ENHANCED QUALITY SCALABILITY FOR JPEG2000 CODE-STREAMS BY THE CHARACTERIZATION OF THE RATE-DISTORTION SLOPE

Francesc Aulí-Llinàs, Joan Serra-Sagristà, Joan Bartrina-Rapesta, Jose Lino Monteagudo-Pereira [†]

Universitat Autònoma de Barcelona (SPAIN) - Dep. of Information and Communications Engineering

ABSTRACT

Quality scalability is a fundamental feature of JPEG2000, achieved through the use of quality layers. Two points, related with the use of quality layers, may need to be addressed when dealing with JPEG-2000 code-streams: 1) the lack of quality scalability of single quality layer code-streams, and 2) the non rate-distortion optimality of windows of interest transmission. This paper introduces a new rate control method that can be applied to already encoded code-streams, addressing these two points. Its main key-feature is a novel characterization that can fairly estimate the rate-distortion slope of the coding passes of code-blocks without using any measure based on the original image or related with the encoding process. Experimental results suggest that the proposed method is able to supply quality scalability to already encoded code-streams achieving a near-optimal coding performance. The low computational costs of the method makes it suitable for use in interactive transmissions.

Index Terms— Image coding, image communication

1. INTRODUCTION

1.1. Quality scalability in JPEG2000

Quality scalability and quality progression are fundamental features of image coding systems. The quality progression, for example, allows the truncation of a code-stream at different bit-rates without penalizing the quality of the decoded image. The quality scalability, for example, is needed in interactive image transmissions to allow the delivery of Windows of Interest (WOI) at increasing qualities.

The JPEG2000 standard [2] achieves quality scalability and quality progression through the use of quality layers and smart progression orders. The rich syntax of the code-stream allows the identification of the quality layers and, using a quality primary progression order, the decoding of the code-stream at the quality layer boundaries yields an optimal rate-distortion representation of the image. Within a JPEG2000 code-stream, one quality layer is a collection of packets, where each packet contains the code-stream segments, possibly null, belonging to some coding passes of the code-blocks within one precinct. The tier-1 coding stage of JPEG2000 encodes each code-block independently, and the fractional bit-plane coder defines three coding passes for each bit-plane, referred to as SPP for the Significance Propagation Pass, MRP for the Magnitude Refinement Pass, and CP for the Cleanup Pass.

The definition of quality layers is an efficient mechanism to achieve quality scalability. However, their use needs to address two points. The first one is that code-streams containing a single quality layer have a lack of quality scalability and quality progression. If the code-stream needs to be truncated or interactively transmitted, the coding performance of the decoded images can be as large as 10 dB worse than when the code-stream contains an adequate number of quality layers [1].

The second point comes up when the image is interactively transmitted. Even though the code-stream may contain an adequate number of quality layers, these layers are constructed to optimize the rate-distortion of the complete spatial area of the image but, in interactive image transmissions, only the WOIs required by the client are transmitted and decoded. In this case, the distribution of the code-stream segments belonging to the WOIs may be not well distributed among quality layers, penalizing the rate-distortion optimality of the WOIs transmission [3].

1.2. Rate-distortion optimization

The rate-distortion optimality of a code-stream is determined by the number, and by the bit-rate allocation, of its quality layers. A smart allocation algorithm has been recently proposed in [4] to optimize a code-stream under an expected multi-rate distortion measure. However, although the allocation of quality layers is optimal, they may still be not well suited for WOIs transmission.

On the other hand, one might think that the lack of quality scalability and progression of single quality layer code-streams could be overcome re-building the code-stream, adding more quality layers. However, once the code-stream is already encoded, the number and allocation of quality layers is fixed without possibility of modifications. The following three paragraphs clarify this statement.

Typically, quality layers are constructed in the encoding process. An important step of their construction is the selection of the code-stream segments included in each one. This step is carried out by the rate-distortion optimization method used in the coder, which is able to optimize the quality for a target bit-rate or, conversely, to optimize the bit-rate for a target quality. Given the bit-rates at which to allocate the quality layers, or given the bit-rate of the overall code-stream, the rate-distortion optimization method selects the optimal code-stream segments, in terms of rate-distortion, for each target bit-rate.

The main reference of rate-distortion optimization in JPEG2000 is the Post Compression Rate-Distortion optimization (PCRD) method introduced in EBCOT [5]. It achieves optimal results but, in its original formulation, it compels to encode the complete image even when few coding passes are included in the final code-stream. With the aim to reduce the computational load of the tier-1 coding stage when applying the PCRD method, more than 26 different rate-distortion optimization methods have been proposed in the last five years. Some of them are [6, 7, 8, 9, 10, 11, 12] (an extensive review and comparison among several methods can be found in [1]).

Almost all of these methods use distortion measures based on

---

[*] This work has been partially supported by the Spanish Government (MEC), by FEDER, and by the Catalan Government, under Grants TSI2006-14005-C02-01 and SGR2005-00319.

[†] The authors would like to thank D. Taubman, M. Marcellin, P. Salembier and R.M. Figueras for their valuable comments and suggestions in the review of the Ph.D. dissertation [1], which is the source of this work.

the original image or information related to the encoding process. Therefore, none of them can be used to re-build the number and allocation of quality layers once the code-stream is already encoded. Only the method described in [13], that can avoid the use of distortion measures based on the original image, could be used; however, it penalizes the coding performance more than 1 dB when compared to the optimal PCRD.

The research presented in this paper is motivated because, neither a good allocation strategy of quality layers, nor any of the rate-distortion optimization methods of the literature, are able to properly overcome the two points related with the use of quality layers.

We have recently presented two rate control methods able to provide quality scalability to already encoded code-streams, even if the code-streams contain a single or few quality layers. The Coding Passes Interleaving (CPI) method [14] uses a fixed scanning order that includes first those coding passes that are at the highest bit-planes. The drawback of CPI is that its coding performance fluctuates continuously from 0.001 to more than 0.5 dB worse than the optimal PCRD method. The Reverse subband scanning Order and coding passes Concatenation (ROC) method [15] introduces three simple modifications to CPI in order to achieve a regular coding performance among all bit-rates. This is commonly achieved but, for some images, ROC does not improve the coding performance of CPI [1].

Section 2 introduces a novel characterization of the rate-distortion slope that does not need any distortion measure based on the original image. This characterization is used to develop a new rate control method able to provide quality scalability to already encoded code-streams, outperforming the results achieved by both CPI and ROC, as assessed in Section 3. A simple implementation strategy to reduce the computational costs of the method is also proposed in Section 3. Last section summarizes this work.

2. THE CORD RATE CONTROL METHOD

2.1. Characterization of the rate-distortion slope

The approach used in the presented rate control method is similar to the one used by the optimal PCRD method, but estimating, instead of actually computing, the rate-distortion slope.

The characterization of the rate-distortion slope is based on two important characteristics of the encoding process of a code-block. The first one 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, the rate control method should *concatenate the coding pass* MRP of a code-block with the following CP. We explain this characteristic using the rate-distortion model from [9], which estimates the decrement in distortion and the increment in bit-rate of a code-block at the 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 ($p = 0$ denotes the lowest bit-plane). Taking into account that the CP has a run mode that might encode four insignificant coefficients with a single bit, we slightly modify this model to distinguish between the coding passes of type MRP and CP as follows:

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

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 1 shows, at different bit-planes and for coding passes of type CP, the percentage of coefficients that have become significant. Percentages are reported on average for the code-blocks sets formed by code-blocks with the same number of magnitude bit-planes within a subband. Each one of these code-block sets is denoted by $b_{l,s}$ and K, where K stands for the number of magnitude bit-planes, l stands for the resolution level ($l = 0$ denotes the lowest one) and s stands for the subband, with $s = \{0 \text{ for HL/LH subbands}, \ 1 \text{ for HH subband}\}$.

Table 1. *For coding passes of type CP, percentage of coefficients becoming significant; Musicians image of the ISO/IEC 12640-1 corpus (2048x2560, gray scaled, lossy compression, 5 DWT levels).*

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%

The second characteristic of the encoding process of a code-block is named the *balloon effect*. The balloon effect is based on the following assumption: the coding passes that encode the largest number of significant coefficients have the greatest rate-distortion slope values. This assumption relies on the meaningful difference between the large decrement in distortion, compared to the small increment in bit-rate, when a significant coefficient is encoded. For coding passes of type SPP and CP, the rate-distortion model proposed in [9] estimates this according to

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

and, although the encoding of a significant coefficient increases twice the estimated bit-rate of the encoding of an insignificant coefficient, $\triangle D$ 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 is, specially at high bit-planes.

Our purpose is to study the distribution of the number of significant coefficients encoded from the highest to the lowest bit-plane of a code-block. Table 2 shows the average 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 issue 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 for each K, 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. We can see this as a balloon, where the width represents the number of significant coefficients encoded at the bit-plane.

The second remarkable issue of this table is that the number of significant coefficients at the highest bit-plane depends on the magnitude bit-planes of the code-block. This is, at their highest bit-plane, the code-blocks that have the lowest number of magnitude bit-planes encode more significant coefficients. In Table 2 this relation is always respected.

Table 2. *Average number of significant coefficients encoded in each bit-plane of the code-block sets within the subband $b_{3,0}$; Candle image of the ISO/IEC 12640-1 corpus (2048x2560, gray scaled, lossy compression, 5 DWT 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

2.2. Algorithm

Based on the characterization above, we are able to compute a theoretical rate-distortion slope for every coding pass of all the code-blocks of an image, just considering the number of magnitude bit-planes of the code-blocks within a subband. 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 bit-plane and coding pass unequivocally, computed as $c = 3p + cp$ with $cp = \{2 \text{ for SPP}, 1 \text{ for MRP}, \ 0 \text{ for CP}\}$. In order to assure that the 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 highest bit-plane, where $\mathcal{F}_{MRP} = 0.99$. For coding passes of type SPP and CP, $\mathcal{F}_{\{SPP|CP\}}$ represents the balloon effect within each subband, and is calculated as

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

where K_{max} denotes the maximum K in the subband to which the code-block belongs. 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.

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

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 are given in Table 3, where $\#K = K_{max} - K_{min}$ with K_{min} denoting the minimum K of the subband to which the code-block belongs. These choices have been determined experimentally.

Knowing the theoretical rate-distortion slope and the bit-rate of coding passes, the selection of coding passes to yield a target bit-rate can be performed straightforward. The rate control method that uses the Characterization of the Rate-Distortion slope is named CoRD.

Table 3. *Choices of parameters $\mathcal{F}_{init}$, $\mathcal{F}_{inc}$ and $\mathcal{F}_{dec}$.*

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

3. EXPERIMENTAL RESULTS

In order to assess the coding performance of CoRD, we present first a comparison among CPI, ROC and CoRD. In this comparison, the Fruit Basket image of the ISO/IEC 12640-1 corpus has been encoded with a single quality layer code-stream and then it has been decoded using CPI, ROC and CoRD at 600 uniformly distributed bit-rates along 0.001 to 5 bps, computing the PSNR between the decoded image and the original one. At the same bit-rates, the image has been encoded using the PCRD method, computing the PSNR difference between PCRD and CPI/ROC/CoRD. Recall that the PCRD method is applied in the encoding process, while CPI, ROC and CoRD are applied once the code-stream is already encoded. In all graphics the coding performance of PCRD is depicted with the top straight line to identify the maximum coding performance that can be obtained with JPEG2000.

In all experiments, Kakadu v4.5 has been used to construct the code-stream with the optimal PCRD method, and CPI, ROC and CoRD have been implemented in BOI v1.2[1]. The parameters of both applications are set to: lossy compression, 5 levels of DWT, derived quantization, code-blocks of size 64×64 and the RESTART coding variation. This coding variation is used to identify the bit-rate of coding passes just decoding the packet headings.

Figure 1 depicts the results obtained by CPI, ROC and CoRD when providing quality scalability to the single quality layer code-stream belonging to the Fruit Basket image. It is worth noting the regularity achieved by CoRD, compared to the continuous fluctuations of CPI and ROC. For this image, CoRD is, on average, only 0.048 dB worse than the optimal PCRD method.

The performance achieved by CoRD has to be compared too to the use of quality layers. This comparison considers two common allocation strategies of quality layers: 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 of quality layers allocation, each image of the ISO/IEC 12640-1 corpus has been encoded containing 20, 40, 80 and 120 quality layers. Then, the code-streams have been decoded at 600 uniformly distributed bit-rates, computing the PSNR with the original images. Figure 2 depicts the best results obtained by both allocation strategies, and by CoRD applied to single quality layer code-streams. This figure reports the average among all images of the corpus. From 0.001 to 5 bps, the equivalently spaced quality layers obtain, on average, a coding performance 0.058 dB worse than the PCRD method. The difference between PCRD and CoRD is, on average, 0.051 dB.

The above results suggest that CoRD is able to provide quality scalability to code-streams achieving near-optimal results. However, in order to apply CoRD in an interactive transmission using, for instance, the JPIP protocol defined in JPEG2000 Part 9 [16], it is important that CoRD is able to extract WOIs using a very low compu-

---

[1] See http://www.kakadusoftware.com and http://www.gici.uab.cat/BOI .

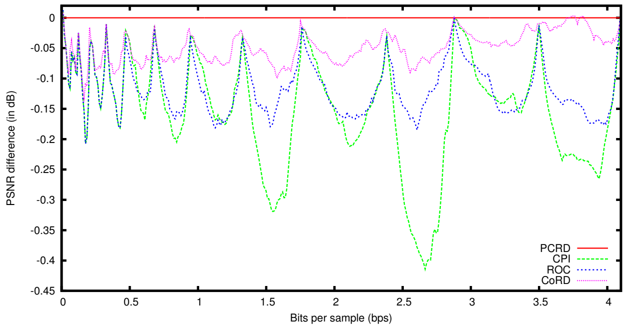

Fig. 1. Coding performance evaluation of the Fruit Basket image of the ISO/IEC 12640-1 image corpus (2048x2560, gray scaled).

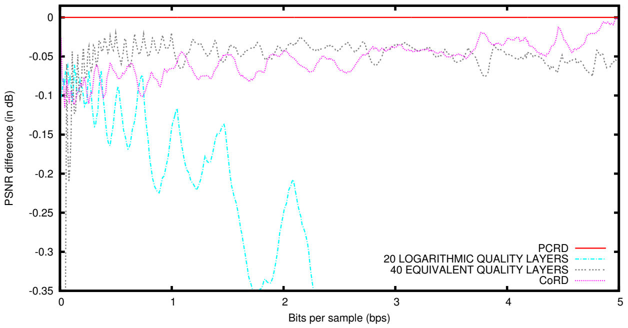

Fig. 2. Average coding performance evaluation of the ISO/IEC 12640-1 image corpus (2048x2560, gray scaled).

tational load. Although the approach used in CoRD is similar to the one used by the PCRD method, CoRD can widely reduce its computational costs taking into account that, all the code-blocks within a subband that have the same number magnitude bit-planes, are estimated equally. The implementation can group all these code-blocks as if they were a single one, reducing the number of code-blocks to consider when the method is applied. For the Cafeteria image of the ISO/IEC 12640-1 corpus, for instance, the PCRD method considers 5124 code-blocks of size 32×32; with this implementation strategy, CoRD considers only 57 different code-block sets. Besides, the single operation needed to apply CoRD is the decoding of the packet headings, but this usually takes less than 1% of the time employed to decode the whole code-stream and, in a client/server application, this operation should only be carried out at the very beginning of the connection.

On the other hand, CoRD can also be applied to the encoding process to reduce the computational load of the tier-1 coding stage. At a target bit-rate of 0.0625 bps, CoRD is able to reduce the computational load of the tier-1 coding stage in 94%. Compared to other rate control methods, CoRD achieves competitive results in terms of coding performance and computational complexity reduction [1]. However, the main drawback of CoRD applied to the encoding process is that it compels to maintain the image in memory to allow the stop and re-start encoding of code-blocks.

4. CONCLUSIONS

Aimed to address the lack of quality scalability of single quality layer code-streams and the non rate-distortion optimality of WOIs transmission, this paper presents a rate control method conceived from a novel Characterization of the Rate-Distortion slope (CoRD). The rate control method CoRD does not use distortion measures based on the original image or related with the encoding process, therefore it can be applied once the code-stream is already encoded. Experimental results suggest that CoRD is able to provide quality scalability to single quality layer code-streams achieving an efficiency close to the obtained with the use of quality layers. Besides, the very low computational complexity of CoRD makes it suitable to control interactive image transmissions.

5. REFERENCES

[1] F. Auli-Llinas, *Model-based JPEG2000 rate control methods*, Ph.D. thesis, Univ. Autònoma de Barcelona, Dec. 2006, Available: http://www.deic.uab.cat/∼francesc/docs/auli-phd.pdf.

[2] D. Taubman and M. Marcellin, *JPEG2000 Image compression fundamentals, standards and practice*, Kluwer Academic Publishers, Norwell, Massachusetts 02061 USA, 2002.

[3] D. Taubman and R. Rosenbaum, "Rate-distortion optimized interactive browsing of JPEG2000 images," in *Proc. IEEE ICIP*, Sept. 2003, pp. 765–768.

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

[5] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Trans. IP*, vol. 9, no. 7, pp. 1158–1170, July 2000.

[6] D. Taubman, "Software architectures for JPEG2000," in *Proc. IEEE DSP*, July 2002, vol. 1, pp. 197–200.

[7] C. Parisot, M. Antonini, and M. Barlaud, "High performance coding using a model-based bit allocation with EBCOT," in *Proc. of EURASIP XI*, Sept. 2002.

[8] W. Du, J. Sun, and Q. Ni, "Fast and efficient rate control approach for JPEG2000," *IEEE Trans. CE*, vol. 50, no. 4, pp. 1218–1221, Nov. 2004.

[9] K.N. Vikram, V. Vasudevan, and S. Srinivasan, "Rate distortion estimation for fast JPEG2000 compression at low bit-rates," *IEE Elec. Lett.*, vol. 41, no. 1, pp. 16–18, Jan. 2005.

[10] T. Kim, H. Mun Kim, P. Tsai, and T. Acharya, "Memory efficient progressive rate-distortion algorithm for JPEG2000," *IEEE Trans. CSVT*, vol. 15, no. 1, pp. 181–187, Jan. 2005.

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

[12] Y.M. Yeung and Oscar C. Au, "Efficient rate control for JPEG2000 image coding," *IEEE Trans. CSVT*, vol. 15, no. 3, pp. 335–344, Mar. 2005.

[13] T. Masuzaki, H. Tsutsui, T. Izumi, T. Onoye, and Y. Nakamura, "JPEG2000 adaptive rate control for embedded systems," in *Proc. IEEE ISCS*, May 2002, vol. 4, pp. 333–336.

[14] F. Auli-Llinas, J. Serra-Sagrista, J.L. Monteagudo-Pereira, and J. Bartrina-Rapesta, "Efficient rate control for JPEG2000 coder and decoder," in *Proc. IEEE DCC*, Mar. 2006, pp. 282–291.

[15] F. Auli-Llinas and J. Serra-Sagrista, "Low complexity JPEG-2000 rate control through reverse subband scanning order and coding passes concatenation," *IEEE Signal Processing Letters*, Apr. 2007, in Press.

[16] ISO/IEC 15444-9, "Information technology - JPEG2000 image coding system - Part 9: Interactivity tools, APIs and protocols," Dec. 2005.