Low Complexity JPEG2000 Rate Control through Reverse Subband Scanning Order and Coding Passes Concatenation

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

***Abstract*— This letter introduces a new rate control method devised to provide quality scalability to JPEG2000 codestreams containing a single or few quality layers. It is based on a Reverse subband scanning Order and a coding passes Concatenation (ROC) that does not use distortion measures based on the original image. The proposed ROC method allows a flexible rate control when the image has already been encoded, using negligible computational resources and obtaining the same efficiency as when using quality layers. Besides, the proposed ROC can be used in the encoding process to reduce the coder complexity, avoiding to encode unnecessary coding passes and achieving a competitive performance in terms of MSE.**

***Index Terms*— JPEG2000 standard, rate distortion optimization, quality scalability, interactive transmissions.**

I. Introduction

JPEG2000 is a powerful standard to encode, transmit and manipulate images. The Part 1 [1] of the standard describes the core coding system and the specification of the file syntax. The JPEG2000 coding scheme is wavelet based with a two-tiered coding strategy built on an Embedded Block Coding with Optimized Truncation (EBCOT) [2]. The tier-1 stage includes a fractional bit plane coder and the MQ arithmetic coder, and the tier-2 stage considers the coding of block contributions to each quality layer. One valuable capability of the coding system is the ability to create a codestream at a target bitrate providing the best recovering. Typically, the codestream construction is controlled by the rate control method of the encoding process, which uses distortion measures based on the original image in order to identify the distortion contribution of each truncation point. Using these distortion measures, the Post Compression Rate Distortion (PCRD) method of EBCOT describes how to obtain the best codestream for a target bitrate. Although this process achieves the optimal results, for instance in terms of Mean Square Error (MSE), in its original formulation it lacks in efficiency because it compels to encode the whole image even if only some coding passes are included in the final codestream. Considering that the tier-1 stage represents more than 60% of the encoding process [3], the optimization of the tier-1 stage could widely decrease the whole coder complexity.

New rate control strategies have been developed in the last four years that need to encode fewer coding passes than the optimal PCRD method while maintaining the coding performance. An efficient heap-based rate allocation algorithm that allows a selective encoding of the code-blocks included in the final codestream is presented in [4]; several non-obvious implementation strategies for software architectures are described in [3]; in [5] the optimal quantization step sizes for a target MSE are computed before encoding; an efficient slope computation with an interleaving strategy is used in [6] to select the coding passes to be encoded; in [7] two algorithms are proposed that efficiently control the rate distortion of multi-component images using an incremental encoding; in [8] the rate distortion is computed before the encoding process thanks to a prediction of the truncation points lengths; in [9] an accumulation of the code-block slopes is used to identify when to stop encoding; in [10] three efficient methods using different computational resources are proposed. A comparative table among several rate control methods is also provided in [8]. At a target bitrate of 0.0625 bits per pixel (bpp), some of these techniques can save more than 94% of the time spent by the tier-1 stage, while the coding performance is reduced about only 0.1 dB compared to the optimal PCRD method.

Another important capability of the JPEG2000 standard is the ability to manipulate a codestream without needing to re-encode the image. This capability is related to three important features [11]: the resolution scalability, the spatial random access, and the quality scalability. The resolution scalability is supplied by the dyadic decomposition of the wavelet transform in subbands grouped into resolution levels. The spatial random access is supplied by the division of each subband in small blocks of coefficients that are encoded independently and organized into precincts. The quality scalability is supplied by the quality layers. These quality layers provide a good recovering when decoding only a segment of the whole codestream and they allow to retrieve selected spatial regions at different qualities.

All the rate control methods described above allow the construction of quality layers. Their construction does not involve additional computational resources and entails minor costs for the coding performance, therefore, quality layers should always be used. However, the standard does not establish the number and bitrate of the quality layers that a codestream should contain, letting the user specify them. A

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

This work was supported in part by the Spanish Government (MEC), by FEDER, and by the Catalan Government, under Grants TIC2003-08604-C04-01, TSI2006-14005-C02-01, SGR2005-00319, and 2006FIC-00189.

recent study [12] reexamines the rate distortion optimality of a JPEG2000 codestream under an expected multirate distortion measure considering uniform, exponential and Laplacian rate distribution functions. In this approach the quality layers construction uses dynamic programming, obtaining an optimal codestream under the expected multirate distortion measure.

The above rate control methods are all devised for the encoding process and most of them use distortion measures based on the original image, therefore, once the codestream is constructed, there is no possibility to reconstruct the quality layers. Only if the whole encoding process is performed again –which implies a great effort–, these rate control methods could be used.

The main purpose of this research is to obtain a rate control method which provides quality scalability for already encoded JPEG2000 codestreams, even if they contain a single or few quality layers. The method is introduced in Section II an it provides an efficient mechanism to control the rate distortion of a codestream without needing to re-encode the whole image. In order to assess the performance of the proposed method, in Section III we present some experimental results. Section IV contains the conclusions.

II. Analysis of the PCRD method and design of the new rate control strategy

A. PCRD compared to a simple interleaving strategy

We have recently proposed the rate control method Coding Passes Interleaving (CPI) [13]. This simple interleaving strategy selects the coding passes included in the final codestream using a fixed scanning order based on the coding levels (to be defined next). The fractional bit plane coder of JPEG2000 encodes each bit plane in three coding passes: the Significance Propagation Pass (SPP), the Magnitude Refinement Pass (MRP) and the Cleanup Pass (CP). The coding level $c_{i,j}$ of a code-block B_i is a fractional bit plane, with $j = (bitPlane * 3) + CoPType$, where $CoPType = \{SPP$=2, MRP=1, CP=0$\}$. C_i stands for the highest coding level of the code-block B_i, while the highest and lowest coding level of the image will be referred to as $C_{max} = max(C_i)$ and $C_{min} = 0$ respectively. CPI encodes the coding passes of code-blocks belonging to the same coding level, from C_{max} to C_{min} until the target bitrate is achieved. In each coding level, the coding passes are selected from the lowest resolution level $\mathcal{L}_0$ to the highest resolution level $\mathcal{L}_L$. The set of subbands belonging to the resolution level $\mathcal{L}$ are referred to as $b_{\mathcal{L},s}$ and they are scanned following the order $sOrder = [HL, LH, HH]$. Table I (top) formulates the CPI algorithm for the encoding process when a target bitrate has to be attained.

There is no use of any distortion measure based on the original image. Therefore, CPI allows the division of quality layers and the extraction of selected image areas at different qualities from a codestream containing one or more quality layers, thus solving the problem that can cause the lack of quality scalability in single or few quality layers codestreams.

The main theoretical assumption behind CPI is the spread belief that the coding passes situated at high bit planes recover the original image better than those situated at low bit planes.

TABLE I

Top: CPI algorithm. Bottom: CPI simple modifications (ROC)

set *bitRate* $\leftarrow$ *0*
for each *coding level* C **from** C_{max} **to** C_{min} **do**
  for each *resolution level* $\mathcal{L}$ **from** $\mathcal{L}_0$ **to** $\mathcal{L}_L$ **do**
    set $sOrder \leftarrow \begin{cases} [LL] & \text{if } \mathcal{L} = \mathcal{L}_0 \\ [HL, LH, HH] & \text{otherwise} \end{cases}$
    for each *subband* $b_i \in b_{\mathcal{L},s}$ **following** *sOrder* **do**
      for each *code-block* $B_i \in b_i$ **do**
        ENCODE coding level C *of* B_i
        set *bitRate* $\leftarrow$ *bitRate + length(C)*
        if *bitRate* $\geq$ *targetBitRate* **then**
          STOP encoding
        endif
      endfor
    endfor
  endfor
endfor

From a particular coding level onward,

1. The scanning order of the resolution levels goes from $\mathcal{L}_L$ to $\mathcal{L}_0$.
2. The inclusion of a MRP is concatenated by the CP of the same code-block.
3. The scanning order of the coding levels with coding passes of type MRP or CP follows the $sOrder = [HH, LH, HL]$.

The scanning order followed by CPI is also used in other coding schemes that achieve a regular performance among all bitrates. Therefore, it may be expected that CPI should also obtain a coding performance similar to that of the optimal PCRD method, but it does not: the performance of CPI is not well-balanced and, for some bitrates, it is 0.5 dB worse. This fact arises two questions: when these differences occur and why they are produced.

The first question can be readily answered comparing the optimal PCRD method to CPI. All tests of this Section have been performed for all the eight images of the ISO/IEC Standard 12640-1 corpus. Although only the results obtained for the Cafeteria image are reported here, the remaining images have very similar results. In the first test, each image has been encoded at 2000 different bitrates using the optimal PCRD method and CPI. When computing the Peak Signal to Noise Ratio (PSNR) difference between both methods, all the images exhibit a similar coding performance among all bitrates: the performance of CPI fluctuates about 0.001 to 0.5 dB worse than the PCRD method. To better appreciate these irregularities, Figure 1(a) depicts these results. This detailed comparison answers the first question about when the differences happen: when scanning coding levels containing coding passes of type SPP, or when scanning two consecutive coding levels containing coding passes of type MRP and CP.

Identifying when the differences occur gives us the clue to answer why they are produced. Note that in Figure 1(a) the coding performance of the optimal PCRD method and CPI coincide in several bitrates. An accurate observation on these bitrates disclose that both methods select practically the same coding passes when CPI ends the scanning of a coding level that contains coding passes of type SPP, or at the end of a coding level that contains coding passes of type CP. Therefore, our attention is focused on these bitrates. We use the weighted Mean Square Error as the distortion measure, computing the distortion of the coding level $c_{i,j}$ of the code-block B_i as

TABLE II

ANALYSIS FOR THE CAFETERIA IMAGE. DISTORTION CONTRIBUTION OF EACH RESOLUTION LEVEL, EVALUATED PER CODING LEVELS.

cod. level	$\mathcal{L}_0$	$\mathcal{L}_1$	$\mathcal{L}_2$	$\mathcal{L}_3$	$\mathcal{L}_4$	$\mathcal{L}_5$
20 (SPP)	0.2%	2.4%	10.1%	**26.5%**	**43.6%**	**17.3%**
19 (MRP)	0.7%	1.1%	2.4%	4.3%	5.8%	1.9%
18 (CP)	0.0%	0.0%	0.7%	7.0%	**30.9%**	**45.2%**
17 (SPP)	0.0%	0.7%	4.0%	14.6%	**39.4%**	**41.3%**
16 (MRP)	0.5%	1.0%	2.7%	6.4%	11.8%	9.1%
15 (CP)	0.0%	0.0%	0.2%	2.2%	**14.8%**	**51.4%**
14 (SPP)	0.0%	0.3%	1.9%	8.7%	**30.4%**	**58.7%**
13 (MRP)	0.4%	0.9%	2.8%	7.6%	**17.7%**	**21.2%**
12 (CP)	0.0%	0.0%	0.1%	0.9%	7.3%	**41.3%**

$D_i^j = w_{b_i}^2 \sum_{k \in B_i} (\hat{s}_i^j[k] - s_i[k])^2$ where w_{b_i} is the weight of the subband b_i to which code-block B_i belongs, $s_i[k]$ are the original samples of the code-block B_i, and $\hat{s}_i^j[k]$ are the samples quantized at coding level j.

We evaluate the distortion contribution of each resolution level between the bitrates where the coding performance of the PCRD method and CPI coincide. Let D^j be the sum of distortions of all code-blocks of the image at the coding level j, denoting the distortion contribution of coding level j as $\triangle D^j = D^{j+1} - D^j$. Let $D^j_{\mathcal{L}}$ be the sum of distortions of all code-blocks belonging to the resolution level $\mathcal{L}$ at the coding level j, denoting the distortion contribution of $\mathcal{L}$ at the coding level j as $\triangle D^j_{\mathcal{L}} = D^{j+1}_{\mathcal{L}} - D^j_{\mathcal{L}}$. Between the bitrates at which both methods coincide, CPI scans a coding level with coding passes of type SPP, or two consecutive coding levels with coding passes of type MRP and CP. Table II considers the distortion contribution of each resolution level for these coding levels in percentage (i.e. $\frac{\triangle D^j_{\mathcal{L}}}{\triangle D^j}$)[1]. The resolution levels that have a major contribution in terms of rate distortion are emphasized in bold font. Note that when including two consecutive coding levels with coding passes of type MRP and CP, the coding levels with coding passes of type MRP usually have equal or minor contributions than the coding levels with coding passes of type CP. Note also that the major contributions in the listed coding levels are situated at resolution levels $\mathcal{L}_4$ and $\mathcal{L}_5$.

B. A more elaborated interleaving strategy

Based on the above analysis and on practical experimentation, we propose three modifications to CPI (Table I bottom). The incorporation of these modifications to CPI does not modify its main structure, allowing to keep the characteristics of the original CPI method while increasing its performance.

These modifications have to be applied to CPI at the bit planes where the distortion contribution of the highest resolution levels is larger than the distortion contribution of the lowest resolution levels. However, this approach would use distortion measures based on the original image, so it would be only useful at encoding time. Experimental evidence suggests that another suitable measure is the number of code-blocks that belong to the highest resolution level and have to be encoded in the current coding level: the modifications to CPI should be applied when at least 55% of the code-blocks belonging to $\mathcal{L}_L$ are included in the current coding level. The resulting algorithm will be referred to as Reverse subband scanning Order and coding passes Concatenation (ROC). To evaluate the performance of the proposed modifications, Figure 1(a) depicts the results obtained for the Cafeteria image. Note the improvement obtained for almost all bitrates.

Without needing to re-encode the image, ROC allows the reconstruction of a codestream modifying the number and bitrate of the quality layers that it contains, thus providing quality scalability to codestreams containing a single or few quality layers. The interleaving algorithm of ROC can be applied just decoding the packet headers, taking negligible costs in terms of memory consumption and computational complexity. Therefore, ROC can also be used to control the interactive image transmission of a single (or few) quality layers codestream –at the server side– without need to embed quality layers in it. When applying ROC in the encoding process it minimizes the time spent by the tier-1 stage, although the whole image as well as some information used by the MQ coder has to be maintained in memory.

III. EXPERIMENTAL RESULTS

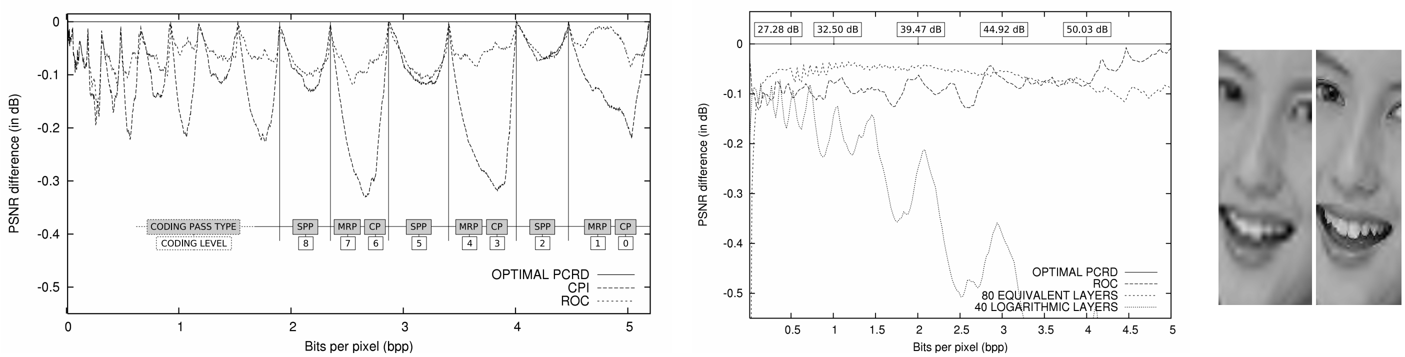
This Section reports two types of results: the coding performance of ROC compared to the optimal PCRD method, and the performance when decoding only a segment of a codestream, compared to the use of quality layers. The tests have been done for the eight images of the ISO 12640-1 corpus (gray scaled, size 2048∗2560), computing an average among them. Kakadu software (v4.5) has been used to construct codestreams with single or more quality layers, our BOI implementation of JPEG2000 (Part 1) for the proposed ROC[2].

To compare ROC with the optimal PCRD method, each image has been encoded at 200 uniformly distributed bitrates along 0.001 to 5 bpp using both ROC and PCRD methods, and the PSNR difference between them has been computed. The straight line of Figure 1(b) depicts the performance obtained by the optimal PCRD method; the ROC line depicts the performance of ROC method. Note the regular performance among all bitrates; the average difference is 0.077 dB!

To compare ROC with the use of quality layers, two common options of quality layers construction have been analyzed. The first option constructs the quality layers representing compressed bitrates logarithmically spaced along 0.001 to 5 bpp, yielding high performance at very high compression ratios. The second option constructs the quality layers representing equivalent bitrates and, in order to yield better coding performance at low bitrates, finer quality layers in terms of bitrate are distributed from 0.001 to 0.5 bpp, and coarser quality layers from 0.5 to 5 bpp. To perform the comparison, for both options, each image of the corpus has been encoded containing 20, 40, 80 and 120 quality layers.

[1]When including two consecutive coding levels with coding passes of type MRP and CP, the percentage is computed considering both coding levels.

[2]Compression options are the following: lossy mode of JPEG2000 with derived quantization and RESTART mode for the MQ coder to allow the identification of coding passes lengths, 5 levels of discrete wavelet transform. http://www.kakadusoftware.com, http://www.gici.uab.cat/BOI



(a) PCRD, CPI, and ROC performance

(b) PCRD, ROC, and quality layers performance

(c) Visual analysis

Fig. 1. (a) Detailed comparison among the optimal PCRD, CPI, and ROC for the Cafeteria image. The inclusion of the coding levels among different bitrates is depicted by the white boxes. (b) ROC compared to the optimal PCRD method and to the use of quality when decoding only a segment of the codestream. Average performance of the ISO 12640-1 corpus. (c) Detail of the Woman image recovered at 0.03125 bpp. Left: 1 layer, 23.07 dB; Right: 1 layer+ROC, 24.7 dB (same PSNR as with 40 logarithmic layers).

Then, the codestreams have been decoded at 200 uniformly distributed bitrates and the difference obtained compared to the optimal PCRD method when encoding at that particular target bitrate has been computed. Regarding the performance of ROC, a single quality layer codestream has been constructed and, for the same 200 bitrates, the proposed ROC method has selected the segment of the codestream to be decoded and then compared to the optimal PCRD method. Figure 1(b) depicts the obtained results. Obviously, the performance of ROC now is the same as the obtained by ROC at encoding time, thus both plots are drawn in one line. To ease the visual interpretation, only the best results obtained for both options of layer construction are plotted (i.e., 40 logarithmic quality layers and 80 equivalent quality layers). Note that the regular performance of ROC is very similar to the best choice of quality layers: at high compression ratios, ROC obtains a performance practically equivalent to both options of quality layers construction; at medium and low compression ratios, ROC performance is close to that obtained with equivalent layers. The largest differences between ROC and the best choice of quality layers do not exceed 0.06 dB and, for very low compression ratios, ROC obtains the best performance. The average PSNR difference obtained with the logarithmically and equivalently spaced quality layers compared to the optimal PCRD method is 0.55 dB and 0.074 dB respectively. However, the performance obtained at low bitrates is usually of major interest: from 0.001 bpp to 1 bpp the average PSNR difference of logarithmic and equivalent quality layers is 0.13 dB and 0.094 dB respectively; the average difference of ROC in this bitrate range is 0.096 dB. Regarding the qualitative analysis, Figure 1(c) provides a visual comparison for the Woman image (ISO 12640-1 corpus), at a compression factor of 256:1.

IV. Conclusions

In this letter a low complexity rate control method for JPEG2000 has been proposed. It is based on a simple interleaving strategy that uses a Reverse subband scanning Order and a coding passes Concatenation (ROC). The lack of distortion measures based on the original image allows the applicability of the proposed ROC to control the rate distortion of JPEG2000 codestreams, even if they contain a single quality layer. Experimental results suggest that ROC performance when decoding a segment of a codestream is comparable to the performance obtained with the use of quality layers. Taking negligible computational resources, ROC may represent an alternative to quality layers in some scenarios, in particular to control interactive image transmissions, for instance using the JPIP client-server protocol. Besides, the proposed ROC can be applied to the coder, constructing a compliant JPEG2000 codestream and reducing the computational complexity of the encoding process.

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 Processing*, vol. 9, no. 7, pp. 1158–1170, July 2000.

[3] ——, “Software architectures for JPEG2000,” in *Proc. IEEE DSP’02*, vol. 1, July 2002, pp. 197–200.

[4] 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.

[5] C. Parisot, M. Antonini, and M. Barlaud, “Stripe-based MSE control in image coding,” in *Proc. IEEE ICIP’02*, vol. 2, Sept. 2002, pp. 649–652.

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

[7] P. Kulkarni, A. Bilgin, M. Marcellin, J. C. Dagher, T. Flohr, and J. Rountree, “Reduced memory multi-layer multi-component rate allocation for JPEG2000,” in *Proc. SPIE IVCP’05*, vol. 5685, Jan. 2005, pp. 139–150.

[8] 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.

[9] 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.

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

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

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

[13] F. Auli-Llinas, J. Serra-Sagrista, J. Monteagudo-Pereira, and J. Bartrina-Rapesta, “Efficient rate control for JPEG2000 coder and decoder,” in *Proc. IEEE DCC’06*, Mar. 2006, pp. 282–291.