Adapting JPEG2000 bit allocation to preserve features of interest

Cédric Verleysen, Nathalie Merlin and Christophe De Vleeschouwer
ICTEAM institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium
cedric.verleysen@uclouvain.be

***Abstract*—Nowadays, the demands for higher quality and flexibility in accessing compressed images has led to a gradual replacement of the ubiquitous JPEG standard by the recent JPEG2000 compression standard in various domains. Similar to other compression standards, this new algorithm tends to deteriorate some of the intrinsic geometrical features of the image. For example, the direction of the edges are not well conserved after compression. This papers shows that the optimization metric generally used in JPEG2000, the mean squared error (MSE), is responsible for the non-uniform deterioration of the shapes' angle. To address this problem, a general method of modifying the MSE into any other additive metric is presented and applied to the conservation of underlying shapes' directional information. This modification is handled in a way that does not change the intrinsic implementation and functionalities of any JPEG2000 encoder. This general method of balancing the shapes' deterioration builds on the definition of a directional metric and the computation of edge direction based on a modified Hough Transform, which allows to determine the shapes' slopes with an error of less than the degree, even with low PSNR pictures (up to 15dB). Finally, it is illustrated how our contribution increases the performance of the JPEG2000, when all singularities of an image must be conserved in spite of high compression rates.**

***Index Terms*—Image compression; JPEG2000; directional metric; Hough Transform.**

I. Introduction

Currently, the domain of image compression is more and more faced with the conservation of specific image's features instead of global visual quality. This paper investigates how the bit allocation strategy of an encoder can be adapted to better preserve some specific features of interest in the image. More precisely, the proposed method adapts the Rate-Distortion (RD) strategy of a recent image compression ISO standard called JPEG2000 [1], [2]. Both the higher quality of compressed images and the ability to scale them depending on the user's needs make the JPEG2000 more attractive than its predecessor, the JPEG standard, in a growing number of professional markets such as medical imagery [3], space imagery [4], digital cinema [5], audio-visual archives [6], security and remote sensing.

The principal asset of JPEG2000 is linked to the use of a Wavelet Transform to decompose a picture into uncorrelated versions at lower resolutions. After spatially dividing these subband images into code-blocks and coding them into an embedded code-stream, both a specific spatial and spectral information can be extracted from it, according to the user's needs (in term of region of interest, resolution, color, etc.) and available resources (bandwidth, display, etc.). The coding step is supervised by rate allocation mechanisms that aim to minimize the default mean squared error (MSE) for a given compression rate. However, some applications require a conservation of specific features (e.g. singularities, edges, etc.) instead of minimizing the MSE over the whole image.

In general, the alteration of the image features during compression depends (i) on the representation method adopted to define the image by a limited number of coefficients, and (ii) on the distortion metric adopted to control the allocation of rate among those significant coefficients. Hence, two categories of approaches can be identified to tune the impairments resulting from compression, e.g. so as to preserve some specific features compared to others. On the one hand, the representation of the image can be modified, e.g. by replacing the standard JPEG2000 Daubechies wavelets by 2D directional wavelet filters, or by investigating other kinds of representation such as the ones offered by curvelets, ridgelets or even matching pursuits. Then, given the additive property of MSE, RD optimality is achieved simply by balancing the MSE increment per unit of rate over the whole image [7]. However, such modification of the underlying representation of the image dramatically impacts the design and implementation of the encoder, making it incompatible with the original JPEG2000 standard. On the other hand, the distortion metric used to allocate the available bit budget among the wavelet coefficients can be adapted to favour some specific features of interest. This solution is compatible with the JPEG2000 standard definition, and is thus the one promoted in this paper.

A natural way to replace the MSE metric by another one is to express this new metric in the wavelet domain and to modify the encoder's implementation accordingly. This approach has been followed by [8], in which the authors approximate the $SSIM$ (Structural Similarity Metric) in the wavelet domain. However, not all kinds of distortion metrics can be expressed in the form of an (additive) function of wavelet coefficients. In particular, metrics that are supposed to verify the preservation of local features (e.g. edges) in the spatial domain are obviously defined on a narrow spatial support, which induces a wide support in the wavelet domain, but also dependencies between subbands.

This paper proposes a more generic methodology to skew the JPEG2000 rate allocation strategy towards the preservation of some spatially located image features. In essence, the method aims at allocating the available bit budget so as to minimize a distortion metric that is supposed to reflect the degradation of the (local) feature of interest in the image. Therefore, our approach first uses training samples to learn how the local content of an image affects the rate-distortion performance of the encoder, both in terms of the MSE and of the alternative distortion metric. This knowledge is then exploited to tune the RD trade-off targeted by the conventional JPEG2000 MSE-based rate allocation. This adaptation is done as a function of the local content of the image, so as to get the RD optimality in terms of the alternative metric instead of MSE. Interestingly, because our method interacts with the last flexible step of a JPEG2000 encoder, it does not change its intrinsic features and implementation.

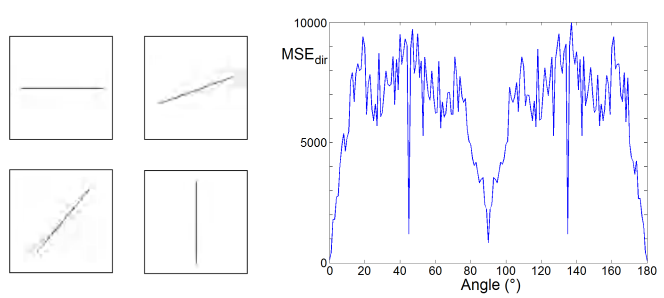
This paper is organized as follows. Section II illustrates a motivation of modifying the MSE metric into another by showing that a traditional JPEG2000 encoder deteriorates some directions of shapes more than others. Section III presents the general strategy associated with the change of metric in JPEG2000. Finally, Section IV elaborates an estimator for the shapes' direction and merges it with the proposed general approach to hedge the MSE bitrate allocation in order to minimize a directional metric.

II. Limitation of the MSE metric in JPEG2000

Despite all its advantages, JPEG2000 strongly deteriorates slanting singularities at high compression rates, as illustrated on Fig.1(a). A way to confirm this observation is to define a directional metric that measures dissimilarity between the shapes' directions of the original I_0 and compressed I_c images:

$$MSE_{dir} = \frac{1}{MN} \sum_{E} (|\arg(\nabla I_0(m,n))| - |\arg(\nabla I_c(m,n))|)^2 \quad (1)$$

with $E = \{(m,n) \mid \|\nabla I_0(m,n)\| \neq 0\}$ and $\#E = MN$ where ∇ denotes the gradient operator in one pixel of an image and $\|\nabla I\|$ the amplitude of a 2D gradient. Using Sobel filters to compute the horizontal and vertical gradients (∇_x and ∇_y respectively) and $\arg(\nabla I) = atan(\frac{\nabla_y}{\nabla_x})$ to determine the direction of the gradient, Fig. 1(b) illustrates how JPEG2000 favours shapes' directions $\theta \ mod \ 45° = 0$ compared to others.



(a) (b)

Fig. 1. (a) Compression of lines of 0°, 20°, 50° and 90° with *OpenJPEG* with rate=70. (b) MSE_{dir} as a function of the lines' angle with rate=40

Preservation of edges might be of primary importance for some applications. This is for example the case when dealing with space imagery [9] or medical images [10]. The next sections will explain how to change the bit allocation mechanisms in JPEG2000 to preserve the edge direction information.

III. Change of quality metric in JPEG2000

In this section, the key principles of a JPEG2000 encoder are surveyed, with a focus on its underlying rate allocation mechanisms. We then explain how bit allocation can be adapted to minimize another (additive) distortion metric than the conventional MSE. This result will be exploited in the Section IV to preserve edges by minimizing a distortion metric that measures the deterioration of shapes' directions during compression.

A. How does a JPEG2000 compression work?

The Wavelet Transform reorganizes the information in resolution levels. Starting from the original full resolution, each successive level transforms the input image into a four times smaller image plus detail coefficients [7]. A grid partition splits the low-pass image and the subbands into code-blocks. Binary representation of the code-block coefficients defines a set of bit-planes. Each code-block is processed independently by the EBCOT algorithm [7] which successively encodes the bit-planes, from the most significant to the least significant bits and send them to a MQ arithmetic coder. Then comes the bit allocation process that, for each code-block, produces an embedded codestream and truncates it at various rates (corresponding to a different amount of bit-planes) to get a RD curve called a convex hull. From this convex hull, it determines the bit-planes to reject to satisfy the specified rate constraint while minimizing the image distortion. Thus, the optimization problem consists in finding the best allocation **b*** (among the set **b** of admissible bits), which assigns the bits b_i^* to code-block i, such that the total distortion (sum of the distortions D_i over N code-blocks) is minimized under a global rate constraint R_t, i.e.:

$$\mathbf{b}^* = \arg\min_{\mathbf{b}} \quad \sum_{i=1}^{N} D_i(b_i)$$

$$\text{subject to} \quad R_t \approx \sum_{i=1}^{N} R_i(b_i)$$

This problem can be solved using Lagrange multipliers:

$$\mathbf{b}^*(\lambda) = \arg\min_{\mathbf{b}} \left(\sum_{i=1}^{N} D_i(b_i) + \lambda \sum_{i=1}^{N} R_i(b_i) \right)$$

If the distortion metric is additive (i.e. the contribution provided by multiple code-blocks to the entire image distortion is equal to the sum of the distortions computed for each individual code-block), the minimization can be done independently on each code-block i:

$$b_i^*(\lambda) = \arg\min_{q} \left(D_i(b_i) + \lambda R_i(b_i) \right)$$

Zeroing the first order derivative shows that the optimal truncation points (that determine the bit-planes to conserve) have to be selected such that

$$\lambda = -\left.\frac{dD_i}{dR_i}\right|_{b_i^*} \qquad \forall \text{ code-block } i \tag{2}$$

In other words, the slope of the tangent to the RD curve at the truncation points has to be the same for all code-blocks. In this way, the increase (decrease) of distortion per removed (added) bit is the same for all code-blocks, and it is not possible to reduce the overall distortion by shifting one bit from one code-block to another. In practice, it means that all the bit-planes of a code-block for which the increment of quality per unit of bits is a smaller slope than λ on the RD convex hull will be rejected.

B. General methodology for a change of quality metric

As we have seen so far, a rate constrained Lagrange optimization can be applied to minimize the image distortion for a given rate constraint. When the distortion is additive over the code-blocks, this optimization implies the selection of truncation points that correspond to the same slope λ on the RD convex hulls of all code-blocks. Hence, the optimization of the encoder with respect to a new metric called $Metric'$ can simply be done by selecting a number of bit-planes b_i^* for the code-block i such that

$$\lambda' = -\left.\frac{dMetric'_i}{dR_i}\right|_{b_i^*} \qquad \forall \text{ code-block } i$$

However, in practice the alternative metric is not necessarily easy to define in terms of wavelet coefficients. Therefore, we propose to express the constraint imposed to $dMetric'_i/dR_i$ in terms of $dMSE_i/dR_i$. Let's define f_i to be a bijective function between $dMetric'_i/dR_i$ and $dMSE_i/dR_i$ for the i^{th} code-block, such as

$$\frac{dMetric'_i}{dR_i} = f_i\left(\frac{dMSE_i}{dR_i}\right)$$

In principle, when $Metric'$ can be computed on a code-block basis, the function f_i for the i^{th} code-block can be derived as depicted in Fig.2, by merging the derivatives of the RD curves obtained for each metric when truncating the bitrate of this code-block at different rates. Once f_i has been estimated for each code-block, it can be used to turn the $Metric'$ contraint into a MSE-based constraint. Formally, the contraint writes

$$\lambda_i = f_i^{-1}(\lambda') \quad \text{with} \quad \lambda_i = -\left.\frac{dMSE_i}{dR_i}\right|_{b_i'^*} \tag{3}$$

where $b_i'^*$ is the $Metric'$-optimal bit allocation of the code-block i and λ_i the specific slope associated only with this code-block.

To reduce the computational load associated to the derivation of the RD curves for all code-blocks of the image, we propose to define clusters of code-blocks that have similar RD behaviors, such that a code-block i is associated to a cluster

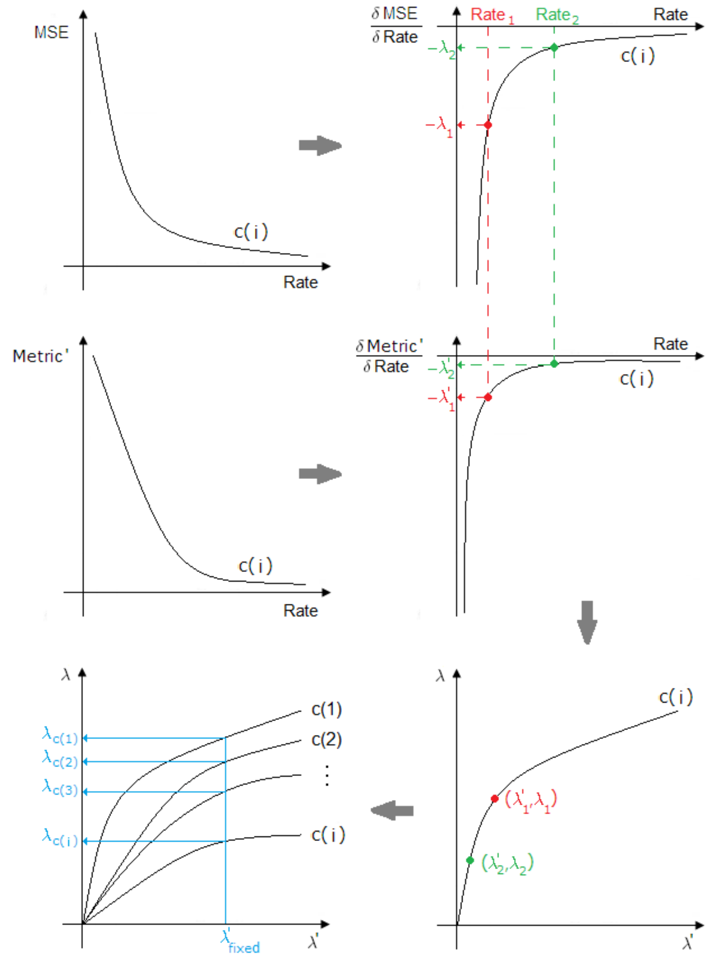


Fig. 2. The relation between $\lambda_{c(n)}$ and λ' for a given kind of code-block can be determined from the RD curves

n according to $n = c(i)$. The definition of a bijective function F_c relating $dMSE/dR$ and $dMetric'/dR$ on the c^{th} cluster and its estimation (based on a representative code-block of the cluster) enables setting f_i to $F_{c(i)}$ for all code-blocks, with $c(i)$ denoting the cluster index associated to code-block i. In practice however, as told above, $Metric'$ might not have a direct interpretation in terms of code-block coefficients, e.g. because it is defined in the pixel domain to capture a signal that is spread on several wavelet subbands. In this case, we decide to estimate the relations F_c on representative blocks (of pixels) of the image, and to restrict our framework to the use of a constant $\lambda_{c(i)}$ for all code-blocks associated to a given block of pixel. In Section IV, a detailled illustration of how such a clustering can be derived in the particular case of a directional metric is presented. We explain there that the representative set of functions F_c associated to the clusters of pixel blocks has only to be computed once, for all kind of images.

IV. JPEG2000 ADAPTATION TO A DIRECTIONAL METRIC

To better preserve edge's directions, this section considers the directional metric of Equation (1) to replace the conventional MSE. In this case, the clustering of the code-blocks should be done according to the direction of the shape they represent, and a representative block of pixels for cluster n is defined by two homogeneous regions separated by a straight line with direction θ_n. The researched relation (3) then becomes:

$$\lambda_{\theta_n} = F_{\theta_n}^{-1}(\lambda') \tag{4}$$

with θ_n denoting the direction of the cluster n. This relation is used for all code-blocks corresponding to a block of pixels assigned to the n^{th} cluster.

In the remaining of the section, we first explain how to assign a direction to an arbitrary block of pixels, and then derive the function $F_{\theta_n}^{-1}$ for a discrete set of directions θ_n.

A. Analysis of shape's directions

The Hough Transform [11] is a common feature extraction method that enables the estimation of the shapes' directions. In principle, this transform projects the direct space (image space) on a feature space (Hough space) that is representative of the analyzed shape. If we assume that a shape can be approximated as a line at the block level, the Hough Transform takes all the points of an extracted shape (thanks to a shape detector) and projects their coordinates (x_s, y_s) on a curve in the 2D space $\{(r, \theta) |\ r \in \mathbb{R},\ \theta \in [0; 180°[\}$ following the relation:

$$r = x_s \cos(\theta) + y_s \sin(\theta)$$

In other words, each shape's point (x_s, y_s) in the direct space is described by a beam of lines crossing this point and having an euclidian distance with the image origin of r and a direction θ. The peak value observed in the Hough space defines the r and θ features of the main line constituting the shape. The application of this method on extracted shapes (using a Laplacian of Gaussian filter with $\sigma = 0.8$ and with a threshold at half of the maximum intensity) shows an estimation error less than one degree on noisy images (gaussian noise) of PSNR as low as 15dB. The tested shapes are entire lines, discontinuous lines and arbitrary non-closed curves.

When dealing with real images blocks, the number of representative edges on a given block is not known a priori. To circumvent the computational complexity related to the accurate computation of all the edges in the block, we restrict ourselves to discriminate between homogeneous blocks (no edge), blocks with a single dominant edge and blocks with complex multi-edges patterns. In practice, a decision tree is applied to estimate if there are zero, one or more shapes. Thanks to the good performances of the LoG filter, the first node computes if there exists a shape on the picture by looking if there is any sufficiently positive pixel in the image of the extracted edge. If it is the case, an observation of the two first maxima in the Hough space predicts at the second node if one edge or more are represented on the block.

B. Determination of the relation between λ' and λ_{θ_n}

Now that we have explained how to estimate the dominant orientation of a block, the function $F_{\theta_n}^{-1}$ should be defined for each orientation θ_n. In practice, the function is estimated in a discrete set of orientations. As explained in Section III-B, this relation can be determined from the RD curves of both the MSE and MSE_{dir}. For a given θ_n, we consider the RD curves corresponding to blocks composed by two homogeneous areas separated by a sharp edge of orientation θ_n. To get rid of the actual position of the edge, the average of the observed distortions is computed on a set of blocks corresponding to translated parallel edges.

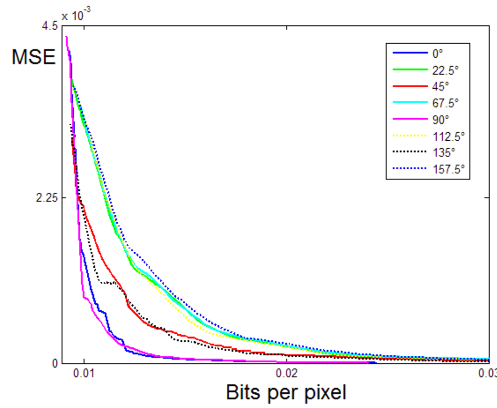


Fig. 3. A code-block representing a slope of θ_n shows the same RD curve than one of θ_m if $\theta_m\ mod\ 45° = \theta_n\ mod\ 45°$

As illustrated for the MSE on Fig.3, the symmetries (in relation to $\theta\ mod\ 45° = 0$) for the MSE and the MSE_{dir} enables the restriction of the previous relation's analysis to the interval $[0°; 45°]$.

By fitting these curves with exponential and hyperbolic regressions and taking their derivatives, the link between λ_{θ_n} and λ' for a given direction is plotted on Fig.4:

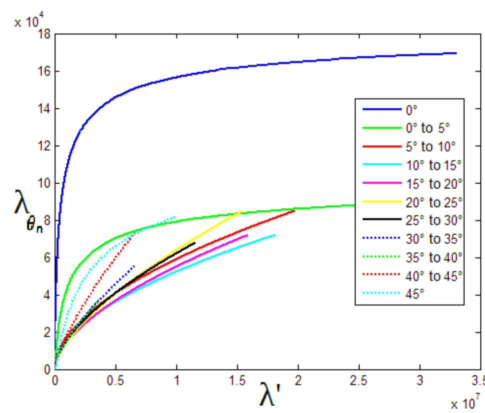


Fig. 4. The application of the methodology presented in Section III-B gives the relation between λ' and λ_{θ_n}

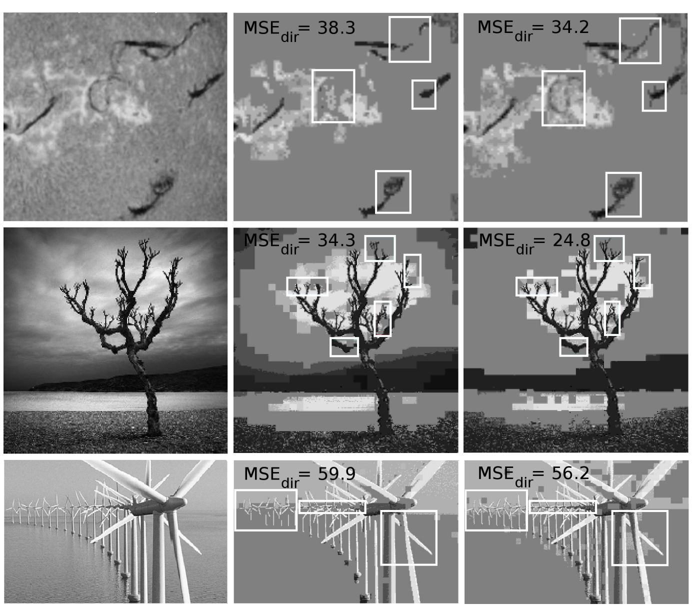
This figure confirms our intuition: the update of λ on each code-block is done in such a way as to compensate the (dis)advantage that JPEG2000 gives to certain angles.

Specifically, for a given λ', directions that are close from $\theta \ mod \ 45° = 0$ are assigned to a larger λ, which ends up in rejecting more bit-planes for those blocks compared to what JPEG2000 would traditionally do. This is in accordance with the observation made in Fig.1, which shows that the MSE preserve too much vertical and horizontal directions compared to others. Knowing this information, the rule for updating the truncation points is the following:

- Zero shapes: the code-block can be highly compressed, as a code-block with direction $0°$ would be.
- One shape: the code-block truncation point has to be adapted according to the direction of the shape.
- More than one shape: because it is difficult to know the worst case direction, the truncation point is fixed as though if it should contain the worst direction.

C. Validation on specific and general images

The proposed adaptation of JPEG2000 to the directional metric (1) has been tested on real case images. As a first application, space imagery is a prominent domain in which a strong image compression is always requested to reduce the energy consumption of the image transmission from the satellite to the Earth. Moreover, preservation of the image's edges is often essential to enable a physical interpretation of the photographed space phenomena. The first row of Fig.5 illustrates that the proposed method satisfies this request for the next solar mission of the European Space Agency (ESA) [9] because the solar singularities (framed in white) are better preserved with our method than with a traditional JPEG2000 compression. The second and third rows of Fig.5 show that the generality of the method enables a better preservation of edges' directions on all kinds of images.



Original image | Optimized MSE | Optimized MSE_{dir}

Fig. 5. A comparison between the usual and adapted JPEG2000 compression shows that the proposed method enables a better preservation of edges' directions for all kinds of images (confirmed by the values of MSE_{dir})

V. Conclusion

In this paper, the impact of the minimization of the MSE metric in the JPEG2000 image compression standard has been studied. The investigation has shown that this distortion metric might be an inappropriate choice if the user aims to preserve some features of interest in the image. As an illustration, the definition and use of a directional metric has shown that JPEG2000 tends to preserve singularities with angles $\theta \ mod \ 45° = 0$ better than others. The main contribution of this paper is an adaptation of the rate-distortion optimization mechanisms of JPEG2000 in order to replace the usual MSE by an other (additive) metric without changing the intrinsic proprieties of any JPEG2000 encoder. In essence, a general methodology that corrects the code-blocks truncation points to preserve the features of interest during compression has been proposed. The application of this general method to the conservation of the shapes' directionality has led to

i) the application of the Hough Transform to estimate edges' directions with a precision higher than one degree for images with PSNR up to 15dB. The Hough coefficients are then used to classify the image blocks according to their content in terms of edges.

ii) the exploitation of the edge content information of a block to adapt the JPEG2000 bitstreams truncation points in a way that optimizes the new directional metric, instead of the conventional MSE.

The C implementation of these steps and their application on real images have demonstrated the effectiveness of our method: the horizontal and vertical edges, which used to be favored by a traditional MSE-based bit allocation, have now transfered a part of their precision bits to the coefficients describing edges that where penalized by the MSE optimization.

Acknowledgment

The authors thank L. Jacques and A. Descampe for their insightful advices. The research reported herein was partly funded by the Belgian NSF (F.N.R.S and F.R.I.A).

References

[1] M. Rabbani and R. Joshi, *An overview of the JPEG2000 still image compression standard*, Signal Processing: Image Com., 17:3-48, 2002.

[2] D. Taubman and M.W. Marcellin, *JPEG2000: Image compression fundamentals standards and practice*, Kluwer Academic, Boston, 2002.

[3] D.H. Foos, E. Muka, R.M. Slone, B.J. Erickson, M.J. Flynn, D.A. Clunie, L. Hildebrand, K.S. Kohm and S.S. Young, *JPEG2000 compression of medical imagery*, SPIE Proc., 3980:85-96, 2003.

[4] D. Van Buren, *A high-rate JPEG2000 compression system for space*, IEEE Aerospace Conf., pp.1-7, 2005.

[5] S. Fossel, G. Fottinger and J. Mohr, *Motion JPEG2000 for high quality video systems*, IEEE Trans. on Consumer Elec., 49(4):787-791, 2003.

[6] J. Janosky and R.W. Witthus, *Using JPEG2000 for enhanced preservation and web access of digital archives*, IST Archiv. Conf., pp.145-149, 2004.

[7] D. Taubman, *High performance scalable image compression with EBCOT*, IEEE Trans. on Image Processing, 9(7):1158-1170, 2000.

[8] T. Richter and K.J. Kim, *A MS-SSIM optimal JPEG2000 encoder*, Data Compression Conference 2009, pp.401-411, 2009.

[9] Next Solar mission of ESA (Solar Orbiter): www.sci.esa.int/solarorbiter/

[10] M. Hu, C. Zhang, J. Lu and B. Zhou, *A multi-ROIs medical image compression algorithm with edge feature preserving*, 3rd International Conf. on Intelligent System and Knowledge Engineering, 2008.

[11] R.O. Duda and P.E. Hart, *Use of the Hough Transformation to detect lines and curves in pictures*, ACM Com., 15:11-15, 1972.