evaluation

Dyadic VIF	RDO VIF Inc.	RDOH+ VIF F.inc.
0.13	-0.07	0.18
0.21	0.17	0.10
0.28	0.24	0.04
0.34	0.24	0.05
0.39	0.25	0.01
0.45	0.22	0.02
0.49	0.23	0.02
0.51	0.23	0.01
0.53	0.23	0.01
0.56	0.22	0.01
0.64	0.18	0.01
0.66	0.17	0.01
0.68	0.18	0.00
0.69	0.19	0.01
0.71	0.19	0.01
0.74	0.17	0.01
0.80	0.13	0.00
0.82	0.12	0.00

TABLE XIII
ARTIFICIAL, MAX. DECOMP. 7

Dyadic PSNR	RDO PSNR Inc.	RDOH+ PSNR F.inc.
22.78	-3.77	3.87
24.67	-0.24	1.14
25.85	0.39	0.83
27.17	0.45	0.54
28.37	0.10	0.83
29.21	0.32	0.66
30.09	0.53	0.40
30.83	0.57	0.26
31.51	0.55	0.36
32.28	0.66	0.22
32.82	0.70	0.19
33.37	0.67	0.20
33.89	0.71	0.20
34.43	0.67	0.13
34.87	0.69	0.09
35.33	0.70	0.08
35.70	0.79	0.12
36.30	0.71	0.11

TABLE XIV
BARBARA, MAX. DECOMP. 7

VI. CONCLUSION

Rate distortion optimal (RDO) wavelet packet basis (WPB) selection for JPEG2000 has been presented. In this work, the RDOH+WPB algorithm has been presented and discussed; this algorithm enables the concise selection of the best WPB in the JPEG2000 coding framework. Thus we are are able to report the upper bound of performance improvements achievable with custom isotropic wavelet packet bases in JPEG2000.

In terms of compression performance our results show that for highly textured data, RDO wavelet packet bases perform significantly better than the dyadic decompostion. Even in the lossless case compression performance improvements can be reported. For higher decomposition depths the consideration of the header cost in JPEG2000 WPB optimization is favorable.

Dyadic VIF	RDO VIF Inc.	RDOH+ VIF F.inc.
0.11	-0.08	0.09
0.16	-0.02	0.04
0.21	0.00	0.03
0.25	0.02	0.02
0.31	-0.01	0.03
0.35	-0.01	0.03
0.38	0.01	0.01
0.40	0.03	0.01
0.45	0.03	0.01
0.50	0.02	0.00
0.51	0.03	0.01
0.54	0.03	0.01
0.57	0.02	0.01
0.59	0.02	0.01
0.60	0.05	0.00
0.62	0.07	0.00
0.65	0.06	0.00
0.70	0.03	0.00

TABLE XV
BARBARA, MAX. DECOMP. 7

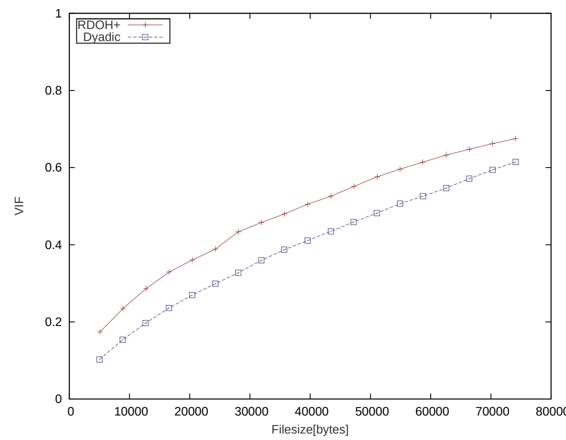


Fig. 10. A 20% subset of the Brodatz database: VIF evaluation

The improved PSNR performance of RDO[H+] WPB also leads to improvements in terms of state-of-the-art objective image quality assessment. For highly textured image data significant compression performance improvements are shown.

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 Transactions on Image Processing*, 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.

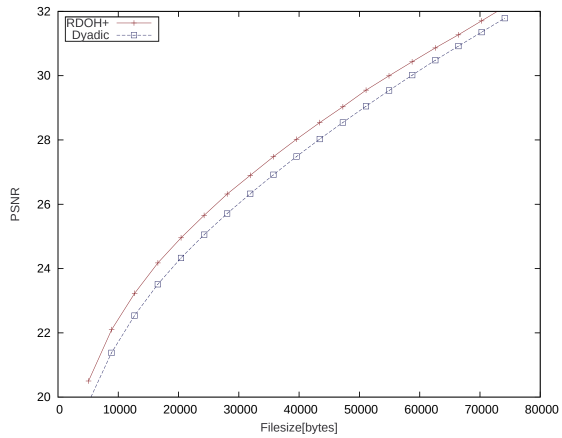


Fig. 11. The Brodatz database: PSNR evaluation

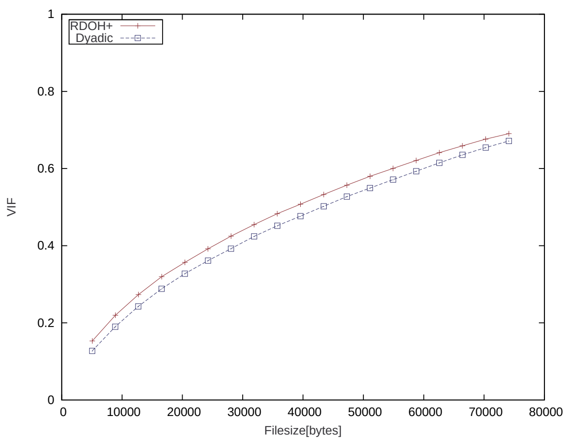


Fig. 12. The Brodatz database: VIF evaluation

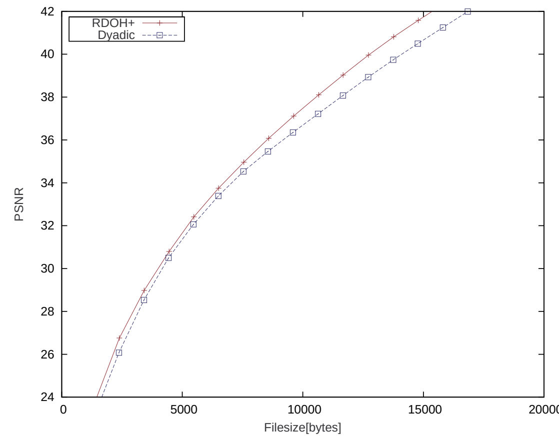


Fig. 13. FVC2004 DB4B: PSNR evaluation

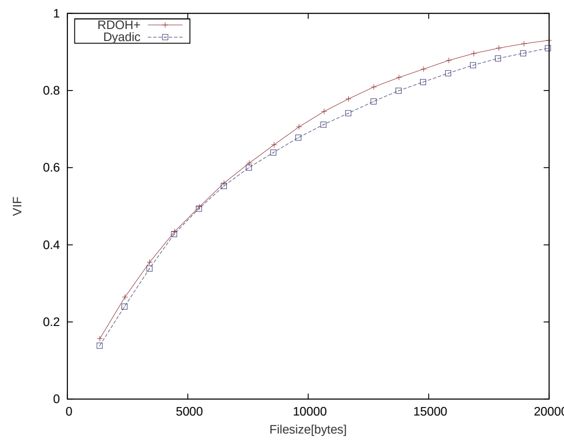


Fig. 14. FVC2004 DB4B: VIF evaluation

[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] Thomas Stütz, Bernhard Mühlbacher, and Andreas Uhl, "Best wavelet packet bases in a JPEG2000 rate-distortion sense: The impact of header data," in *Proceedings of the IEEE International Conference on Multimedia & Expo, ICME '10*, Singapore, July 2010, pp. 19–24, IEEE.

[17] B. Mühlbacher, T. Stütz, and A. Uhl, "JPEG2000 Part 2 wavelet packet subband structures in fingerprint recognition," in *Visual Communications and Image Processing 2010 (VCIP'10)*, P. Frossard, H. Li, F. Wu, B. Girod, S. Li, and G. Wei, Eds., Huang Shan, China, July 2010, number 7744 in Proceedings of SPIE, pp. 77442C–1 – 77442C–10, SPIE.

[18] P. Vandewalle, J. Kovacevic, and M. Vetterli, "Reproducible research in signal processing - What, why, and how," *IEEE Signal Processing Magazine*, vol. 26, no. 3, pp. 37–47, Mar. 2009.

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

[20] 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.

[21] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling, *Numerical Recipes in C : The Art of Scientific Computing*, Cambridge University Press, October 1992.

[22] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, May 2006.