JPEG2000 Quality Scalability Without Quality Layers

Francesc Aulí-Llinàs, *Member, IEEE,* and Joan Serra-Sagristà, *Member, IEEE*

***Abstract*—Quality scalability is a fundamental feature of JPEG2000, achieved through the use of quality layers that are optimally formed in the encoder by rate-distortion optimization techniques. Two points, related with the practical use of quality layers, may need to be addressed when dealing with JPEG2000 code-streams: 1) the lack of quality scalability of code-streams containing a single or few quality layers, and 2) the rate-distortion optimality of windows of interest transmission. Addressing these two points, this paper proposes a mechanism that, without using quality layers, provides competitive quality scalability to code-streams. Its main key-feature is a novel characterization of the code-blocks rate-distortion contribution that does not use distortion measures based on the original image, or related with the encoding process. Evaluations against the common use of quality layers, and against a theoretical optimal coding performance when decoding windows of interest or when decoding the complete image area, suggest that the proposed method achieves close to optimal results.**

***Index Terms*—JPEG2000 standard, rate-distortion optimization, quality scalability, interactive image transmission.**

I. Introduction

JPEG2000 is a powerful standard structured in 12 parts addressing different aspects of coding, transmission, security, and manipulation of images and video. The Part 1 [1] of the standard, published in December 2000, defines a basic file format and the core coding system, which is the basis of the standard and is used in all other parts. Among other features, the JPEG2000 core coding system provides scalability by quality, spatial location, resolution, and component.

These four types of scalability are reached through a wavelet-based coding system built on an Embedded Block Coding with Optimized Truncation (EBCOT) [2] that partitions the image in small blocks of coefficients, called code-blocks. Scalability by component is achieved identifying image components separately, and scalability by resolution is supplied by the dyadic decomposition of the wavelet transform. The independent coding of the code-blocks, combined with a smart code-stream organization, provides scalability by spatial location. Scalability by quality is provided by the optimized truncation of the code-streams produced for each code-block, and by the definition of quality layers. Quality layers are formed by collections of code-stream segments optimally selected using rate-distortion optimization techniques.

In a multi-purpose standard like JPEG2000, quality scalability is fundamental. On the one hand, the identification of layers of quality within the code-stream enables the definition of progression orders primarily by quality, which provides optimal rate-distortion representations of the image when the code-stream is decoded at quality layer boundaries. This is essential, for example, to truncate or transmit the code-stream at different bit-rates without penalizing the quality of the decoded image. On the other hand, when quality scalability is combined with random code-stream access and the other types of scalability, the image can be interactively transmitted using the JPEG2000 Interactive Protocol (JPIP) defined in Part 9 [3], which effectively minimizes the amount of transmitted information. In this case, quality scalability is essential for the delivery of Windows Of Interest (WOI) at increasing qualities.

Another important benefit of quality scalability achieved through quality layers is that the target bit-rate, or target quality, need not to be known at encoding time and, in addition, the image does not have to be compressed multiple times to attain one or several target bit-rates or target qualities. Furthermore, the identification of quality layers within the code-stream does not require the decoding of the image, and, thus, manipulations of the image such as quality reduction or the delivery of WOIs at increasing qualities are carried out in the compressed domain, requiring low computational effort.

Beyond any doubt, the definition of quality layers is a sound mechanism of JPEG2000 to provide quality scalability. However, its practical use must take into account that, once the final code-stream is constructed, the rate-distortion contributions of code-stream segments within one quality layer can not be distinguished unless the code-stream is decoded several times and techniques such as [4] are applied. This means that, in terms of rate-distortion, it is neither possible to differentiate code-stream segments among them, nor to further truncate segments optimally, and thus the number and rate distribution of quality layers is fixed without possibility of modifications.

This fact arises two points regarding the rate-distortion optimality of JPEG2000 code-streams. Let a code-stream primarily progressive by quality contain N quality layers allocated at bit-rates $R^0, R^1, ..., R^{N-1}$, with $R^l < R^{l+1}$. When the code-stream needs to be truncated at bit-rate $\hat{R}$, with $R^l < \hat{R} < R^{l+1}$, the decoded image may have a non optimal, or even a poor, coding performance. Although this first point can be avoided using an adequate allocation strategy of quality layers [5], code-streams containing few or a single

The authors are with the Department of Information and Communications Engineering, Universitat Autònoma de Barcelona, ETSE-UAB, 08290 Cerdanyola del Vallès, Spain (e-mail: fauli@deic.uab.es; joan.serra@uab.es).

This work has been partially supported by the Spanish Government (MEC), by FEDER, and by the Catalan Government, under Grants MEC_BPD-2007-1040, TSI2006-14005-C02-01 and SGR2005-00319. A preliminary version of this research was presented in the 14th IEEE International Conference on Image Processing (September 2007 - San Antonio, Texas).

quality layer, or code-streams constructed through inadequate allocation strategies, may penalize the quality of the decoded image in more than 10 dB (see Figures 9b, 9d).

The second point that needs to be addressed is concerned with the optimality achieved when decoding WOIs from a code-stream. Let $O^0, O^1, ..., O^{N-1}$ denote the increasing bit-rate of the code-stream segments belonging to a WOI that are allocated in each quality layer, and let $\triangle O^l$ denote the bit-rate increment of the WOI's code-stream segments in layer l, computed as $\triangle O^l = O^l - O^{l-1}$. When the encoder selects and truncates the code-stream segments to form quality layers, the complete spatial area of the image is considered, and therefore the overall bit-rate of quality layers, referred to as $\triangle R^l$, can be adequately determined. However, it is easy to see that this may be not fulfilled for $\triangle O^l$, which has a bit-rate in the range $\triangle O^l \in [0, \triangle R^l]$, and, therefore, may have lack of precision.

The aim of this research is to propose a mechanism able to provide competitive quality scalability to JPEG2000 code-streams when decoding WOIs, or the complete image area, even when the code-stream contains few or a single quality layer. This is achieved by means of a characterization that can fairly estimate rate-distortion contributions of code-blocks without needing to decode the image. Based on this characterization, we propose a method able to further truncate and optimally select code-stream segments, achieving close to optimal results.

This paper is structured as follows. Section II briefly reviews the main coding stages of the JPEG2000 core coding system and describes the code-stream organization, giving the grounds for the succeeding sections. State-of-the-art rate-distortion optimization methods and allocation strategies are described in Section III. Section IV introduces the main insights of the proposed mechanism, and Section V describes the characterization of the rate-distortion slope, suggesting one algorithmic approach. Extensive experimental results assessing the performance of the proposed method are presented in Section VI. Last section points out some remarks.

II. Overview of JPEG2000

The core coding system of JPEG2000 is constituted by four main stages: sample data transformations, sample data coding, code-stream re-organization, and rate-distortion optimization. The first three stages are considered as the coding pipeline, whereas rate-distortion optimization may entail different techniques in different stages and, since JPEG2000 Part 1 only standardizes the decoder, in the literature there have appeared several rate-distortion optimization methods, which are reviewed in the following section.

Figure 1 depicts the JPEG2000 encoding pipeline. The first *sample data transformations* stage compacts the energy of the image and sets the range of the sample values. Then, the image is logically partitioned in code-blocks that are independently coded by the *sample data coding* stage, called Tier-1.

The purpose of Tier-1 is to produce a code-stream containing first the data that has the greatest distortion reductions. This is achieved through a fractional bit-plane coder and the arithmetic coder MQ, encoding each coefficient of the code-block $\mathcal{B}_i$ from the highest bit-plane $p = K_i - 1$ to the lowest

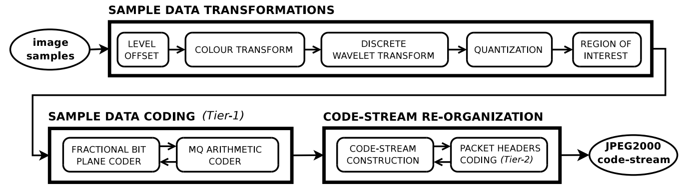


Fig. 1: Stages and operations of the JPEG2000 encoding pipeline.

bit-plane $p = 0$, K_i denoting the minimum magnitude of bit-planes needed to represent all coefficients of $\mathcal{B}_i$. In each bit-plane, Tier-1 carries out three sub-bit-plane coding passes called Significance Propagation Pass (SPP), Magnitude Refinement Pass (MRP), and Cleanup Pass (CP); each coefficient is scanned in only one of these sub-bit-plane coding passes. SPP and CP coding passes encode whether insignificant coefficients become significant in the current bit-plane or not. The main difference between SPP and CP is that the former scans those coefficients that are more likely to become significant. MRP coding pass refines the magnitude of those coefficients that have become significant in previous bit-planes. A valuable advantage of this sub-bit-plane coding is that it produces an embedded code-stream with a large collection of potential truncation points (one at the end of each coding pass) that can be used by the rate-distortion optimization techniques.

The last stage of the coding pipeline is the *code-stream re-organization*, which codes the auxiliar data needed to properly identify the content of quality layers through the Tier-2 and organizes the final code-stream in containers that encapsulate and sort the code-streams segments using one or several progression orders. Within the code-stream, the auxiliar data encoded by the Tier-2 stage is encapsulated in the packet headers that are included in quality layers. For each code-block, packet headers contain: 1) whether or not the code-block contributes to the quality layer, 2) number of included coding passes, 3) length of encoded data and, 4) number of the magnitude bit-planes K_i, which is efficiently encoded through a tag-tree technique exploiting redundancy among code-blocks within precincts (described below). Note that, usually, the length of individual coding passes is not encapsulated in packet headers.

As Figure 2 depicts, the containers within the code-stream are closely related with the partitioning system defined by JPEG2000. The first partition of the image is the tile, which defines rectangular regions of same size that are processed as completely independent images, encapsulating their data in the *tile-stream* container. Each component in a tile is called tile-component, and it is represented with $L + 1$ distinct resolutions after the sample data transformation stage, L denoting the number of levels of Discrete Wavelet Transform (DWT) applied in that tile-component. Through the dyadic decomposition allowed in JPEG2000 Part 1, each DWT level is achieved by the successive application of the wavelet filter bank to the low-frequencies subband, producing the subbands LL, HL, LH, HH that contain the Low and High frequencies in the horizontal and vertical direction respectively.

Before logically partitioning in code-blocks, JPEG2000

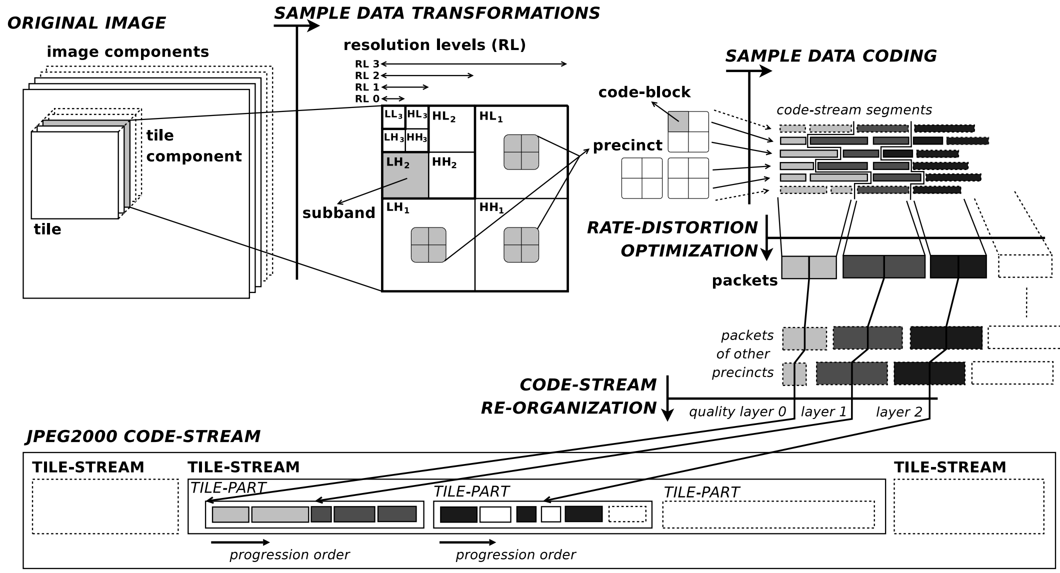


Fig. 2: JPEG2000 partitioning system and code-stream organization.

introduces the concept of precincts. In the DWT domain, one precinct is defined as the same spatial region of subbands HL, LH, HH of one resolution level. Although precincts are further partitioned in code-blocks, they define the smallest accessible spatial locations of the image because their data is enclosed in the ultimate container of the code-stream, the *packet*.

One packet encapsulates some code-stream segments of code-blocks belonging to one precinct. In order to support quality scalability, each precinct P can have several packets, referred to as $\mathcal{T}_P^l$, with $0 \leq l < N$, N denoting the number of quality layers. This organization conceptually defines the packet as a quality increment of one spatial location of one resolution level, and allows the identification of quality layers within the code-stream as the collection of packets $\mathcal{T}_P^l$, from different precincts, with equal l. The construction of a code-stream primarily progressive by quality is achieved ordering packets in increasing order $l = [0, 1, ..., N-1]$, although the standard defines five different progression orders that can be changed in each *tile-part*.

Wide possibilities of random access and re-organization are allowed in the compressed domain through this code-stream organization; for example, the modification of the image resolution, rotation, flipping or cropping of the image, and even the change of progression orders or precinct sizes without needing to decode the image. Note that all operations are carried out without modifying the content of packets, which is determined at encoding time through rate-distortion optimization and allocation strategies.

III. Rate-distortion optimization and allocation strategies

A. Rate-distortion optimization in the encoder

In JPEG2000, rate-distortion optimization is needed to select the code-stream segments –included in packets–, that minimize the overall distortion of the image attaining a target bit-rate, say R^{max}. This is used to yield a bit-rate for the final code-stream, or to allocate quality layers at different bit-rates using, for example, rate distribution functions. The first approach addressing this issue is defined in EBCOT as the Post Compression Rate-Distortion optimization (PCRD) method. The main idea behind PCRD is to use the bit-rate and the distortion of every truncation point of code-blocks to approach the optimization problem through a generalized Lagrange multiplier for a discrete set of points [6].

Let n_j denote the potential truncation points of the code-stream produced for code-block $\mathcal{B}_i$, with $0 \leq j < T_i$, T_i denoting the number of coding passes of $\mathcal{B}_i$, and let $R_i^{n_j}$ and $D_i^{n_j}$ denote, respectively, the bit-rate and distortion of these truncation points, with $R_i^{n_j-1} \leq R_i^{n_j}$. PCRD computes first the rate-distortion slope $S_i^{n_j} = \triangle D_i^{n_j} / \triangle R_i^{n_j}$, with $\triangle D_i^{n_j} = D_i^{n_j-1} - D_i^{n_j}$ and $\triangle R_i^{n_j} = R_i^{n_j} - R_i^{n_j-1}$, to identify those truncation points with strictly decreasing rate-distortion slope, i.e. those truncation points lying on the convex hull. Computing the total distortion of the image and the total bit-rate of the code-stream as $D = \sum_i D_i^{n_j}$ and $R = \sum_i R_i^{n_j}$ respectively, and considering only the truncation points lying on the convex hull, PCRD approaches the rate-distortion optimization problem as follows: $\{n_j^\lambda\}$ stands for the set of truncation points minimizing

$$(D(\lambda) + \lambda R(\lambda)) = \sum_i (D_i^{n_j^\lambda} + \lambda R_i^{n_j^\lambda}) \ ,$$

where the value of λ that minimizes this expression yielding $R(\lambda) = R^{max}$ represents the optimal solution.

Although this method achieves optimal results in terms of rate-distortion, in its original formulation it compels to fully encode the image even when few coding passes are included in the final code-stream. When encoding images at low bit-rates, this causes that Tier-1 consumes more computational resources than those strictly necessary. With the aim to reduce the computational load of Tier-1 when applying PCRD, several rate-distortion optimizations methods have been proposed in the literature in the last five years.

One common approach is to carry out the sample data coding and rate-distortion optimization simultaneously [7]–[10], encoding only those coding passes included in the final code-stream and achieving near-optimal results since the rate-distortion slope of coding passes can still be calculated. The drawback of such methods is that the wavelet data needs to be maintained in memory to be able to stop and restart the encoding of code-blocks. This is overcome in [11]–[14], collecting statistics from the already encoded code-blocks to decide which coding passes need to be encoded in the remaining code-blocks. These methods also achieve near-optimal results in terms of rate-distortion, but the computational load reduction is not as large as in the previous ones.

Another approach is to estimate the rate-distortion contributions of code-blocks before the encoding process [15]–[17], although the non optimal accuracy of estimations may penalize the coding performance. In [18], the computational load of Tier-1 is reduced by attaining the target bit-rate through the determination of different step sizes for each subband and, in the same vein, more recently a general rate control approach for wavelet data has been introduced in [19]. Other approaches based on variations of the Lagrange multiplier have been proposed in [20]–[22].

The Tier-1 computational load reduction has also been considered for hardware-based applications [23], and the complementary problem of the optimization of the bit-rate for a target quality is addressed in [24]–[26] reducing the computational load of Tier-1 too. On the other hand, the application of specific techniques of rate-distortion optimization is also needed in scenarios such as scan-based applications [27], the coding of hyperspectral data [28], implementations of motion JPEG2000 [29]–[31], and for images containing tiles [32]–[34].

Some of these methods achieve highly competitive results; for instance, at a target bit-rate of 0.0625 bits per sample (bps), they can save more than 94% of the time spent by the Tier-1, while the coding performance is reduced less than 0.1 dB compared to the optimal PCRD method [9], [10], [17].

An extensive review and comparison of rate-distortion optimization for JPEG2000 is found in [35]. However, in spite of the competitiveness of these approaches, it is important to stress that, from the point of view of providing quality scalability to already encoded code-streams, most methods can not be employed due to the use of distortion measures based on the original image, or of information related with the encoding process, since this information is not kept in the code-stream and therefore is not available once the image has been already encoded. Only the PSRA method described in [9], and the fast approximation approach in [19], could be adapted to address this issue, and are discussed in the following section.

B. *Allocation strategies of quality layers*

The use of an allocation strategy of quality layers is necessary in JPEG2000 to construct adequate code-streams in terms of rate-distortion. It is clear that if the bit-rates at which the code-stream is going to be decoded are known at encoding time, the code-stream can be optimally constructed. However, this is not usually the case and allocation strategies must construct code-streams that work reasonably well for most applications and scenarios.

Some recommendations on the number and bit-rate of quality layers are given in [36, Chapter 8.4.1] based on experience. More recently, the rate-distortion optimality of JPEG2000 code-streams has been evaluated under an Expected Multi-Rate Distortion measure that weights the distortion of the image recovered at some bit-rates by the probability to recover the image at those bit-rates [5]. Under this measure and considering uniform, logarithmic, and Laplacian distributions, a smart algorithm able to optimally construct code-streams is proposed. Although this research is the first one proposing an optimal construction of JPEG2000 code-streams, experimental results suggest that the improvement achieved by the proposed method is usually small. This is explained by the authors due to the optimality of PCRD and the almost convex rate-distortion curve that Tier-1 already originates.

C. *Transmissions of WOIs in the framework of JPIP*

Allocation strategies consider the complete image area, meaning that code-streams may be optimally constructed for the whole image; nevertheless, the decoding and transmission of WOIs still needs to be addressed. In JPEG2000, this issue is first considered in the framework of JPIP to improve the quality of the transmitted images [37]. In this approach a re-sequencing of the packets belonging to the WOI is carried out in order to place, in the first quality layers, those packets that contribute the most to the distortion decrease. The key-feature of the method is the use of a window scaling factor –related to the number of coefficients in a code-block that belong to the WOI and to the energy gain factor of subbands–, to determine the rate-distortion contribution of each packet.

Similar techniques are used in [38] in the context of telemedicine when transmitting volumetric images, and in [39] to enhance the browsing experience through user navigation models, however, none of these techniques considers the further truncation of the code-stream segments contained in each packet. Only recently, this is considered in a method that recovers distortion estimates [40] of the code-stream segments of a code-block. This method extrapolates rate-distortion slopes at continuous bit-rate through the rate-distortion slope thresholds associated to each quality layer and through an auxiliary transition slope determined for each subband, which have to be explicitly transmitted.

IV. Proposed mechanism

A. *Main insight*

The main idea behind the proposed method is to estimate the rate-distortion slope of coding passes in order to conceptually apply the PCRD over these estimations. This would allow the further truncation of the code-stream segments within packets and, therefore, the optimal decoding of WOIs or of the complete image area. However, in order to apply PCRD over rate-distortion slope estimations, coding passes need to be identified within the code-stream, which is an operation that requires the decoding of the image, because packets contain one or more coding passes, probably without keeping the length of each one.

To overcome this difficulty maintaining JPEG2000 compliance, we use the restart coding variation of Tier-1 [36, Chapter 12.4]. The intention of this coding variation is to restart the MQ coder at the beginning of each coding pass to facilitate the parallel implementation of coding passes. To our purposes, the main consequence of the restart coding variation is that, since it produces one codeword for each coding pass, it compels to explicitly code the length of coding passes in the header of packets. This allows the recovery of coding passes lengths just decoding packet headers, which is an operation with a low computational complexity (it spends less than 2% of the decoding time) and, furthermore, it almost does not penalize the bit-rate of the final code-stream (less than 1%). To evaluate the proposed approach when no ancillary information is transmitted, we also report the results obtained when the MQ restart coding variation is not used, and the coding passes lengths are *not* kept in the JPEG2000 code-stream; this approach implies modifications slightly beyond the scope of the standard, since the encoder must store coding passes lengths to an external file independent of the code-stream. However, note that both approaches construct fully compliant JPEG2000 code-streams and perfectly fit in the framework of interactive image transmissions. Note also that the creation of the external file of lengths, or the introduction of the restart coding variation to already encoded code-streams, is needed only once, without requiring the full decoding of the image.

The straight approach to estimate rate-distortion slopes of code-blocks is to consider that coding passes situated at high bit-planes have larger rate-distortion contributions than coding passes situated at low bit-planes. Taking into account that 97% of the truncation points lying on the convex hull only contain one or two coding passes [41], one might think that the successive encoding of coding passes from the highest bit-plane of the image to the lowest one may obtain competitive results.

This is the main idea used in PSRA [9] except for the last step that uses rate-distortion slopes and, in the context of providing quality scalability to already encoded code-streams, this idea has also been used in the Coding Passes Interleaving (CPI) method [42]. CPI defines a coding level c as the coding pass of all code-blocks of the image at the same height, given by $c = (p \cdot 3) + cp$, where p stands for the bit-plane and cp stands for the coding pass type with $cp = \{2 \text{ for SPP},\ 1 \text{ for MRP},\ 0 \text{ for CP}\}$. Coding passes are scanned from the highest coding level of the image to the lowest one until the target bit-rate is achieved. In each coding level, coding passes are selected from the lowest resolution level to the highest one, and in each resolution level, subbands are scanned in order $[HL, LH, HH]$.

Similar scanning orders are used in other image coding systems achieving competitive coding performance [43], therefore it may be expected that CPI should also obtain a competitive coding performance, similar to the optimal PCRD method. However, experimental evidence suggests that CPI achieves a coding performance that, at some bit-rates, is more than 0.5 dB worse than PCRD (see Figure 3). The Reverse subband scanning Order and coding passes Concatenation (ROC) method [44] introduces three modifications to CPI in order to improve coding results but, for some images, this is not achieved [35].

B. Evaluation of our hypothesis

CPI and ROC define a fixed scanning order considering resolution levels and subbands, but within a subband, all code-blocks are considered equally. The underlying idea of the method proposed here is to precisely distinguish code-blocks within subbands. Our hypothesis is that, just identifying the number of magnitude bit-planes of code-blocks, which can be obtained decoding packet headers, rate-distortion slopes of coding passes can be fairly estimated. Similarly, the fast approximation approach in [19], which also requires only the maximum coefficient magnitude in a group of coefficients, efficiently computes an estimation of the slope that is used to model the rate-control.

Within a subband, code-blocks with different number of magnitude bit-planes will have different estimations, whereas the code-blocks with the same number of magnitude bit-planes will have the same estimation, so they are grouped in a code-block set $\mathcal{G}$ that is identified by $b_{r,s}$ and K, where K stands for the number of magnitude bit-planes, and $b_{r,s}$ stands for the resolution level r ($r = 0$ for the lowest one) and the subband type s ($s = \{0 \text{ for } HL/LH \text{ subbands, } 1 \text{ for } HH \text{ subband}\}$). We do not make distinctions among code-blocks of subbands HL and LH, since they have the same energy gain factor and thus the signal difference among them just corresponds to the vertical and horizontal details of the image. Besides, experimental evidence suggest that differentiating HL and LH does not improve coding results. This type of subband arrangement is also used in [15].

To validate our hypothesis and for comparison purposes, we first evaluate the coding performance that can be optimally achieved when considering the actual rate-distortion slope of code-block sets in the JPEG2000 encoder. Let D_i^c denote the distortion of code-block $\mathcal{B}_i$ at coding level c, and $D_{\mathcal{G}}^c$ denote the sum of distortions of code-block set $\mathcal{G}$, given by

$$D_{\mathcal{G}}^c = \sum_{\mathcal{B}_i \in \mathcal{G}} D_i^c \ .$$

Let R_i^c be the length of the coding pass corresponding to coding level c of $\mathcal{B}_i$, and let $R_{\mathcal{G}}^c$ be the sum of these lengths for the code-block set $\mathcal{G}$. Using $D_{\mathcal{G}}^c$ and $R_{\mathcal{G}}^c$, we calculate the rate-distortion slope of code-block sets as

$$S_{\mathcal{G}}^c = \frac{\triangle D_{\mathcal{G}}^c}{\triangle R_{\mathcal{G}}^c} = \frac{D_{\mathcal{G}}^{c+1} - D_{\mathcal{G}}^c}{R_{\mathcal{G}}^c - R_{\mathcal{G}}^{c+1}}$$

to identify the operational rate-distortion curve of $\mathcal{G}$. Then, the search of the optimal truncation points for the code-blocks sets attaining a specified bit-rate can be straightforwardly carried out using the Lagrange multiplier. We refer this method to as PCRD over Code-block Sets (PCRD-CS).

Figure 3 depicts the coding performance achieved with PCRD-CS when compared to PCRD for the Fruit Basket

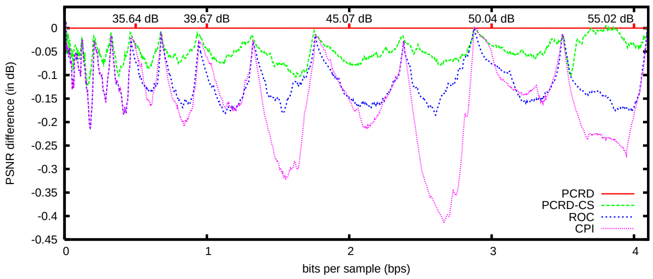


Fig. 3: Evaluation of the coding performance achieved with PCRD-CS, ROC and CPI compared to PCRD. Results for the Fruit Basket image (gray scaled, 2560×2048). Coding parameters: 9/7 DWT 5 levels, derived quantization, 64×64 code-blocks, no precincts, restart coding variation.

image of the ISO/IEC 12640-1 corpus [45]. In this evaluation, the image has been encoded using both methods at 2000 target bit-rates equivalently spaced between 0.001 to 4.1 bps, decoding the code-streams and computing the Peak Signal to Noise Ratio (PSNR) with the original image. The results are given as the PSNR difference between PCRD and PCRD-CS in order to enhance the visual comparison. The straight line of this graphic depicts the coding performance achieved by PCRD; the remaining plots depict the difference (in terms of dB) achieved at the same bit-rate between PCRD and the evaluated method. Besides the comparison between PCRD and PCRD-CS, this figure also reports the results achieved with CPI and ROC to evaluate the gains that can be expected from the grouping of code-blocks in sets. PCRD-CS obtains an almost regular coding performance along all bit-rates, with a penalization of less than 0.05 dB on average compared to PCRD, whereas CPI and ROC are, at some bit-rates, more than 0.4 and 0.2 dB worse than PCRD respectively. These results also hold for the other images of the corpus. It is worth noting that PCRD-CS achieves competitive results considering few number of code-block sets; for the images of the ISO/IEC 12640-1 corpus, for instance, the number of code-block sets in the two highest resolution levels ranges from 4 to 7, and in the two lowest resolution levels, it ranges from 1 to 3.

V. Characterization of the Rate-Distortion Slope

PCRD-CS is only of interest from a theoretic point of view, since it can only be applied at encoding time. However, it suggests that the grouping of code-blocks within subbands may obtain near-optimal results in terms of coding performance. The challenge here is to fairly characterize the rate-distortion slope of code-blocks to carry out estimations that can obtain similar results. Our characterization of the rate-distortion slope is based on two important characteristics of the encoding process of a code-block.

A. Concatenation of coding passes

The first characteristic is that, at the same bit-plane, coding passes of type MRP often have smaller rate-distortion slopes than coding passes of type CP. Therefore, a coding pass of type MRP of a code-block should be concatenated with the following CP.

We explain this characteristic using the rate-distortion model proposed in [17], which estimates the decrement in distortion and the increment in bit-rate of a code-block at bit-plane p according to

$$\triangle D = (N_{\text{sig}} + 0.25 N_{\text{ref}})(2^p)^2 \quad \triangle R = 2N_{\text{sig}} + N_{\text{ref}} + N_{\text{insig}} \ ,$$

where $N_{\text{sig}}, N_{\text{ref}}, N_{\text{insig}}$ denote, respectively, the number of significant, refinement and insignificant coefficients at bit-plane p. This model approaches the distortion through the Mean Squared Error (MSE), basing the factor 0.25 of $\triangle D$ on the expected distortion reduction of the refinement coefficients, derived in [46]. The estimation of the bit-rate assumes that both a refinement and an insignificant coefficient are coded with 1 bit, and that a significant coefficient is coded with 2 bits to include the sign. It is worth noting that, although this model fairly approximates distortions, bit-rates are roughly approximated since the arithmetic coder MQ encodes the incoming data efficiently, thus producing a shorter code-stream.

The authors proposing this model consider insignificant those coefficients that are *not* coded through the run mode defined in the coding pass CP. The run mode encodes four insignificant coefficients with a single bit and, mostly at the highest bit-planes of code-blocks, most coefficients are encoded through the run mode, so in order to distinguish between coding passes of MRP and CP we slightly modify this model as follows:

$$\begin{array}{ll} \triangle D^{MRP} = (0.25 N_{\text{ref}})(2^p)^2 & \triangle R^{MRP} = N_{\text{ref}} \\ \triangle D^{CP} = N_{\text{sig}}(2^p)^2 & \triangle R^{CP} = 2N_{\text{sig}} + 0.25 N_{\text{insig}} \ . \end{array}$$

with N_{insig} denoting all the insignificant coefficients, also including the ones encoded through the run mode.

Through these estimations, we calculate when the rate-distortion slope of coding passes of type CP, referred to as S^{CP}, is greater than the rate-distortion slope of coding passes of type MRP, referred to as S^{MRP}, by

$$S^{CP} > S^{MRP} \equiv \frac{\triangle D^{CP}}{\triangle R^{CP}} > \frac{\triangle D^{MRP}}{\triangle R^{MRP}} \rightarrow N_{\text{sig}} > 0.125 N_{\text{insig}}$$

inferring that $S^{CP} > S^{MRP}$ when at least 12.5% of the coefficients encoded in a coding pass of type CP are significant. Table I shows, at different bit-planes, the percentage of coefficients that have become significant in coding passes of type CP, emphasizing in bold font those ones with a percentage greater than 12.5%. Except for the three or four highest bit-planes, this is always fulfilled in all code-block sets. Experimental evidence in [36, Chapter 8.3.3] also suggests that coding passes of type MRP lie on the convex hull more often at high bit-planes than at medium and low bit-planes.

B. The balloon effect

The second important characteristic of the encoding process of a code-block is named the balloon effect. We explain the balloon effect through the following assumption: the coding passes that encode the largest number of significant coefficients have the greatest rate-distortion slope values. This

TABLE I: For Coding Passes of Type CP, Percentage of Coefficients Becoming Significant. Results for the Musicians Image (Gray Scaled, 2560×2048, 9/7 DWT 5 Levels). These Code-block Sets Account For 98% of the Image Code-blocks.

p	$b_{3,0}$ K=9	$b_{3,1}$ K=8	$b_{4,0}$ K=7	$b_{4,1}$ K=7	$b_{5,0}$ K=6	$b_{5,1}$ K=6
8	1.00%					
7	1.33%	0.36%				
6	1.15%	1.15%	0.11%	0.43%		
5	2.39%	1.84%	1.80%	2.00%	0.59%	0.29%
4	**12.51%**	**15.83%**	**14.30%**	**13.91%**	7.51%	3.33%
3	**17.84%**	**27.84%**	**22.51%**	**23.47%**	**16.75%**	12.47%
2	**22.86%**	**30.00%**	**26.24%**	**17.15%**	**17.13%**	**15.55%**
1	**41.43%**	**25.87%**	**75.00%**	**30.56%**	**20.77%**	**22.97%**
0	-	-	-	**50.00%**	**32.95%**	**39.81%**

TABLE II: Number of Significant Coefficients Encoded in Each Bit-plane of Code-block Sets Within Subband $b_{3,0}$. Results for the Candle Image (Gray Scaled, 2560×2048, 9/7 DWT 5 Levels).

	CP			SPP		
p	K=11	K=10	K=9	K=11	K=10	K=9
10	1					
9	23	15		5		
8	238	110	27	68	25	
7	118	159	129	392	195	50
6	53	112	163	418	451	266
5	71	77	125	505	583	507
4	94	55	82	609	621	646
3	16	27	33	616	550	669
2	1	10	11	403	421	514
1	0	9	8	198	282	365
0	0	1	1	139	184	240

assumption relies on the meaningful difference between the large decrement in distortion, compared to the small increment in bit-rate. For coding passes of type SPP and CP, the rate-distortion model proposed in [17] estimates this according to

$$\triangle D = N_{\text{sig}}(2^p)^2 \qquad \triangle R = 2N_{\text{sig}} + N_{\text{insig}} \ ;$$

the encoding of a significant coefficient increases twice the estimated bit-rate of the encoding of an insignificant coefficient, however $\triangle D$ is decremented by $(2^p)^2$. Therefore, it is expected that, as more significant coefficients are encoded in a coding pass, greater its rate-distortion slope, especially at high bit-planes [16]. Table II shows the number of significant coefficients encoded at each bit-plane of the code-block sets belonging to the subband $b_{3,0}$, distinguishing coding passes SPP and CP.

Two issues are worth noting in this table. The first one is that, for coding passes of type CP, the number of significant coefficients encoded at each bit-plane increases from the first to the third highest bit-plane of each code-block set, and then decreases progressively. The same property holds for coding passes of type SPP, but the increase is from the first to the fifth (or seventh) highest bit-plane. Figure 4 represents this as a balloon, where the width represents the number of significant coefficients encoded at the bit-plane.

The second remarkable issue of Table II is that the number of significant coefficients at the highest bit-plane depends on the magnitude bit-planes of the code-block set. This is, at their

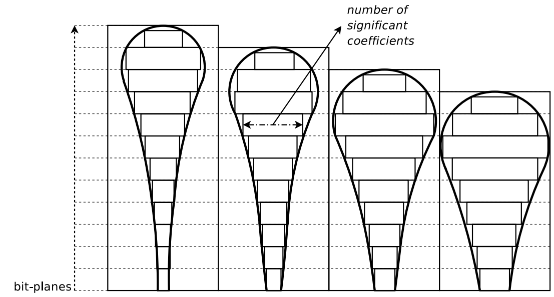


Fig. 4: The number of significant coefficients encoded in each bit-plane of the code-blocks within one subband can be represented as a balloon.

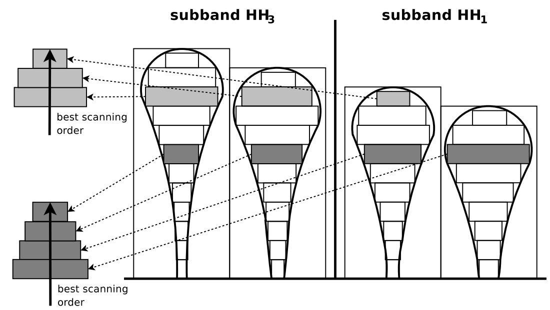


Fig. 5: The balloon effect among different subbands.

highest bit-plane, the code-blocks that have the lowest number of magnitude bit-planes encode more significant coefficients, and this relation is always respected in Table II. Figure 4 depicts this as the shape of the balloon: the shorter the balloon is, the wider.

We explain why the balloon effect occurs as follows. First, we know that the largest number of significant coefficients are usually found at high bit-planes, especially for coding passes of type CP. This could be caused for the application of the high-pass filter of the DWT, which sets most coefficients of a subband to a null value apart from the areas where high frequencies are detected. These high frequencies areas usually have large values concentrated in the same spatial locations. Second, the scan performed in coding passes of type CP visits almost all the coefficients of the code-block at the highest bit-planes, therefore, it is expected that it discloses these areas of high frequencies. The more significant coefficients the CP encodes, the less coefficients at the following CP bit-plane it visits, and consequently the less significant coefficients it encodes. At the lowest bit-planes, almost all the coefficients still insignificant are neighbours of already significant coefficients, so while CP does not visit almost any coefficient, the SPP encodes all the coefficients of the medium and lowest bit-planes. This also explains why the balloon effect is more emphasized in coding passes of type CP than in coding passes of type SPP.

The balloon effect is observed in all subbands of an image except for the LL, but experimental evidence suggests that the consideration of subband LL separately does not improve results significantly (this subband contains few code-blocks). Considering code-blocks among subbands, a consequence of the balloon effect for strategies using scanning orders is that, although at the highest bit-planes the best scanning order is achieved starting from the lowest resolution level to the highest one, from a particular bit-plane onwards, the best scanning is achieved reversing this order (see Figure 5). This observation is validated experimentally in [44]. However, note that the lack of specification of a scanning order, such as the proposed method CoRD does, gives a finer approach that may improve results and, furthermore, it avoids the need to determine the point where the scanning order has to be reversed [35].

C. CoRD algorithm

Based on the characterization above, we propose an algorithmic approach to estimate rate-distortion slopes of the coding passes of code-block sets. The rate-distortion slope is computed as

$$S^c = \begin{cases} c + \mathcal{F}_{SPP} & \text{for SPP coding passes} \\ c + \mathcal{F}_{MRP} & \text{for MRP coding passes} \\ c + 1 + \mathcal{F}_{CP} & \text{for CP coding passes} \end{cases},$$

where c identifies the coding level. In this equation, c compels to select coding passes from the highest to the lowest coding level of the image and, within each coding level, coding passes are selected upon the value of $\mathcal{F}$. In order to assure that coding passes of type MRP are always concatenated with the consecutive coding pass of type CP, we set $\mathcal{F}_{MRP} = 0$, except for the MRP of the second highest bit-plane, where $\mathcal{F}_{MRP} = 0.99$. For coding passes of type SPP and CP, $\mathcal{F}_{\{SPP|CP\}}$ attempts to approximate the balloon effect within each subband by

$$\mathcal{F}_{\{SPP|CP\}} = \begin{cases} \mathcal{F}_{init}(\mathcal{F}_{inc})^{K_S - p} & \text{if } p \geq K_{balloon} \\ 1 - (\mathcal{F}_{dec}(K_{balloon} - p)) & \text{otherwise} \end{cases},$$

where $K_S = K - 1$ for CP coding passes and $K_S = K - 2$ for SPP coding passes. This expression increases $\mathcal{F}_{\{SPP|CP\}}$ exponentially from the highest bit-plane $K - 1$ to the bit-plane $K_{balloon}$, and decreases $\mathcal{F}_{\{SPP|CP\}}$ linearly from the bit-plane $K_{balloon} - 1$ to the lowest bit-plane 0. $K_{balloon}$ is set to the bit-plane that causes $\mathcal{F}_{\{SPP|CP\}} \geq 1$, i.e.

$$K_{balloon} = P \text{ such that } \nexists \;\; p > P, \;\; \mathcal{F}_{init}(\mathcal{F}_{inc})^{K_S - p} \geq 1 \; .$$

In this way, the values of $\mathcal{F}_{\{SPP|MRP|CP\}}$ are restricted to the interval $[0, 1)$. $\mathcal{F}_{init}$ must reflect the rate-distortion slope initialization at the highest bit-plane, in other words, the width of the top of the balloon. Good choices for these three parameters, which have been determined experimentally, are given in Table III, where $\#K = K_{max} - K_{min} + 1$ with K_{max} and K_{min} respectively denoting the maximum and minimum K in the subband to which the code-block belongs. With our choice of $\mathcal{F}_{dec}$, the expression for $\mathcal{F}_{\{SPP|CP\}}$ could be simplified, although we keep $\mathcal{F}_{dec}$ to retain generality. When this algorithm is used to decode WOIs, K_{max} and K_{min} denote the maximum and minimum K of the code-blocks within the subband *and* belonging to the WOI. Recall from Section II that the number of magnitude bit-planes K is coded through the Tier-2 stage for each code-block and is encapsulated in packet headers. K_{max} and K_{min} can be straightly identified decoding the headers of the subband's packets.

TABLE III: Choices of Parameters $\mathcal{F}_{init}$, $\mathcal{F}_{inc}$ and $\mathcal{F}_{dec}$.

$\mathcal{F}_{CP}$	$\mathcal{F}_{SPP}$
$\mathcal{F}_{init} = \frac{0.075}{\#K}(K_{max} - K + 1)$	$\mathcal{F}_{init} = \frac{0.05}{\#K}(K_{max} - K + 1)$
$\mathcal{F}_{inc} = 10$	$\mathcal{F}_{inc} = 4$
$\mathcal{F}_{dec} = \frac{1}{K_{balloon} + 2}$	$\mathcal{F}_{dec} = \frac{1}{K_{balloon} + 2}$

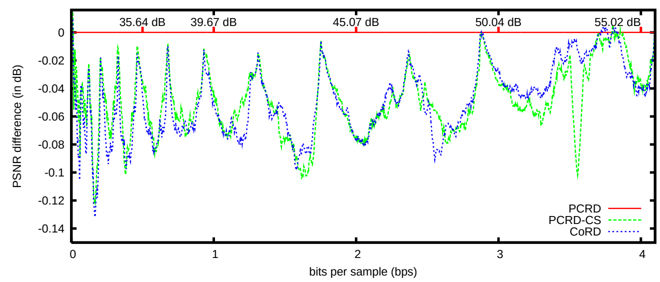


Fig. 6: Evaluation of the coding performance achieved with PCRD-CS and CoRD compared to PCRD. Results for the Fruit Basket image (gray scaled, 2560×2048). Coding parameters: 9/7 DWT 5 levels, derived quantization, 64×64 code-blocks, no precincts, restart coding variation.

Using these estimations, the selection of coding passes to yield a target bit-rate can be carried out without needing information related with the encoding process, or distortion measures based on the original image. Only packet headers need to be decoded. We name this method as the Characterization of the Rate-Distortion slope (CoRD).

To evaluate the accuracy of this algorithm, Figure 6 depicts the coding performance achieved when using CoRD to decode the Fruit Basket image at 2000 target bit-rates using a single quality layer code-stream (see next section for a larger description about how the experiments have been conducted). Results are given as the PSNR difference between PCRD and CoRD. Performance achieved with PCRD-CS is also depicted. Note that CoRD achieves almost the same coding performance as PCRD-CS in all bit-rates. For most images of the ISO/IEC 12640-1 corpus, the same results hold [35].

VI. Experimental results

A. Decoding the complete image area

To assess the coding performance of the proposed method, we present first a comparison between CoRD and the use of quality layers when the complete spatial area of the image is decoded. CoRD is evaluated decoding the image at specified target bit-rates from a code-stream containing a single quality layer and the restart coding variation. This method is referred to as CoRD+r. When the length of coding passes is kept in an external file and thus the restart coding variation is not used, the method is referred to as CoRD-r. The bit-rate of the external file keeping the length of coding passes in CoRD-r is *not* considered in experiments since the intention of CoRD-r is to evaluate the proposed approach when no ancillary information is transmitted.

The evaluation of the coding performance achieved using quality layers considers two common allocation strategies: 1) to distribute quality layers *logarithmically* spaced in terms of bit-rate and, 2) to distribute quality layers *equivalently* spaced in terms of bit-rate. To enhance the results of the second allocation strategy at low bit-rates, the layers are finely distributed from 0.001 to 0.5 bps and coarsely from 0.5 to 5 bps. For both strategies, code-streams containing 20, 40, 80 and 120 quality layers have been constructed, although graphics only report the best results for both strategies to ease the visual interpretation.

Results are given as the difference between the PSNR obtained with CoRD and the PSNR obtained with PCRD when encoding at the same bit-rate. For quality layers, the recovered image is obtained when the code-stream is truncated and decoded at the specified bit-rates. In all experiments, the construction of code-streams containing quality layers and the encoding with PCRD does not use the restart coding variation. The top straight line of graphics depicting PCRD identifies the maximum coding performance that can be obtained with JPEG2000, however note that this performance is mostly of interest from a theoretical point of view, since PCRD can only be applied at encoding time.

Kakadu [47] has been used to construct code-streams containing quality layers and to encode with the PCRD method[1]. CoRD has been implemented in our JPEG2000 Part 1 implementation BOI [48]. Parameters of both applications are set to: lossy compression, 5 levels of DWT, derived quantization, code-blocks of size 64×64, maximum size of precincts. All images of the ISO/IEC 12640-1 corpus have been evaluated using 600 target bit-rates uniformly distributed among 0.001 to 5 bps.

Figure 7 depicts the results obtained for the Cafeteria image. The best coding performance achieved with quality layers is when the code-stream contains 20 logarithmically, or 40 equivalently spaced quality layers. On average, these allocation strategies are 0.61 dB and 0.15 dB worse than PCRD, respectively. Needing no information of rate-distortion, CoRD+r is only 0.08 dB worse than the best allocation strategy of quality layers, and CoRD-r outperforms both allocation strategies in almost all bit-rates, achieving an average coding performance of only 0.04 dB worse than PCRD. Note that both CoRD+r and CoRD-r use same estimations, therefore differences in their results are caused because CoRD+r is using the restart coding variation that keeps coding passes lengths within the code-stream, whereas CoRD-r does not need to transmit ancillary information.

Figure 8 reports the coding performance average among all images of the corpus. On average for all bit-rates and all images, CoRD-r is 0.05 dB worse than PCRD and CoRD+r is 0.18 dB worse than PCRD, whereas the allocation strategy forming 40 equivalently spaced quality layers is 0.16 dB worse than PCRD.

Same experiments have been carried out using an image corpus belonging to the Remote Sensing (RS) community,

[1]To avoid penalizing the coding performance, Kakadu has been forced to use 32-bit data representations for floating point operations and the comment markers introduced by default have been deactivated.

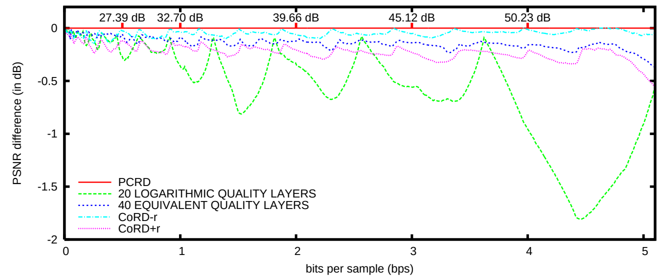


Fig. 7: Evaluation of the coding performance achieved with CoRD and the use of quality layers compared to PCRD. Results for the Cafeteria image (gray scaled, 2048×2560).

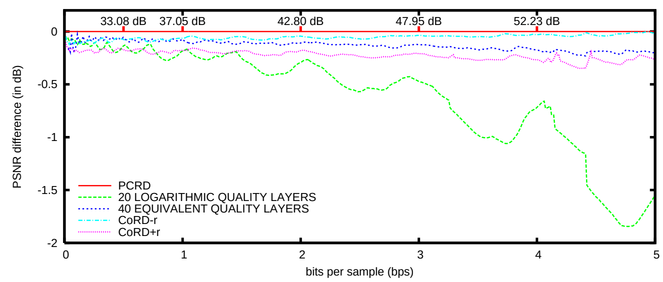
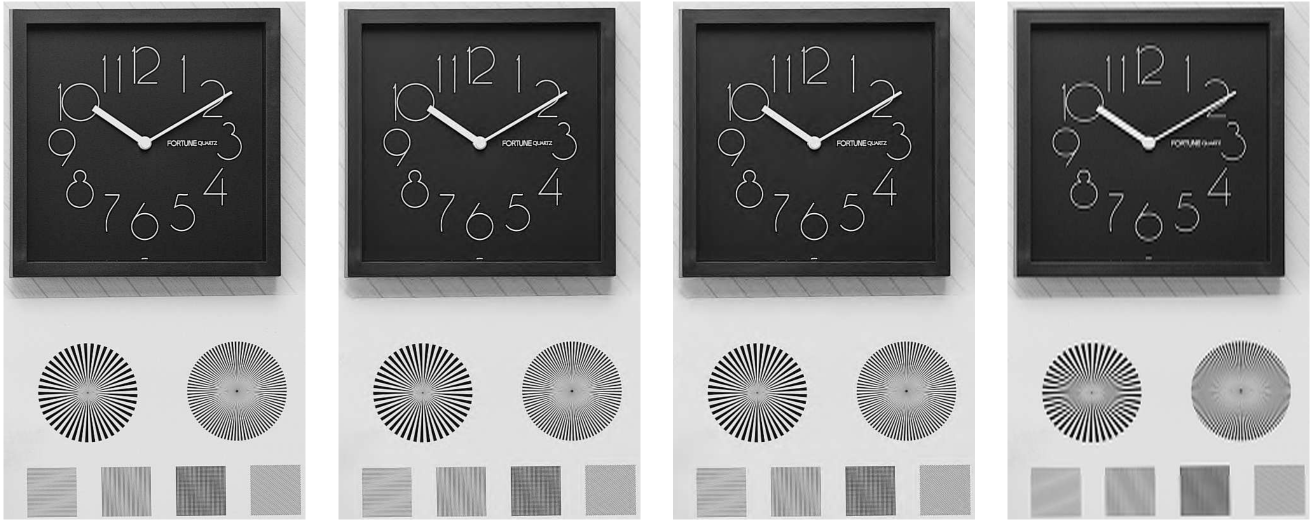


Fig. 8: Evaluation of the coding performance achieved with CoRD and the use of quality layers compared to PCRD. Average results for the images of the ISO/IEC 12640-1 corpus (gray scaled, 2560×2048).

and an image corpus belonging to the medical community. The RS corpus has four aerial images provided by the Cartographic Institute of Catalonia (ICC) [49], which cover vegetation and urban areas (gray scaled, size of 4096×4096). The medical corpus has four images provided by UDIAT Centre Diagnostic [50], including one computer radiology and three radiologies (gray scaled, sizes varying from 480×640 to 2700×3913). Table IV reports the results achieved by the best allocation strategies of quality layers and by CoRD at different bit-rate ranges. The PCRD row indicates the PSNR achieved when encoding with PCRD at the highest bit-rate of the range; the remaining rows indicate the average PSNR difference along the specified bit-rate range. Bit-rate ranges for the medical corpus are smaller than for the other corpora since the encoding of medical images achieves same PSNR results at lower bit-rates. The coding performance results are similar for all corpora: CoRD-r achieves the best performance and CoRD+r is slightly worse than an allocation strategy using equivalently spaced quality layers. These results also hold for other coding parameters [35].

Regarding the visual comparison, Figure 9 shows an area of the Bicycle image (ISO/IEC 12640-1 corpus) decoded at 0.5 bps from a code-stream containing 20 logarithmically spaced quality layers and from a single quality layer code-stream decoding with and without CoRD. The images decoded with CoRD and from the code-stream containing 20 quality layers are practically equivalent, whereas the image decoded from the single quality layer code-stream is clearly blur.



(a) ORIGINAL

(b) 20 LOGARITHMIC QUALITY LAYERS - 34.42 dB

(c) SINGLE QUALITY LAYER with CoRD+r - 34.35 dB

(d) SINGLE QUALITY LAYER - 22.70 dB

Fig. 9: Visual comparison between CoRD and the use of quality layers. Bicycle image (gray scaled, 2048×2560) decoded at 0.5 bps (compression factor 16:1). An area of 475×815 is showed.

TABLE IV: Average Coding Performance Achieved with CoRD and the Use of Quality Layers Compared to PCRD. Average Results, in Different Bit-rate Ranges, for the Images of Corpora: ISO/IEC 12640-1, RS and Medical.

		bit-rate range (in bps)			
		$(0, 0.5]$	$(0.5, 1]$	$(1, 2]$	$(2, 4]$
ISO/IEC	*PCRD*	*32.97 dB*	*36.98 dB*	*42.74 dB*	*52.53 dB*
	20 log	-0.12	-0.19	-0.29	-0.64
	40 equiv	-0.08	-0.09	-0.11	-0.14
	CoRD-r	-0.08	-0.07	-0.06	-0.04
	CoRD+r	-0.16	-0.17	-0.18	-0.22
RS	*PCRD*	*26.97 dB*	*30.11 dB*	*35.49 dB*	*46.78 dB*
	40 log	-0.08	-0.10	-0.19	-0.80
	80 equiv	-0.06	-0.06	-0.07	-0.12
	CoRD-r	-0.03	-0.03	-0.03	-0.04
	CoRD+r	-0.07	-0.08	-0.12	-0.19
		$(0, 0.25]$	$(0.25, 0.5]$	$(0.5, 1]$	$(1, 2]$
medical	*PCRD*	*41.77 dB*	*44.68 dB*	*48.42 dB*	*54.02 dB*
	20 log	-0.11	-0.14	-0.16	-0.36
	40 equiv	-0.14	-0.10	-0.10	-0.12
	CoRD-r	-0.07	-0.09	-0.08	-0.05
	CoRD+r	-0.14	-0.17	-0.17	-0.19

B. *Decoding WOIs*

The aim of this evaluation is to fairly assess and compare the coding performance achieved with CoRD and the use quality layers when decoding WOIs. Although the transmission of WOIs has been commonly evaluated in the framework of JPIP [37]–[39], we slightly change this point of view to carry out the evaluation considering only the optimality of CoRD and quality layers. In our evaluation, WOIs are selected having an *exact cropping* in the compressed domain or, in other words, corresponding to some specific code-blocks of the image without trespassing code-block boundaries. The specification of random-defined WOIs, which is necessary in the JPIP's framework, would compel to use a window scaling factor [37] for both CoRD and quality layers. Such evaluation is discarded since our purpose is to consider CoRD and quality layers mechanisms as they are, without implicating other techniques required in interactive transmissions. Furthermore, it is expected that a window scaling factor influences both mechanisms similarly.

The following procedure is carried out to assess the coding performance achieved with quality layers: 1) the image is encoded using an allocation strategy of quality layers, 2) the WOI is cropped in the compressed domain by extracting those packets belonging to the WOI spatial area, 3) the WOI packets are used to construct a new code-stream and, 4) the quality layers of the new code-stream are re-sequenced eliminating those layers not holding any packet. Through this process, the new code-stream contains the image area corresponding to the WOI with the original distribution of packets among quality layers. This code-stream is truncated and decoded at the specified bit-rates, repeating the process for both allocation strategies proposed in the previous section and with code-streams containing 20, 40, 80, and 120 quality layers, although only the best results are reported. CoRD has been evaluated when decoding the WOI from a code-stream containing a single quality layer.

Results are given as PSNR differences among 500 bit-rates uniformly distributed from 0.001 to 4 bps (bit-rate is computed considering the size of the WOI). To evaluate the coding performance achieved by PCRD, the WOI is cropped in the original image and encoded using PCRD at the same target bit-rates used for the evaluation of quality layers and CoRD. In order to avoid the areas affected by the wavelet transform boundaries of the original image, in the comparison between images, a sufficient boundary for the WOI has been considered. This is the cause that in some bit-rates PCRD is not strictly optimal.

An aerial image provided by ICC covering a region of

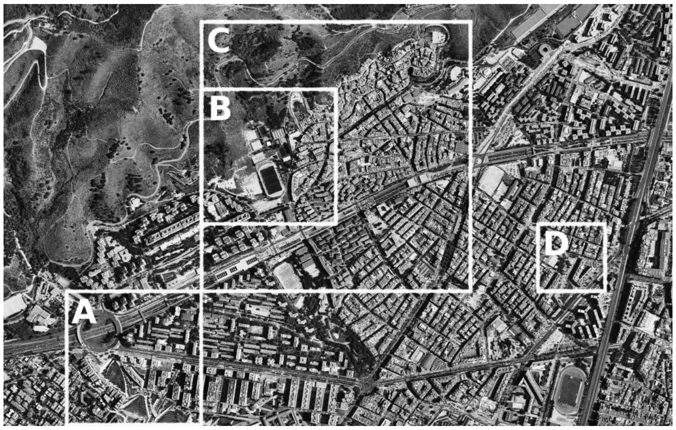


Fig. 10: Area of the image containing the evaluated WOIs. The size of the complete image is 7168×4096; size of WOIs A and B is 1024×1024, size of WOI C is 2048×2048, and size of WOI D is 512×512.

Barcelona including vegetation and urban areas has been used to report these experiments. We present the evaluation when decoding the four WOIs showed in Figure 10. Coding parameters are the following: lossy compression, 4 levels of DWT, derived quantization, code-blocks of size 32×32, and precincts of the same size as the WOI for quality layers, and with the maximum size for CoRD. As well as in the previous experiments, the restart coding variation is only used for CoRD+r.

Figure 11 depicts the obtained results. The coding performance achieved with quality layers is very similar to the one achieved in the previous evaluation when the complete image area is decoded, suggesting that quality layers are also near-optimal when decoding WOIs. Results of CoRD are similar to the ones obtained in the previous evaluation: CoRD-r outperforms both strategies of quality layers allocation, and CoRD+r achieves almost the same results as the best of such strategies. These results also hold for other WOI sizes, locations, and images.

We would like to stress the accuracy of the proposed method, since CoRD-r achieves practically the same coding performance as the one achieved with PCRD. The inclusion of the length of the coding passes for the restart coding variation penalizes the coding performance of CoRD+r in a similar degree than quality layers.

VII. Conclusions

This paper addresses the quality scalability of JPEG2000 code-streams by means of a mechanism able to distinguish the rate-distortion contributions of individual coding passes of code-blocks. The key-feature of this mechanism is the Characterization of the Rate-Distortion slope (CoRD), which provides insights to fairly estimate the operational rate-distortion curve of the JPEG2000 fractional bit-plane coder, without using distortion measures based on the original image, or measures related with the encoding process.

In addition of CoRD being employed as a mechanism to replace the use of quality layers while still providing competitive quality scalability, practical applications of the proposed method are the optimal truncation of code-streams containing few or a single quality layer, the re-construction of quality layers of a code-stream, and the optimal decoding of WOIs. Since the proposed method has negligible computational costs, requiring only the decoding of packet headers when the image is indexed, it is suitable to use in interactive image transmissions allowing optimal deliveries of WOIs at increasing qualities. With slight modifications, the proposed method may also serve for the fusion of multiple compressed views of a scene [40], or for the modification of video quality in motion JPEG2000.

Besides the enhancement of the quality scalability at the decoder side, this method could also be used in the encoder to reduce the computational load of the Tier-1 when compared to the use of PCRD. Experimental evidence suggests that reductions as large as 94% are achieved at 0.0625 bps [35], while the coding performance is the same as the one achieved in the decoder; only three methods in literature achieve similar speed-ups [9], [10], [17]. However, the proposed method is restricted to applications without memory constrained resources since wavelet data needs to be maintained in memory.

Acknowledgment

The authors would like to thank D. Taubman, M. Marcellin, P. Salembier and R.M. Figueras for their comments and suggestions in the review of the Ph.D. dissertation [35], which is the source of this paper, and to the anonymous referees for their valuable remarks.

References

[1] *Information technology - JPEG 2000 image coding system - Part 1: Core coding system*, ISO/IEC Std. 15 444-1, Dec. 2000.

[2] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Trans. Image Process.*, vol. 9, no. 7, pp. 1158–1170, Jul. 2000.

[3] *Information technology - JPEG 2000 image coding system: Interactivity tools, APIs and protocols*, ISO/IEC Std. 15 444-9, Dec. 2005.

[4] J. G. Apostolopoulos, "Secure media streaming & secure adaption for non-scalable video," presented at the Proc. IEEE International Conference on Image Processing, vol. 3, Oct. 2004, pp. 1763–1766.

[5] X. Wu, S. Dumitrescu, and N. Zhang, "On multirate optimality of JPEG2000 code stream," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2012–2023, Dec. 2005.

[6] H. Everett, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Oper. Res.*, vol. 11, pp. 399–417, 1963.

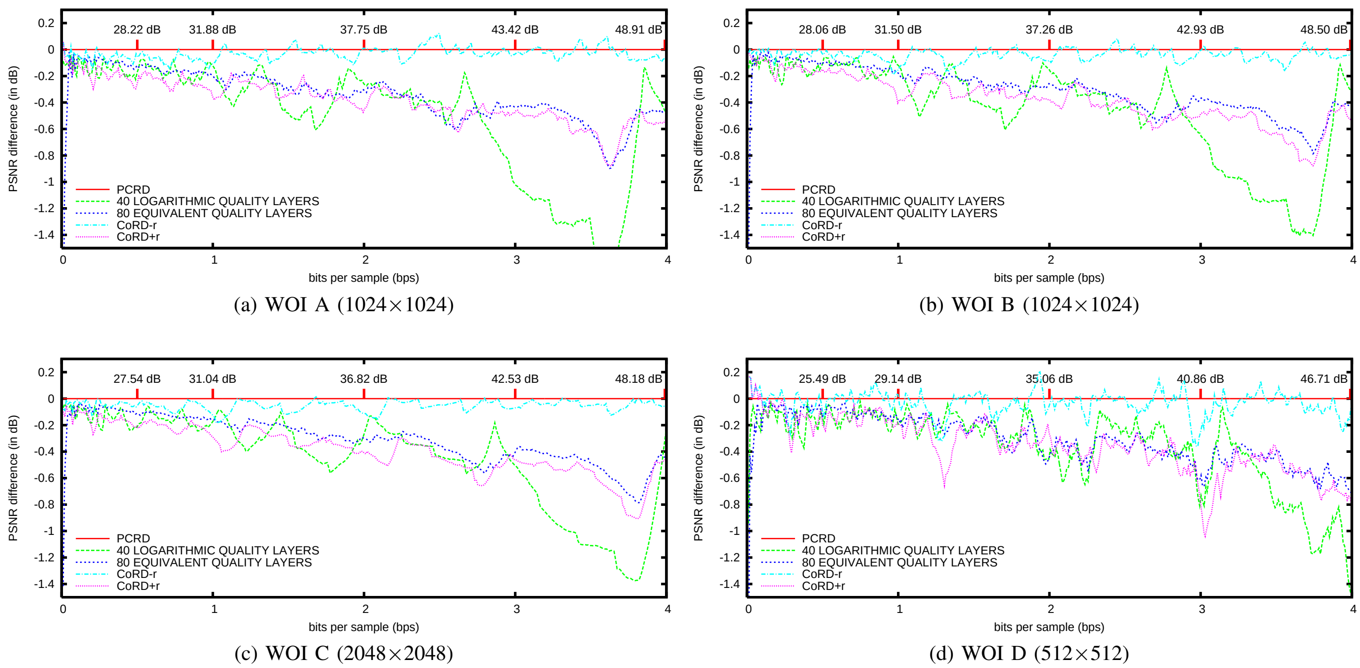
[7] T. Sanguankotchakorn and J. Fangtham, "A new approach to reduce encoding time in EBCOT algorithm for JPEG2000," presented at the Proc. IEEE Conference on Convergent Technologies for Asia-Pacific Region, vol. 4, Oct. 2003, pp. 1338–1342.

[8] T. Kim, H. M. Kim, P.-S. Tsai, and T. Acharya, "Memory efficient progressive rate-distortion algorithm for JPEG 2000," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 1, pp. 181–187, Jan. 2005.

[9] Y. Yeung and O. C. Au, "Efficient rate control for JPEG2000 image coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 3, pp. 335–344, Mar. 2005.

[10] W. Yu, F. Sun, and J. E. Fritts, "Efficient rate control for JPEG-2000," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 5, pp. 577–589, May 2006.

[11] T.-H. Chang, L.-L. Chen, C.-J. Lian, H.-H. Chen, and L.-G. Chen, "Computation reduction technique for lossy JPEG2000 encoding through EBCOT tier-2 feedback processing," presented at the Proc. IEEE International Conference on Image Processing, vol. 3, Jun. 2002, pp. 85–88.



(a) WOI A (1024×1024)

(b) WOI B (1024×1024)

(c) WOI C (2048×2048)

(d) WOI D (512×512)

Fig. 11: Evaluation of the coding performance achieved with CoRD and the use of quality layers compared to PCRD when decoding four WOIs from a JPEG2000 code-stream.

[12] D. Taubman, “Software architectures for JPEG2000,” presented at the Proc. IEEE International Conference on Digital Signal Processing, vol. 1, Jul. 2002, pp. 197–200.

[13] W. Du, J. Sun, and Q. Ni, “Fast and efficient rate control approach for JPEG2000,” *IEEE Trans. Consum. Electron.*, vol. 50, no. 4, pp. 1218–1221, Nov. 2004.

[14] F. Chebil and R. Kurceren, “Pre-compression rate allocation for JPEG2000 encoders in power constrained devices,” presented at the Proc. SPIE International Conference on Image and Video Communications and Processing, Jan. 2006.

[15] T. Masuzaki, H. Tsutsui, T. Izumi, T. Onoye, and Y. Nakamura, “JPEG2000 adaptive rate control for embedded systems,” presented at the Proc. IEEE International Symposium on Circuits and Systems, vol. 4, May 2002, pp. 333–336.

[16] X. Qin, X.-L. Yan, X. Zhao, C. Yang, and Y. Yang, “A simplified model of delta-distortion for JPEG2000 rate control,” presented at the Proc. IEEE International Conference on Communications, Circuits and Systems, Jun. 2004, pp. 548–552.

[17] K. Vikram, V. Vasudevan, and S. Srinivasan, “Rate-distortion estimation for fast JPEG2000 compression at low bit-rates,” *IEE Electronics Letters*, vol. 41, no. 1, pp. 16–18, Jan. 2005.

[18] C. Parisot, M. Antonini, and M. Barlaud, “High performance coding using a model-based bit allocation with EBCOT,” presented at the Proc. of EURASIP XI European Signal Processing Conference, Sep. 2002.

[19] M. D. Gaubatz and S. S. Hemami, “Robust rate-control for wavelet-based image coding via conditional probability models,” *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 649–663, Mar. 2007.

[20] A. Aminlou and O. Fatemi, “A novel efficient rate control algorithm for hardware implementation in JPEG2000,” presented at the Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, Mar. 2005, pp. 21–24.

[21] Y.-Z. Zhang and C. Xu, “Analysis and effective parallel technique for rate-distortion optimization in JPEG2000,” presented at the Proc. IEEE International Conference on Image Processing, Oct. 2006, pp. 2465–2468.

[22] C. Chan, S. Wee, and J. Apostolopoulos, “On optimal embedded schedules of JPEG-2000 packets,” presented at the Proc. IEEE International Conference on Image Processing, Oct. 2006, pp. 785–788.

[23] A. K. Gupta, D. Taubman, and S. Nooshabadi, “Near-optimal low-cost distortion estimation technique for JPEG2000 encoder,” presented at the Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 3, May 2006, pp. 952–955.

[24] Y.-W. Chang, H.-C. Fang, C.-J. Lian, and L.-G. Chen, “Novel pre-compression rate-distortion optimization algorithm for JPEG 2000,” presented at the Proc. SPIE International Conference on Visual Communications and Image Processing, Jan. 2004, pp. 1353–1361.

[25] Z. Liu, L. J. Karam, and A. B. Watson, “JPEG2000 encoding with perceptual distortion control,” *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 1763–1778, Jul. 2006.

[26] Y.-W. Chang, H.-C. Fang, C.-C. Cheng, C.-C. Chen, and L.-G. Chen, “Precompression quality-control algorithm for JPEG 2000,” *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3279–3293, Nov. 2006.

[27] P. Kulkarni, A. Bilgin, M. W. Marcellin, J. C. Dagher, J. H. Kasner, T. J. Flohr, and J. C. Rountree, “Compression of earth science data with JPEG2000,” in *Hyperspectral Data Compression*. USA: Springer, 2006, ch. 12, pp. 347–378.

[28] O. M. Kosheleva, B. E. Usevitch, S. D. Cabrera, and E. Vidal, “Rate distortion optimal bit allocation methods for volumetric data using JPEG 2000,” *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2106–2112, Aug. 2006.

[29] J. C. Dagher, A. Bilgin, and M. W. Marcellin, “Resource-constrained rate control for motion JPEG2000,” *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1522–1529, Dec. 2003.

[30] A. Descampe, F.-O. Devaux, G. Rouvroy, J.-D. Legat, J.-J. Quisquater, and B. Macq, “A flexible hardware JPEG 2000 decoder for digital cinema,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 11, pp. 1397–1410, Nov. 2006.

[31] M. D. Smith and J. Villasenor, “Constant quality JPEG2000 rate control for digital cinema,” presented at the Proc. SPIE International Conference on Visual Communications and Image Processing, Jan. 2007.

[32] S. Battiato, A. Buemi, G. Impoco, and M. Mancuso, “JPEG2000 coded images optimization using a content-dependent approach,” *IEEE Trans. Consum. Electron.*, vol. 48, no. 3, pp. 400–408, Aug. 2002.

[33] E. Ardizzone, M. L. Cascia, and F. Testa, “A new algorithm for bit rate allocation in JPEG2000 tile encoding,” presented at the Proc. IEEE International Conference on Image Analysis and Processing, Sep. 2003, pp. 658–661.

[34] Z. Wu and N. Zheng, “Efficient rate-control system with three stages for JPEG2000 image coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 9, pp. 1063–1073, Sep. 2006.

[35] F. Auli-Llinas, “Model-based JPEG2000 rate control methods,” Ph.D. dissertation, Universitat Autònoma de Barcelona, Dec. 2006, available: http://www.deic.uab.cat/∼francesc/docs/auli-phd.pdf.

[36] D. S. Taubman and M. W. Marcellin, *JPEG2000 Image compression*

fundamentals, standards and practice. Norwell, Massachusetts 02061 USA: Kluwer Academic Publishers, 2002.

[37] D. Taubman and R. Rosenbaum, "Rate-distortion optimized interactive browsing of JPEG2000 images," presented at the Proc. IEEE International Conference on Image Processing, vol. 3, Sep. 2003, pp. 765–768.

[38] K. Krishnan, M. W. Marcellin, A. Bilgin, and M. S. Nadar, "Efficient transmission of compressed data for remote volume visualization," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1189–1199, Sep. 2006.

[39] A. Descampe, C. D. Vleeschouwer, M. Iregui, B. Macq, and F. Marques, "Prefetching and caching strategies for remote and interactive browsing of JPEG2000 images," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1339–1354, May 2007.

[40] D. S. Taubman, "Localized distortion estimation from already compressed JPEG2000 images," presented at the Proc. IEEE International Conference on Image Processing, Oct. 2006, pp. 3089–3092.

[41] W. Yu, "Integrated rate control and entropy coding for JPEG 2000," presented at the Proc. IEEE Data Compression Conference, Mar. 2004, pp. 152–161.

[42] F. Auli-Llinas, J. Serra-Sagrista, J. L. Monteagudo-Pereira, and J. Bartrina-Rapesta, "Efficient rate control for JPEG2000 coder and decoder," presented at the Proc. IEEE Data Compression Conference, Mar. 2006, pp. 282–291.

[43] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 243–250, Jun. 1996.

[44] F. Auli-Llinas and J. Serra-Sagrista, "Low complexity JPEG2000 rate control through reverse subband scanning order and coding passes concatenation," *IEEE Signal Process. Lett.*, vol. 14, no. 4, pp. 251–254, Apr. 2007.

[45] *Graphic technology - Prepress digital data exchange - Part 1: CMYK standard colour image data (CMYK/SCID)*, ISO/IEC Std. 12 640-1, 1997.

[46] J. Li and S. Lei, "An embedded still image coder with rate-distortion optimization," *IEEE Trans. Image Process.*, vol. 8, no. 7, pp. 913–924, Jul. 1999.

[47] D. Taubman. (2007, May) Kakadu software. [Online]. Available: http://www.kakadusoftware.com

[48] Group on Interactive Coding of Images. (2007, May) BOI software. [Online]. Available: http://www.gici.uab.cat/BOI

[49] Institut Cartografic de Catalunya. (2007, May) ICC. Barcelona 08038 (Spain). [Online]. Available: http://www.icc.cat

[50] Corporacio Parc Tauli. (2007, May) UDIAT centre diagnostic. Sabadell 08208 (Spain). [Online]. Available: http://www.parctauli.cat/webcspt/udiat

Francesc Aulí-Llinàs (S'2006-M'2008) received the B.Sc. and B.E. degrees in Computer Management Engineering and Computer Engineering in 2000 and 2002 respectively, both with highest honours from the Universitat Autònoma de Barcelona, Spain. From 2002 to 2006, he held a doctoral fellowship funded by the Catalan Government, receiving the M.S. and Ph.D. degrees in Computer Science in 2004 and 2006 respectively, also from the Universitat Autònoma de Barcelona. He is the main developer of BOI, a free JPEG2000 Part 1 implementation that has been awarded by the Catalan Government. Currently, he holds a postdoctoral fellowship funded by the Spanish Government to carry out research stages with the group of David Taubman in the University of New South Wales (Australia), and with the group of Michael Marcellin in the University of Arizona (USA). His research interests include highly scalable image and video coding systems, rate-distortion optimization techniques, and interactive image and video transmission.

Joan Serra-Sagristà (S'97-M'05) received the B.S., M.S., and Ph.D. degrees in Computer Science from the Universitat Autònoma de Barcelona, Spain, in 1992, 1994, and 1999, respectively. Since 1992 he has been with the Department of Information and Communications Engineering, at the Universitat Autònoma de Barcelona, Spain, where he is currently an Associate Professor and Director of the Group on Interactive Coding of Images. From September 1997 to December 1998, he hold a DAAD research grant at the University of Bonn, Germany. From June to July 2000 he was a visiting researcher at the U.Bonn, Germany. His current research interests include image coding, data compression, vector quantization, and wavelet-based techniques, with special attention to remote sensing and telemedicine applications. He has coauthored several papers in these areas. Dr. Serra-Sagristà is member of SPIE. He has served on the steering committee and on technical program committees of several international conferences, and is a Reviewer for the major international journals in his research field.