JPEG2000 ROI Coding with Fine-Grain Accuracy through Rate-Distortion Optimization Techniques

Joan Bartrina-Rapesta, *Student Member, IEEE*,

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

Abstract—Region Of Interest (ROI) coding is a feature of prominent image coding systems that enables the specification of different coding priorities to certain regions of the image. JPEG2000 provides ROI coding through two mechanisms: either modifying wavelet coefficients, or using rate-distortion optimization techniques. Although ROI coding methods based on the modification of wavelet coefficients provide an excellent accuracy to delimit the ROI area (referred to as fine-grain accuracy), they significantly penalize the coding efficiency. On the other hand, methods based on rate-distortion optimization improve the coding efficiency, but, so far, have not been able to achieve the intended fine-grain accuracy. This paper introduces two ROI coding methods that, using rate-distortion optimization techniques, achieve a fine-grain accuracy, comparable to the one obtained when wavelet coefficients are modified, and are competitive in terms of coding efficiency.

Index Terms—Region of interest coding, JPEG2000, rate-distortion optimization.

I. INTRODUCTION

REGION Of Interest (ROI) coding is a feature provided by modern image coding systems that allows the prioritization of specific ROIs over the rest of the image (the background). The main idea behind ROI coding is to increase the priority of the ROI *within* the encoding process, generating a compressed codestream that, decoded at increasing bitrates, recovers the ROI first and with a higher quality than the background.

Several image-based applications have the need to enhance the quality of specific regions over the rest of the image. For example, in geographic information systems [1], images may contain more meaningful regions; in the medical community [2], special images such as mammographies or computer tomographies contain selected and restricted regions of particular interest; in the remote browsing of images, improving the quality of featured regions is also beneficial [3].

JPEG2000 [4] is a powerful standard for the coding, transmission, and manipulation of images and video. The core coding system is built on the Embedded Block Coding with Optimized Truncation (EBCOT) [5] strategy that, among other characteristics, codes small sets of wavelet coefficients (called codeblocks) independently, generating one quality embedded bitstream for each. These bitstreams are optimally truncated
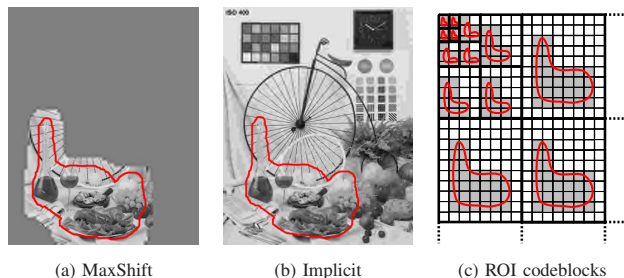
(a) MaxShift   (b) Implicit   (c) ROI codeblocks

Fig. 1. Bicycle image (size 2048×2560, gray-scaled) decoded at a bitrate of 0.01 bps, with a ROI (emphasized in red) prioritized at $\mathcal{R} = 15$.

using rate-distortion optimization techniques in order to attain a target bitrate, or a target quality, for the final codestream.

Natively, JPEG2000 provides two mechanisms for the ROI coding: the first one, performed right after the quantization stage, adjusts the priority of the ROI coefficients multiplying the corresponding quantized wavelet coefficients by a desired priority $\mathcal{R}$, which is commonly chosen to be a power of 2, so the multiplication is implemented as a bit-shift operation and is conceptually seen as a bitplane shift. The standard defines two methods using this *first mechanism*: the MaxShift and the Scaling. The main difference between them is that the Scaling allows the user to choose the ROI priority $\mathcal{R}$, whereas in MaxShift $\mathcal{R}$ is chosen to encode all ROI coefficients before the background, avoiding the transmission of the ROI's shape. Several modifications of these methods have been proposed to more precisely combine the ROI with the background: the Bitplane-by-Bitplane Shift [6] was the first proposing an interleaving of ROI and background bitplanes, and [1], [7], [8] introduced modifications to this interleaving strategy to allow richer combinations of the ROI with the background.

The *second mechanism* used for ROI coding is featured by the rate-distortion optimization carried out in the rate control process of JPEG2000 core coding system. The first such approach was the Implicit method [4, Chapter16.2], which is reviewed in the following section. A modification of the Implicit, aimed to reduce the background priority, is presented in [9], and [10] introduces a differentiation of codeblocks containing ROI and background coefficients to better distinguish the ROI. Also using rate-distortion optimization techniques, methods to automatically determine ROIs have been recently proposed in [11], [12].

Main advantages of ROI coding methods modifying wavelet coefficients are the fine-grain accuracy for the selected ROI area, and the rich combination of the ROI with the background. See, for example, in Figure 1(a) the excellent accuracy achieved by MaxShift for the ROI area: the recovered area is almost the same as that specified by

the user. The main drawback of such methods, however, is that they significantly penalize the coding efficiency, since more bitplanes have to be encoded; for instance, for a ROI covering 25% of the image, [13] reports an increase of an 8% in the length of the final codestream. Methods employing rate-distortion techniques address this drawback but, on the other hand, they do not achieve a fine-grain accuracy due to the rough discrimination between codeblocks belonging to the ROI, and codeblocks belonging to the background. This coarse differentiation is depicted, for example, in Figure 1(b), where the ROI and background are jointly recovered.

The main purpose of this research is to enhance the fine-grain accuracy of ROI methods that employ rate-distortion optimization techniques, yielding an accuracy comparable to that provided by methods modifying wavelet coefficients. We propose two strategies that do not penalize the coding efficiency, and can be easily implemented in any coder, maintaining JPEG2000 compliance.

The paper is structured as follows: Section II introduces our strategies of ROI coding; Section III assesses the performance of the proposed methods through extensive experimental results; and Section IV concludes this work pointing out some remarks.

II. PROPOSED ROI CODING METHODS

A. Main insights

Commonly, JPEG2000 coders employ rate-distortion optimization techniques to select those bitstream segments that minimize the overall distortion while attaining a target bitrate for the final codestream. The Post Compression Rate-Distortion optimization (PCRD) [5] is the most popular method to conduct this optimization process. PCRD uses the bitrate and the distortion of every truncation point of codeblocks' bitstreams to pose the optimization problem, which may be solved through a generalized Lagrange multiplier for a discrete set of points.

Let n_j denote the potential truncation points of the bitstream produced for codeblock $\mathcal{B}_i$, and let $R_i^{n_j}$ and $D_i^{n_j}$ respectively denote the bitrate and distortion of these points, with $R_i^{n_j} \leq R_i^{n_{j+1}}$. 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 $\triangle R_i^{n_j} = R_i^{n_j} - R_i^{n_{j-1}}$ identifying those truncation points with strictly decreasing rate-distortion slope or, in other words, those points lying on the convex hull. When the Mean Squared Error (MSE) is the considered distortion metric, the distortion of each coding segment is determined as $D_i^{n_j} = G_{b_i} \sum_{k \in \mathcal{B}_i}(y[k] - \hat{y}^{n_j}[k])^2$ where $y[k]$ denotes the coefficients of codeblock $\mathcal{B}_i$, $\hat{y}^{n_j}[k]$ denotes the quantized coefficients at truncation point n_j, and G_{b_i} stands for the energy gain factor of subband b_i to which codeblock $\mathcal{B}_i$ belongs. Once the operational rate-distortion curve of each codeblock is identified, PCRD selects those coding segments with the highest rate-distortion slope values until the target bit-rate is attained.

The key-feature of the Implicit ROI coding method is the modification of the distortion to $D_i'^{n_j} = \mathcal{R} * D_i^{n_j}$ (in this case, the ROI priority $\mathcal{R}$ is not restricted to be a power of 2). Since D' is the considered measure by the

Lagrange multiplier, codeblocks containing ROI coefficients are effectively more prioritized than codeblocks containing background coefficients. The unsuitable trait of this approach is that it only distinguishes codeblocks as either belonging to the ROI or to the background, considering codeblocks as ROI codeblocks even when they contain just a single ROI coefficient, thus prioritizing them with the same priority as codeblocks containing exclusively ROI coefficients. It is not a simple task to avoid this malfunction, since the ROI defined in the image domain has to be inversely traced in the wavelet domain, which produces an expanding effect due to the filter length and decomposition levels of the wavelet, eventually emphasizing the malfunction, especially at low resolutions levels (see Figure 1(c)).

Aimed to improve the fine-grain accuracy, we propose a modification on the determination of the distortion contributions of codeblocks. The underlying idea is to consider the number and/or the magnitude of the coefficients within the codeblock that belong to the ROI. Being $\mathcal{R}$ the ROI priority as in the Implicit method, two approaches are feasible:

- Subblock ROI coding method. The actual distortion of each coefficient is computed, *instead* of considering the global codeblock distortion, according to

$$D_i''^{n_j} = G_{b_i} \sum_{k \in \mathcal{B}_i} \begin{cases} ((y[k] - \hat{y}^{n_j}[k])^2 * \mathcal{R}) & \text{if } y[k] \in ROI \\ (y[k] - \hat{y}^{n_j}[k])^2 & \text{otherwise} \end{cases} \tag{1}$$

- Weighted ROI coding method. The distortion is computed taking into account the percentage of coefficients of codeblock $\mathcal{B}_i$ that are related to the ROI (here named ROI Weighting Factor, RWF_i), according to

$$D_i'''^{n_j} = \mathcal{R} * D_i^{n_j} * RWF_i . \tag{2}$$

Subblock ROI coding method modifies the distortion computation at a coefficient level, while the Weighted ROI Coding method modifies the distortion computation at a codeblock level. The second approach uses a technique similar to the Window Scaling Factor in [3], however, note that in the Weighted ROI coding method, the percentage of ROI coefficients multiplies the distortion of each coding pass instead of modifying the rate-distortion slope of coding segments.

B. Other considerations

As most of the ROI coding methods presented in the literature, the proposed Subblock and Weighted methods are able to prioritize more than one ROI at different priorities. Even though the ROI may, or may not, be overlapped in the image domain – in which the user defines ROIs –, the expanding effect when tracing the inverse wavelet transform can produce overlapping of the ROI areas in the wavelet domain. In the experiments carried out in the next section, we consider that when one wavelet coefficient is identified as belonging to more than one ROI, it is prioritized with the highest priority of the ROIs it belongs to.

bps	ROI AREA = 5%				ROI AREA = 15%				ROI AREA = 30%			
	Sub.	Wei.	Imp.	Opt.	Sub.	Wei.	Imp.	Opt.	Sub.	Wei.	Imp.	Opt.
0.0025	22.14	-0.23	-0.99	+1.17	20.61	-0.17	-0.64	+0.64	19.26	-0.11	-0.34	+0.36
0.005	23.75	-0.33	-1.19	+1.13	21.85	-0.17	-0.72	+0.62	20.31	-0.08	-0.30	+0.37
0.01	25.62	-0.22	-1.35	+1.26	23.32	-0.15	-0.77	+0.62	21.55	-0.08	-0.30	+0.34
0.1	36.41	-0.17	-1.43	+1.82	31.71	-0.11	-0.83	+0.80	28.44	-0.07	-0.45	+0.38
1	54.08	0.00	+0.14	+0.28	52.18	-0.04	-0.98	+0.84	45.11	-0.06	-0.91	+1.10
2	55.23	0.00	+0.06	+0.08	55.05	0.00	-0.01	+0.06	54.30	-0.01	-0.61	+0.35

TABLE I

AVERAGE RESULTS FOR THE ISO 12640-1 CORPUS WHEN DIFFERENT ROI PRIORITIES ARE USED, $\mathcal{R}$ VARYING FROM 3 TO 8 ($\mathcal{R}$ IS CHOSEN EQUALLY FOR ALL METHODS).

(a) ROI

(b) background
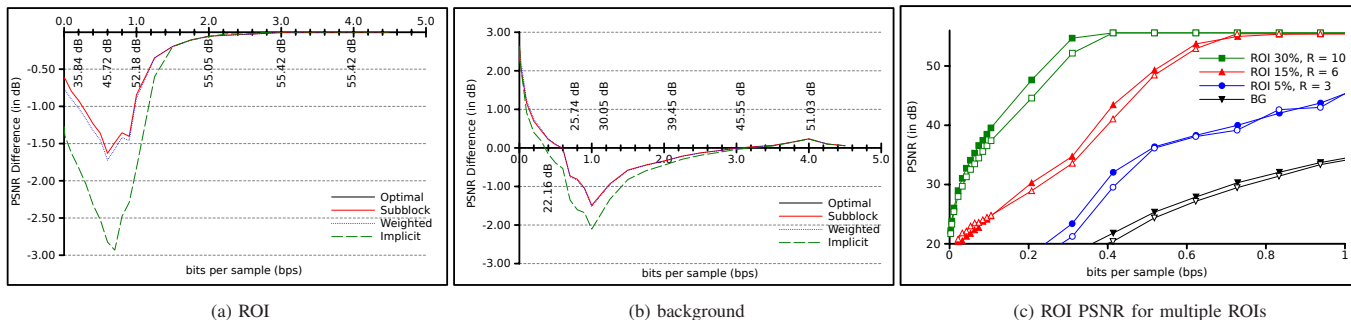
(c) ROI PSNR for multiple ROIs

Fig. 2. Average coding performance achieved for the eight images of the corpus, with a different ROI per each image delineating a relevant region (the ROIs cover 15% of the image area, and $\mathcal{R}$ varies from 3 to 8). (a) PSNR differences of the ROI. (b) PSNR differences of the background. (c) PSNR differences of the ROI for multiple ROIs.

Regarding the computational complexity, it is worth noting that the proposed methods only increase the computational load of the encoder negligibly, especially when compared to methods that modify wavelet coefficients. As reported in the literature [14], the PCRD process represents less than 5% of the computational load of a JPEG2000 encoder, compared to the 65% spent by the coding engine. Recall that methods in the first ROI coding mechanism raise the number of bitplanes to encode, whereas Subblock and Weighted methods only modify slightly the distortion computation that, in addition, might be implemented as a bit-shift operation if $\mathcal{R}$ is a power of 2.

III. EXPERIMENTAL RESULTS

The performance of the two proposed ROI coding methods is assessed comparing Subblock and Weighted to the plain Implicit method and to a theoretical optimal performance. The theoretical optimal performance is computed "up-shifting" the quantized wavelet coefficients belonging to the ROI and encoding the image at the specified target bitrate, in a similar way as the Scaling method does, but allowing arbitrary ROIs and not considering the bitrate needed to transmit the ROI mask. This is not a practical approach but gives us the theoretical optimal performance that may be achieved with methods modifying wavelet coefficients (first ROI coding mechanism). Regarding the second ROI coding mechanism, we only report comparisons to the Implicit method, since all methods using such mechanism differentiate ROI and background codeblocks similarly.

All methods have been implemented in our JPEG2000 Part-1 implementation BOI [15]. Coding parameters are set to: 5 levels of 9/7 DWT, derived quantization, codeblock size of 32×32, and restart coding variation. Although the use of codeblock sizes of 64×64 enlarges the difference between the performance of Subblock and Weighted with respect to Implicit, all reported experiments have been carried

out considering a codeblock size of 32×32 to minimize the penalization of the Implicit method when distinguishing ROI codeblocks from background codeblocks [4, Chapter 16.2]. In the evaluations, the images have been encoded at different target bitrates, decoded, and the Peak Signal to Noise Ratio (PSNR) has been computed separately for the ROI and the background.

Table I reports the average results achieved for all eight images of the ISO 12640-1 corpus when different ROIs are encoded at different priorities. ROIs have been defined in the image domain covering 5%, 15%, and 30% of the image area. The *Sub* column in this table reports the ROI PSNR achieved by Subblock, while remaining columns report the difference between the Subblock with respect to Weighted, Implicit, and Optimal. Experimental results suggest that the Subblock and Weighted methods improve the Implicit's coding performance for the ROI area in all cases. The maximum gain between our methods and the Implicit is produced for the smallest ROI definition, with a peak of 1.43 dB. Although the Subblock and Weighted methods recover the ROI with a higher fine-grain accuracy than the Implicit, this difference is lower when the ROI is larger, since larger ROIs do not penalize so much the codeblocks' differentiation.

Figures 2(a) and 2(b) depict, respectively, the ROI and background PSNR difference between the Weighted, Implicit, and Subblock with respect to the optimal performance. For all bitrates, Subblock and Weighted methods outperform the Implicit, both for the ROI and also for the background: PCRD allocates ROI codestream segments to the final codestream –until the target bitrate is reached– while their slopes are greater than the background slopes; since Subblock and Weighted ROI coding methods discriminate better the ROI codeblocks than the Implicit method, they are able to transmit all ROI codeblocks earlier than the Implicit, leaving room for sending also some background codeblocks.
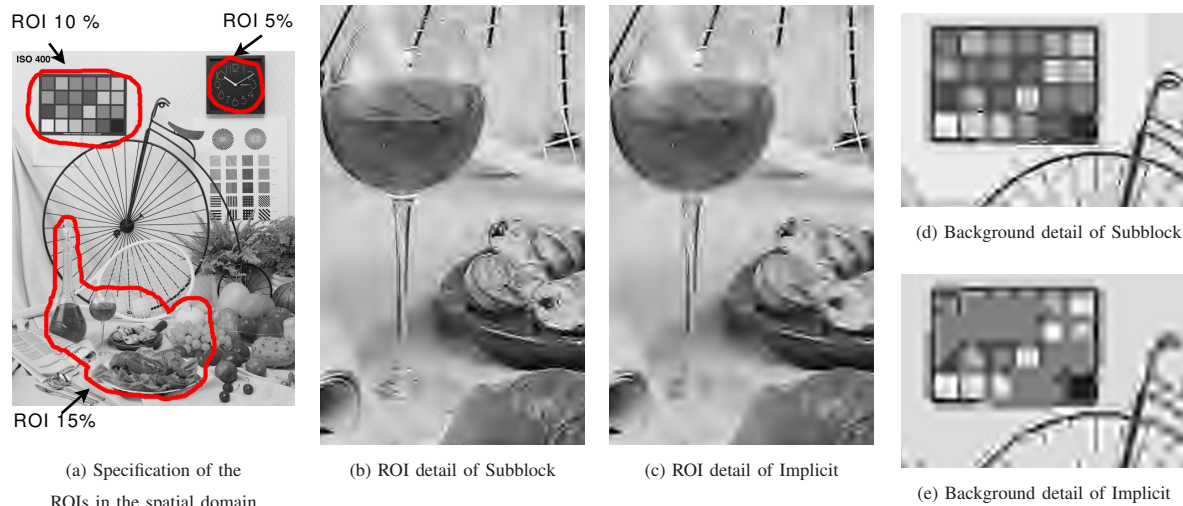
(a) Specification of the ROIs in the spatial domain

(b) ROI detail of Subblock

(c) ROI detail of Implicit

(d) Background detail of Subblock

(e) Background detail of Implicit

Fig. 3. Visual comparison among the Subblock and the Implicit methods. Bicycle image (size 2048×2560, gray-scaled) with the ROI covering 15% of the image area, encoded at a target bitrate of 0.015 bps. The ROI priority is $\mathcal{R} = 8$ for all methods.

The performance difference between Subblock and Weighted is almost equivalent for the ROI and the background, being less than 0.2 dB for all bitrates. Both Subblock and Weighted methods achieve a competitive performance compared to the theoretical optimal one. When the codeblock size is 64×64, the gains of the Subblock with respect to the Implicit method are commonly extended to more than 1 dB compared to results achieved when the codeblock size is 32×32 (ROIs covering 15% of the image).

To compare the coding performance achieved when multiple ROIs are defined, we have encoded all images of the corpus when three ROIs are encoded at different priorities (an example is provided in Figure 3(a)). Figure 2(c) depicts the results achieved by Subblock (solid dots) and by the Implicit (hollow dots). For most bitrates and ROIs, Subblock outperforms the performance of the Implicit method.

Figure 3 depicts a visual comparison for the Bicycle image when the ROI covering 15% of the image area depicted in Figure 3(a) is encoded by Subblock and Implicit at a compression factor of 533:1. The Subblock method enhances the visual quality for both the ROI and the background areas with respect to the Implicit method.

IV. CONCLUSIONS

There exist two mechanisms for the ROI coding in the JPEG2000 standard: either modifying wavelet coefficients, or using rate-distortion optimization techniques. Methods employing the first mechanism achieve a much better fine-grain ROI accuracy, but they fail in coding and computational efficiency. Methods employing the second mechanism achieve a poorer fine-grain ROI accuracy, but a much better coding and computational efficiency. This paper introduces two rate-distortion based ROI coding methods that, without penalizing the coding efficiency, largely improve the fine-grain accuracy, up to that obtained by first mechanism methods. Experimental results suggest that to obtain a good fine-grain accuracy, it is sufficient to know only the number of coefficients belonging to the ROI, without considering their magnitude. In addition our proposals maintain JPEG2000 compliance.

REFERENCES

[1] Z. Li-bao, "New approach to JPEG2000 region of interest image coding and its applications for remote sensing image," in *Proc of SPIE, Geospatial Information, Data Mining, and Applications.*, vol. 6045, Dec. 2005, p. 7.
[2] M. Penedo, W. Pearlman, P. Tahoces, M. Souto, and J. Vidal, "Embedded wavelet region-based coding methods applied to digital mammography," in *Proc. of IEEE International Conference on Image Processing*, vol. 3, Sept. 2003, pp. 14–17.
[3] D. Taubman and R. Rosenbaum, "Rate-distortion optimized interactive browsing of JPEG2000 images," in *Proc. of IEEE International Conference on Image Processing*, vol. 2, Sept. 2003, pp. 765–768.
[4] D. Taubman and M. W. Marcellin, *JPEG2000 Image compression fundamentals, standards and practice*. Norwell, Massachusetts 02061 USA: Kluwer Academic Publishers, 2002.
[5] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Trans. Image Processing*, vol. 9, no. 7, pp. 1158–1170, July 2000.
[6] Z. Wang and A. C. Bovik, "Bitplane-by-bitplane shift (BbBShift)- a suggestion for JPEG2000 region of interest image coding," *IEEE Signal Processing Lett.*, vol. 9, no. 5, pp. 160–162, May 2002.
[7] L. Liu and G. Fan, "A new JPEG2000 region-of-interest image coding method: Partial significant bitplanes shift," *IEEE Signal Processing Lett.*, vol. 10, no. 2, pp. 35–38, Feb. 2003.
[8] Z. Li-Bao and W. Ke, "New approach for JPEG2000 region of interest image coding hybrid bitplane shift," in *Proc. of IEEE International Conference on Machine Learning and Cybernetics*, vol. 6, Aug. 2004, pp. 3955–3960.
[9] H. Yang, M. Long, and H.-M. Tai, "Region-of-interest image coding based on EBCOT," *IEE Proc.-Vis. Image Signal Processing*, vol. 152, no. 5, pp. 590–596, Oct. 2005.
[10] Y. Xie and G.-Q. Han, "ROI coding with separated code block," in *Proc. of IEEE International Conference on Machine Learning and Cybernetics*, vol. 8, Aug. 2005, pp. 5447–5451.
[11] B. Liu, M. Sun, Q. Liu, A. Kassam, C.-C. Li, and R. Sclabassi, "Automatic detection of region of interest based on object tracking in neurosurgical video," in *Proc. of IEEE International Conference on Engineering in Medicine and Biology*, Sept. 2005, pp. 6273–6276.
[12] O. T.-C. Chen and C. Chih-Chang, "Automatically-Determined Region of Interest in JPEG 2000," *IEEE Trans. Multimedia*, vol. 9, no. 7, pp. 1333–1345, Nov. 2007.
[13] C. Christopoulos, J. Askelof, and M. Larsson, "Efficient methods for encoding regions of interest in the upcoming JPEG2000 still image coding," *IEEE Signal Processing Lett.*, vol. 7, no. 9, pp. 247–249, Sept. 2000.
[14] D. Taubman, "Software architectures for JPEG2000," in *Proc. IEEE International Conference on Digital Signal Processing*, vol. 1, July 2002, pp. 197–200.
[15] Group on Interactive Coding of Images, "BOI software," July 2007. [Online]. Available: http://www.gici.uab.cat/BOI