

LOSSY COMPRESSION BY POST-TRANSFORMS IN THE WAVELET DOMAIN

Xavier Delaunay⁽¹⁾, Carole Thiebaut⁽²⁾, Emmanuel Christophe⁽²⁾, Rosa Ruiloba⁽³⁾,
Marie Chabert⁽⁴⁾, Vincent Charvillat⁽⁴⁾, Géraldine Morin⁽⁴⁾

⁽¹⁾*NOVELTIS/CNES/TéSA*

*Parc Technologique du Canal, 2 av. de l'Europe,
31520 Ramonville-Saint-Agne (France)
Email: xavier.delaunay@noveltis.fr*

⁽²⁾*CNES*

*Centre spatial de Toulouse, 18 av. Edouard Belin,
31401 Toulouse Cedex 9 (France)*

⁽³⁾*NOVELTIS*

*Parc Technologique du Canal, 2 av. de l'Europe,
31520 Ramonville-Saint-Agne (France)*

⁽⁴⁾*IRIT/ENSEEIH*

*2 rue Charles Camichel,
31071 Toulouse Cedex 7 (France)*

ABSTRACT

Current state-of-the-art on-board image compression systems are based on the wavelet transform. However, redundancies can still be found between the wavelet coefficients. To exploit these remaining redundancies and increase the compression efficiency, the new post-transform compression scheme applies a second transform on blocks of coefficients. This paper proposes to adapt the post-transform process to on-board constraints using the low complexity Hadamard post-transform and describes how it can be efficiently included in the bit-plane encoder of the CCSDS (Consultative Committee for Space Data Systems) recommendation. Compression performance is provided for each adjustment made to the post-transform process.

INTRODUCTION

With the increasing resolution of Earth observation images, lossy compression has become inescapable. Currently state-of-the-art image coders use the wavelet transform to decorrelate the data. Then, smartly designed entropy coders are in charge of exploiting remaining redundancies between wavelet coefficients [1]. This is the case of EBCOT (Embedded Block Coding with Optimal Truncation points), the coder used in JPEG2000 [2]. In this coder, wavelet subbands are independently coded bit-plane by bit-plane. Each bit-plane is coded in three coding passes. Dependencies between high wavelet coefficients produced near the edges of the image are exploited in the so-called “significance propagation pass”. Zero-tree coders, such as SPIHT (Set Partitioning in Hierarchical Trees) [3] or the BPE (Bit-Plane Encoder) of the CCSDS recommendation [4], use another strategy. They exploit intra- and inter-band dependencies by signalling trees of non-significant wavelet coefficients.

Facing those remaining dependencies between wavelet coefficients, a new approach has been proposed by Peyré and Mallat in [5] to enhance compression efficiency. The idea is still to use the wavelet transform since it has good properties such as multi-resolution, localised basis element and critical sampling which provide good energy compaction. But following this wavelet transform, post-transforms, called bandelet transforms, are applied on small blocks of wavelet coefficients. The bandelet transforms are adaptively chosen in a dictionary known by both the encoder and the decoder. On each block, the best transform is selected based on the minimization of a rate-distortion criterion. Blocks of wavelet coefficients are thus linearly transformed using the bandelet transform with which the compression will be the most efficient. In [6], we have already shown, using a simple adaptive arithmetic coder and a modified version of the bandelet transform, that this compression scheme out-performs the wavelet transform. Moreover, depending on the computational capabilities available on-board, the dictionary can contain one or more post-transforms, from the simple Hadamard transform, to more complex transforms in terms of computational complexity.

In this paper, we study the post-transform compression scheme and describe how it can be included in the BPE of the CCSDS recommendation [4] without huge modifications. First we describe the compression scheme using post-transforms as it was originally proposed. Next, we briefly review the coding of wavelet coefficients in the BPE and propose low complexity solutions to adapt the post-transform compression scheme to the BPE. The first step of this adaptation is a modification of the post-transform selection criterion. Then, we describe how to efficiently include the post-transform information in the bit-stream produced by the BPE in particular for the grandchildren coefficients. Indeed, post-transformed blocks are of size 4x4. Thus by applying the post-transforms only on the high-frequency subbands of the wavelet transform, it can be easily included in the BPE.

COMPRESSION USING POST-TRANSFORMS

In this section, the post-transform compression scheme is described. First, we explain the general principles of the post-transform. Then, we give examples of dictionaries of bases that can be used for the post-transform. Finally, we detail how the bases are selected to post-transform blocks of wavelet coefficients.

Principle of the Post-Transforms

Post-transformations of blocks of wavelet coefficients have first been proposed by Peyré in his thesis [7]. It was derived from the bandelet transform theory as a practical solution to apply it on natural images. The bandelets aim at exploiting redundancies between high magnitude wavelet coefficients produced near the edges in the image. In the bandelet transform, areas following the edges are reshaped and a second transform is applied so as to reduce the redundancies in this area [5].

To decrease the complexity of this process, Peyré proposed to reduce the areas to small blocks of fixed size: 4x4 wavelet coefficients. He also defines a set of directional bases so as to still take into account the possible directions of the edges in the image. The blocks of wavelet coefficients are then transformed in one of the basis of this set. This basis is chosen so as to minimize a rate-distortion criterion after quantization. This is what we call the post-transform process.

This process can be included in a compression scheme as illustrated by Fig. 1. The first step is a classical 2-D multilevel Discrete Wavelet Transform (DWT). In this paper, we use three stages of the 9/7 float transform as recommended by the CSSDS in [8]. Next, blocks of 4x4 wavelet coefficients are post-transformed using bases defined in a dictionary. The coder needs to indicate which basis has been used to post-transform each block. For this reason, side information is necessary. The side information may be viewed as a map of indexes identifying in which basis of the dictionary each block has been post-transformed. Finally, this side information and the post-transformed blocks are entropy coded to form the compressed bit-stream.

The post-transform process on each block of wavelet coefficient is detailed on Fig. 2. On this figure, the block of wavelet coefficient denoted by f is post-transformed in the N_b bases B_b included in the dictionary D . The N_b post-transformed representations of the block f are denoted by f^b with $b \in [1, N_b]$. Those representations are then quantized using a nearly uniform quantization of step q . Finally, the best post-transformed representation of the block f is selected based on a Lagrangian rate-distortion criterion minimization. The quantized coefficients of this best post-transformed representation denoted by $f_q^{b^*}$ are those encoded by the entropy coder and the side information for this block is the index b^* which identify the post-transform used on this block.

Blocks of wavelet coefficients f may not be post-transformed if the quantized wavelet representation f_q minimizes the Lagrangian rate-distortion criterion compared to all the other post-transformed representations f_q^b . In the post-transform framework, this corresponds to a post-transform in the canonical basis. As this post-transform does not require any computation, in the following, the canonical basis will always be implicitly included in the dictionaries of bases. Note that even if this post-transform does not require any computation, the basis index $b=0$, which identify this post-transform, must be transmitted as side information to indicate the blocks on which it is selected.

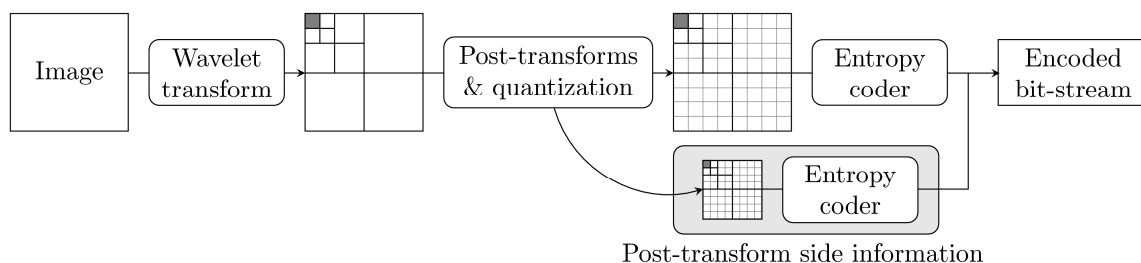


Fig. 1. Compression scheme using post-transforms

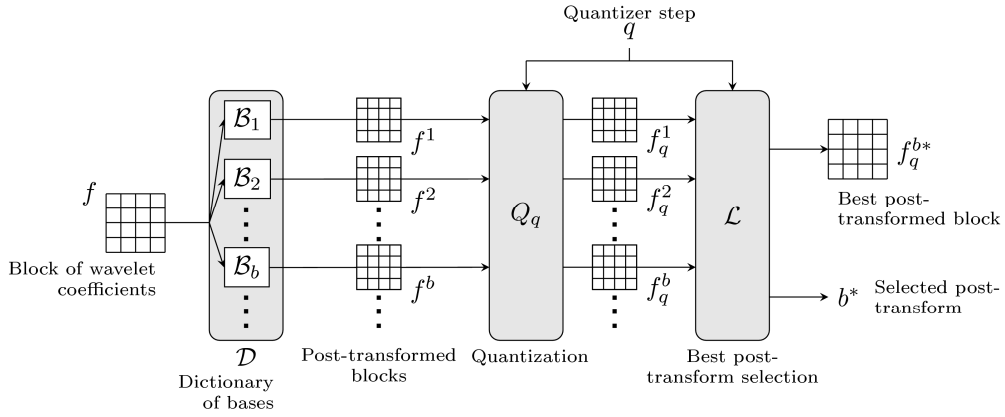


Fig. 2. Post-transform process on one block of wavelet coefficients

Dictionaries of Bases

Post-transforms are linear transforms of blocks of 16 coefficients. The post-transform can thus be viewed as a change of basis in \mathbb{R}^{16} . One post-transform basis B_b is composed of 16 vectors of \mathbb{R}^{16} denoted by ϕ_m^b with $m \in [1, 16]$. Orthonormal bases are preferred so that the distortion after quantization can be computed in the transformed domain. The post-transformed representation f^b of one block of wavelet coefficient f is then simply obtained by the linear transformation:

$$f^b = \sum_{m=1}^{16} \langle f, \phi_m^b \rangle \cdot \phi_m^b = \sum_{m=1}^{16} a[m] \phi_m^b. \quad (1)$$

In [7], Peyré has built a dictionary of 12 directional bases by grouping wavelet coefficients in the same directions as illustrated on Fig. 3. Post-transform bases of \mathbb{R}^{16} are obtained using discrete orthogonal Legendre polynomial bases on each grouping. Fig. 4 illustrates the directional bases #1, #2 and #3. On this figure, each vector of \mathbb{R}^{16} is represented as a 4x4 block. The color gray corresponds to zero coefficients. It can be seen that each vector corresponds to a grouping in one of the grouping configurations displayed on Fig. 3.

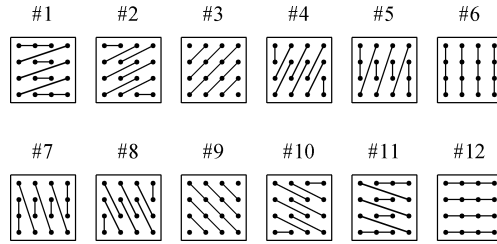


Fig. 3. Directional groupings of wavelet coefficients

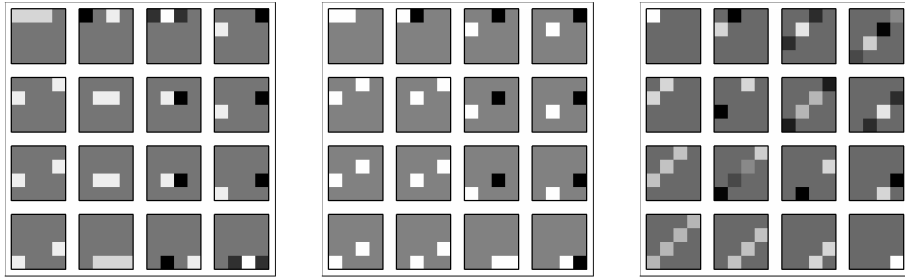


Fig. 4. Directional bandelet bases #1, #2 and #3. Vectors of \mathbb{R}^{16} are represented as 4x4 blocks

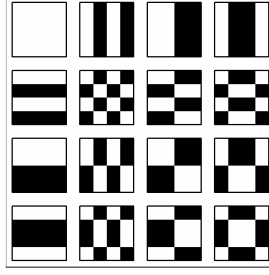


Fig. 5. Hadamard basis

However, as natural images are not only composed of homogeneous area separated by edges but also by textures, additional bases are included in this dictionary: two different Haar bases and the Discrete Cosine Transform (DCT) basis. Those bases may be selected on textured area on which the directional bases are ineffective.

The whole dictionary is thus composed of 15 bases: 12 directional bases and the 3 additional bases. To select the best basis, each block must be post-transformed in the 15 bases of that dictionary. This might not be feasible on-board satellites. Nevertheless, it is possible to reduce the complexity by reducing the number of bases of the dictionary. For example a very simple dictionary can be composed of the Hadamard basis only. Recall that the canonical basis is still in competition. Hadamard basis is represented on Fig. 5. Post-transforms in this basis are very simple since they can be computed by simple additions and subtractions of wavelet coefficients without any multiplication. Only a 2-bit shift (division by 4) is necessary for the normalization. Moreover, all computation can be done in integers. Because of complexity concern, the dictionary of post-transform bases used with the BPE will only be composed of the Hadamard basis.

Best Basis Selection

The post-transformed representation of each block is selected by minimization of the Lagrangian rate-distortion cost:

$$L(f_q^b) = D(f_q^b) + \lambda \cdot R(f_q^b), \quad (2)$$

where $D(f_q^b)$ is the square error between the post-transformed coefficients f^b and f_q^b after quantization with a quantization step q , λ is a Lagrangian multiplier defined in the following, and $R(f_q^b)$ is an estimation of the bit-rate needed to encode the quantized coefficients f_q^b . This bit-rate $R(f_q^b)$ is composed of two terms:

$$R(f_q^b) = R_C(f_q^b) + R_b, \quad (3)$$

where $R_C(f_q^b)$ is the bit-rate associated to the post-transformed coefficients and R_b is the bit-rate needed to indicate the basis $\#b$. $R_C(f_q^b)$ is evaluated based on the zero-order entropy of the quantized post-transformed coefficients and R_b is computed by:

$$R_b = -\log_2 \Pr(b), \quad (4)$$

where $\Pr(b)=0.5$ if $b=0$ that is if no post-transform is applied and $\Pr(b)=0.5/N_b$ for $b \in [1, N_b]$. The Lagrangian multiplier optimization has been studied in [9] and is optimized as a function of the quantization step q by:

$$\lambda = 0,115 \cdot q^2. \quad (5)$$

Compression Results

An adaptive arithmetic coder is used to encode both the post-transformed coefficients and the side information of the chosen post-transform bases on each block. For the results reported in this paper, only the blocks of the first scale of the wavelet transform, i.e. HL_1 , LH_1 and HH_1 , are post-transformed. This choice had been made because with the BPE, the post-transform process can only be applied on the finest scale where 4×4 blocks can be found. Fig. 6 compares compression results of the bandelet transform with the whole dictionary of 15 bases to the compression results of the wavelet transform. These are mean compression results obtained on eleven large 12-bit Earth observation images. These images are simulated PLEIADES images with 70cm resolution and PELICAN airborne sensor images with 20cm resolution. The adaptive arithmetic coder directly encodes the quantized value of the wavelet or bandelet coefficients.

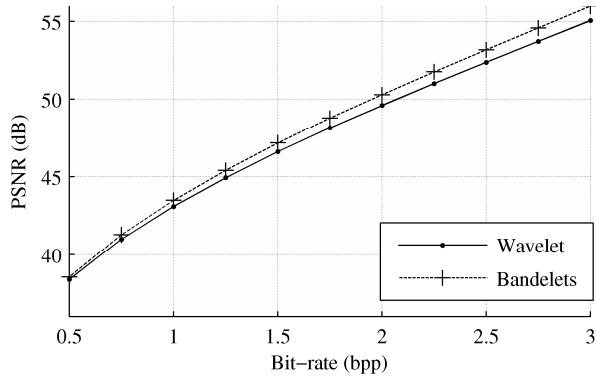


Fig. 6. Mean compression results obtained with the wavelet transform and the bandelet transform with the dictionary of 15 bases

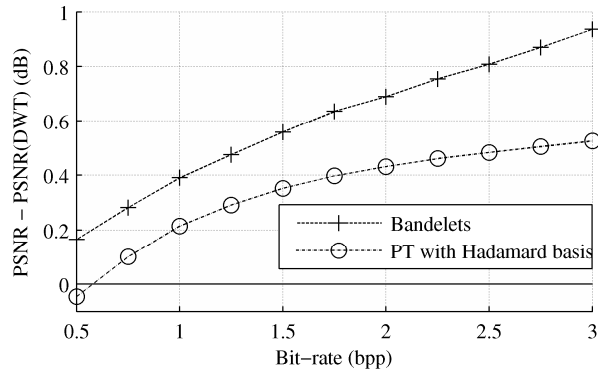


Fig. 7. Bandelet performance with 15 bases and post-transform performance with the Hadamard basis only relative to the wavelet transform results

On Fig. 7, the wavelet transform results are taken as reference. The differences in PSNR to this reference are plotted. It can be seen that the bandelet transform with 15 bases enhance the compression results of 0.2dB at low bit-rates and of 0.9 dB at high bit-rates. Fig. 7 also compares the results obtained with a dictionary composed only of the Hadamard basis. Using this simple dictionary, compression results are increased up to 0.5dB.

USING THE POST-TRANSFORM COMPRESSION SCHEME WITH THE CCSDS BIT-PLANE ENCODER

In this section, we consider the use of the BPE after the post-transform. First, we adapt the post-transform selection process to a bit-plane encoder. Then, we briefly describe how the wavelet coefficients are encoded in the BPE and propose solutions to encode the post-transformed coefficients using the BPE.

Post-Transform Basis Selection for a Bit-Plane Encoder

We have previously seen that the post-transform process requires the knowledge of the quantization step q so as to compute the Lagrangian rate-distortion criterion for the best post-transform basis selection. The first difficulty to adapt the post-transform compression scheme to a bit-plane encoder is that the quantization step q is not defined. Indeed, as the coefficients are progressively encoded until a target bit-rate R_T is reached, the number of bit-planes encoded may be different from one coefficient to the other. At high bit-rates, under high resolution hypothesis, a quantization step q may be inferred from the targeted bit-rate R_T by:

$$q = 2^{H(X) - R_T}. \quad (6)$$

However this expression requires the knowledge of the entropy of the image denoted by $H(X)$. This entropy is not known in general. Moreover, as the produced bit-stream is embedded, it may be truncated at any lower bit-rate. Thus, it is not possible to infer a quantization step from the targeted bit-rate and the post-transform basis selection process needs to be modified so as the quantization step q does not intervene.

The goal of the basis selection process is to find the basis b in which the representation f^b of f minimizes both the distortion and the needed bit-rate. Minimizing the distortion amounts to maximizing the energy kept in the decoded block. Therefore, energy compaction must be favoured, that is to say, most of the energy has to be compacted on a small number of coefficients. Those coefficients will be encoded in the early bit-planes and thus a good amount of the energy of the block will be available at the early stages of the decoding process. But to minimize the bit-rate needed, it is desirable to have many low amplitude coefficients. Indeed, those coefficients will not be significant until the later bit-planes and will not cost much to encode before those bit-planes are reached. The rate-distortion criterion can then be interpreted as a sparsity criterion: it favours representations with a few high energy coefficients and many low amplitude coefficients. A simple sparsity criterion which does not depend on the quantization step q is the l_1 norm:

$$\|f^b\|_1 = \sum_{m=1}^{16} |\langle f, \phi_m^b \rangle| = \sum_{m=1}^{16} |a[m]|, \quad (7)$$

and the basis b^* selected is :

$$b^* = \arg \min_{b \in [1, N_b]} (\|f^b\|_1). \quad (8)$$

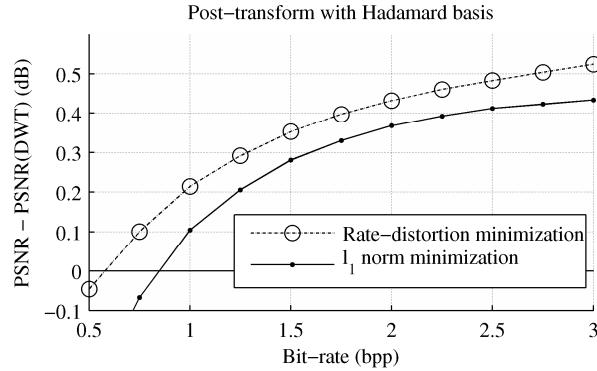


Fig. 8. Comparison of compression results using two different criteria for the selection of the post-transform bases: the rate-distortion minimization criterion and the l_1 norm minimization criterion

By minimizing the sum of the absolute coefficients, this criterion favours representations f^b with many low amplitude coefficients. Thus it minimizes the bit-rate needed to encode f^b . Since the post-transforms used are all orthonormal, the energy of f is the same as the energy of f^b . The search of the representation f^b with the smallest l_1 norm is therefore equivalent to the search of the representation f^b with maximum variance of absolute coefficients. Then, representations with high magnitude coefficients are selected and the distortion is also minimized since a good amount of energy is recovered in the early bit-planes.

Fig. 8 compares the results obtained using the Lagrangian rate-distortion minimization to the one obtained using the l_1 norm minimization. These are post-transform compression results using the Hadamard basis only. Blocks are thus either post-transformed in this basis or not post-transformed. The encoder used is the adaptive arithmetic coder. At bit-rates greater than 1bpp, compression results obtained with the l_1 minimization are less than 0.1dB lower than the ones obtained with the rate-distortion minimization. At low bit-rates, results obtained with the l_1 minimization are not good. This is due to the side information cost. Indeed, with this new criterion, the post-transform basis selection does not depend on the bit-rate and the mean side information cost is 1 bit per block even after the adaptive arithmetic coding. This cost is significant at low bit-rates: it represents 9.3% of the total bit-rate at 0.5bpp and only 1.6% of the total bit-rate at 3bpp. On the contrary, with the Lagrangian rate-distortion criterion, the post-transform basis selection does depend on the quantization step and thus on the bit-rate. The mean side information cost is still 1 bit per block at 3bpp, but at 0.5bpp, it is reduced to 0.6 bit per block and represents only 5.4% of the total bit-rate. However, in a next section, it will be shown that this side information cost can be reduced at low bit-rates using a progressive encoder.

The Bit-Plane Encoder of the CCSDS Recommendation

The BPE is a zero tree coder in which one tree is defined by 64 coefficients: one coefficient from the LL subband, 3 parents p_i from the HL₃, LH₃ and HH₃ subbands, 3 sets C_i of four children from the HL₂, LH₂ and HH₂ subbands and three sets G_i of 16 grandchildren from the HL₁, LH₁ and HH₁ subbands. Moreover, the sets G_i are partitioned into four groups H_{ij} of four coefficients, three descendant sets D_i are defined by the concatenation of the sets C_i and G_i and the block set B is defined as the concatenation of the three sets D_0 , D_1 and D_2 .

The coefficients are encoded bit-plane by bit-plane from the most-significant bit-plane determined by the highest magnitude coefficient of the tree (except the LL coefficient), to a least significant bit-plane depending on the targeted bit-rate. At each bit-plane, transition words are used to indicate sets of coefficients which are still not significant: $tran_B$ is a one bit word indicating if all coefficients in the set B , consisting of all the children and grandchildren, are still not significant in the current bit-plane. If not, $tran_D$ indicates the descendant sets D_i in which all coefficients are still not significant. In the significant sets D_i , $tran_G$ indicates the sets G_i of grandchildren coefficients which are still all not significant. Finally in the significant sets G_i , $tran_{H_i}$ indicates the groups H_{ij} in which all coefficients are still not significant.

Post-Transforms of the Grandchildren Coefficients Sets

The grandchildren coefficients sets are blocks of size 4x4. The post-transform can thus be easily applied on those sets. That is why the post-transform was only performed on the finest scale of the wavelet transform in the previous sections.

Side Information Reduction

In order to reduce the side information, the post-transform basis used on each block is only indicated on blocks of grandchildren which have at least one significant coefficient. In the case of post-transforms in one basis only, this information is one bit per block and is not entropy coded. It has previously been shown that using the l_1 minimization criterion, the mean cost of this side information is 1 bit per block even after the adaptive arithmetic coder. This bit indicating the basis used on each significant grandchildren block is embedded in the bit-stream directly after $tran_G$ words with at least one bit 1 indicating a newly significant grandchildren block of coefficient. Fig. 9 shows the results obtained using this strategy with the Hadamard basis as only possible post-transform. It can be observed that at low bit-rates, the results are close to the ones obtained when the side information cost is not taken into account. Indeed, few grandchildren blocks have significant coefficients; therefore the side information is greatly reduced. At high bit-rates, the results are close to the ones obtained when all the side information is coded: nearly all the grandchildren blocks have significant coefficients.

Energy Compaction Exploitation

Even if the post-transform side information is well embedded in the bit-stream, it can be observed on Fig. 9 that the compression results obtained with the post-transforms are less than 0.2dB better than the results of the DWT and even worse at high bit-rates. This is due to the fact that the post-transforms destroy the zero-tree structure: the inter-band and intra-band dependencies existing between non-significant wavelet coefficients. Post-transformed grandchildren coefficients may become significant anywhere in the block without any relationship to the children significance state.

It is possible to limit this loss by exploiting the energy compaction property of the Hadamard transform. In the subband HL, LH and HH, the energy of the block of wavelet coefficients is not equivalently distributed on the vectors of Hadamard basis. These vectors can be sorted by energy contribution. Hence, it is possible to increase the compression efficiency by grouping the post-transform vectors by energy contribution so that the coefficients in a grandchildren group H_{ij} have roughly the same energy. This artificially recreates intra-band dependencies. The vector orders of Hadamard basis are thus depending on the subband as shown on Fig. 11. Note that even if the orders of the vectors are different in the subbands, the side information cost is still 1bit per block since the post-transform can only be done in one basis: the basis adapted to the subband. The compression performance obtained using this strategy is displayed on Fig. 10. This strategy enables a gain in PSNR of more than 0.15dB at high bit-rates. This figure also compares the performance obtained with the DCT basis to the ones of the Hadamard basis. It can be seen that compression results are enhanced of less than 0.05dB. Therefore, if very low complexity is required, the post-transform in Hadamard basis remains a good choice.

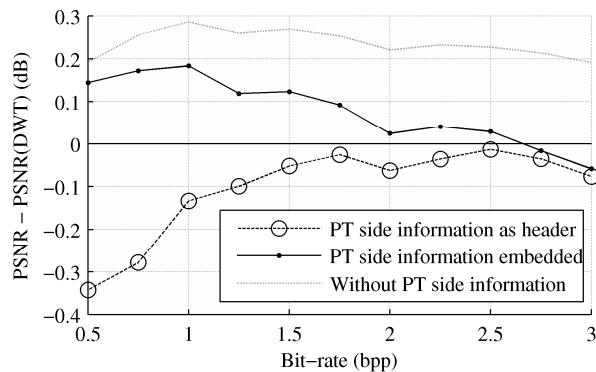


Fig. 9. Post-transform performance when compressed using the BPE, relative to the results of the DWT compressed with the BPE. Comparisons are made for different strategy of coding the post-transform side information: all side information is sent as header of the bit-stream; side information is embedded in the bit-stream and only the useful part is sent; no side information is sent. In this later case, the image cannot be properly recovered

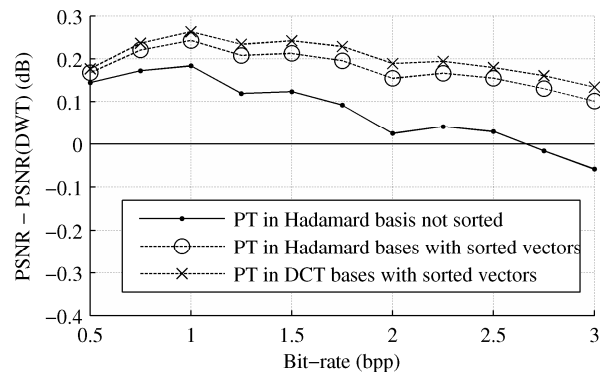


Fig. 10. Comparison of post-transforms performance in Hadamard bases with vectors not sorted and with vectors sorted so as to exploit the energy compaction properties and artificially create groups of grandchildren coefficients with roughly the same energy. Compression results obtained with the Hadamard post-transform are also compared to the results obtained with the post-transform in DCT bases

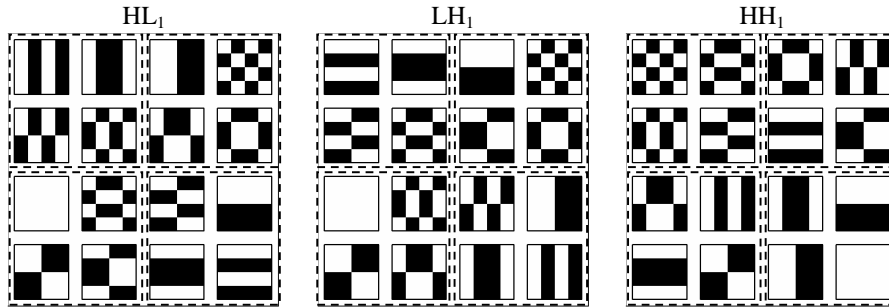


Fig. 11. Hadamard bases with vectors grouped by energy contribution in the subbands HL_1 , LH_1 and HH_1 . Grandchildren groups H_{ij} , as defined in the BPE, are represented by dashed lines

CONCLUSIONS AND FUTURE WORKS

The post-transform is a promising method to increase the performance in compression. In this paper, it has been shown that low complexity solutions can be designed in order to target the use on-board spacecraft. In this purpose, the number of post-transforms in competition has been reduced to one and the retained post-transform, the one in Hadamard basis, is both effective and very simple to compute. Moreover, in order to adapt the post-transform compression scheme to a bit-plane coder, the post-transform selection process has also been modified and the criterion is simpler than the original. The side information indicating the post-transform used on each block has been effectively integrated into the embedded bit-stream so that only the useful information is recovered. Lastly, the energy compaction property of the Hadamard post-transform has been exploited to increase the compression performance.

However, the post-transform destroys the relationships between children and grandchildren. To preserve this relationship exploited by the BPE, blocks of 2×2 children could be post-transformed in Hadamard basis of \mathbb{R}^4 with vectors sorted by energy contribution. Then, the choice would be either to post-transform the whole group of descendant coefficients i.e. the children and grandchildren or no coefficient at all. Hence, the relationships between children and grandchildren would be recreated and even enforced. This would give better results with zero-tree coders such as the BPE.

REFERENCES

- [1] X. Delaunay, M. Chabert, G. Morin and V. Charvillat, "Bit-plane analysis and contexts combining of JPEG2000 contexts for on-board satellite image compression," *Proc. of ICASSP'07*, I-1057-1060, IEEE, April 2007, Honolulu, HI, USA.
- [2] D. Taubman and M. Marcellin. *JPEG2000: Image compression Fundamentals, standards and practice*, Kluwer Academic Publishers, 2001.
- [3] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, pp. 243-250, 1996.
- [4] P-S. Yeh et al. , "The New CCSDS Image Compression Recommendation," *Proc. of IEEE Aerospace Conference*, pp. 1-8, Big Sky, MT, March 5-12, 2005.
- [5] G. Peyré and S. Mallat, "Discrete bandelets with geometric orthogonal filters," *Proc of ICIP'05*, I-658-661, IEEE, Sept. 2005.
- [6] X. Delaunay, M. Chabert, V. Charvillat, G. Morin and R. Ruiloba "Satellite image compression by directional decorrelation of wavelet coefficients," *Proc. of ICASSP'08*, I-1193-1196, IEEE, April 2008, Las Vegas, NV, USA.
- [7] G. Peyré, *Geometrie multi-échelles pour les images et les textures*, Ph. D.thesis, Ecole Polytechnique, 2005.
- [8] CCSDS, *Image Data Compression*, Recommended Standard, CCSDS 122.0-B-1, Blue Book, Nov. 2005.
- [9] X. Delaunay, E. Christophe, C. Thiebaut and V. Charvillat, "Best post-transform selection in a rate-distortion sense", *Proc.of ICIP'08*, IEEE, October 2008, San Diego, CA, in press.