Quality constraint and rate-distortion optimization for predictive image coders

Khouloud Samrouth, François Pasteau, Olivier Deforges

Quality constraint and rate-distortion optimization for predictive image coders

Khouloud Samrouth*[a], François Pasteau[a], Olivier Deforges[a]
[a]IETR Laboratory, UMR CNRS 6164, Rennes France, 00 33 6 23 23 82 86

ABSTRACT

Next generations of image and video coding methods should of course be efficient in terms of compression, but also propose advanced functionalities. Among these functionalities such as scalability, lossy and lossless coding, data protection, Rate Distortion Optimization (RDO) and Rate Control (RC) are key issues. RDO aims at optimizing compression performances, while RC mechanism enables to exactly compress at a given rate. A less common functionality than RC, but certainly more helpful, is Quality Control (QC): the constraint is here given by the quality. In this paper, we introduce a joint solution for RDO and QC applied to a still image codec called Locally Adaptive Resolution (LAR), providing scalability both in resolution and SNR and based on a multi-resolution structure. The technique does not require any additional encoding pass. It relies on a modeling and estimation of the prediction errors obtained in an early work. First, quality constraint is applied and propagated through the whole resolution levels called pyramid. Then, the quantization parameters are deduced considering inter and intra pyramid level relationships. Results show that performances of the proposed method are very close to an exhaustive search solution.

Keywords: Predictive Coder, quantization, quality allocation, rate-distortion optimization

1. INTRODUCTION

The amount of data exchanges over fix and mobile networks is still exponentially increasing. Nowadays, images and videos compose most of these data, and the need for efficient compression techniques is still an important issue. Besides this traditional need for image/video coding systems, new services and applications, in particular for Internet context, suggest new capacities and functionalities. For instance, the diffusion of media over various types of networks in terms of bandwidth would require the use of scalable encoding solutions. The exchanges of images and videos over more or less secured social networks demand to more and more consider problems of privacy and data security by enabling encryption, watermarking and data hiding techniques. Additional requests can concern some specific domains such as medical or art, in which users can prefer reversible (lossless) coding methods, even at the expense of a significant loss of compression efficiency. In this case, a unique solution for both lossy and lossless coding would be suitable. Therefore, it is surprising to notice that JPEG format is still the commonly used compression technique as this method provides restricted functionality. Recently, the JPEG committee has established a Call For Proposal (CFP) in order to promote new technologies more adapted to new requirements. Nevertheless, no solution has been accepted at the end of the experiment process.

The Locally Adaptive Resolution (LAR) coding method has been proposed as a candidate in the context of the CFP. LAR is a global coding framework providing a lot of functionalities such as lossy/lossless compression, resolution and quality scalability, partial/full encryption, Region of Interest coding. Details on the method will be given in the following section. LAR has a multi-layer pyramidal structure. This multi-layer aspect can be an advantage in order to adjust the behavior of the coder according to some constraints such as suitable bit-rates, classes of images (natural or medical). On the other hand, a complex multi-layer structure generally increases the difficulty to well configure the coder as the number of parameters can be significant. It was indeed the main limitation of the LAR for the core experiments: the difficulty to find the optimal / near optimal parameters. The second problem arises with the poor possibility of the method to fix the rate. Therefore, the first problem suggested designing a Rate Distortion Optimization (RDO) strategy, and the second one Rate Control (RC) functionality. This paper presents both aspects, or more precisely a RDO and Quality Control (CQ) both-together.

RDO techniques are generally used in order to automatically find the best configuration of the coder and able to minimize the distortion for a given bit-rate. One well-known solution of RDO technique is the Lagrangien based one

*Khouloud.samrouth1@insa-rennes.fr phone: 00 33 6 37 56 26 70

Image Processing: Algorithms and Systems XI, edited by Karen O. Egiazarian, Sos S. Agaian, Atanas P. Gotchev,
Proc. of SPIE- IS&T Electronic Imaging, SPIE Vol. 8655, 86550D · © 2013 · SPIE-IS&T
CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2001655

embedded in H.264 during the motion estimation/compensation process [1]. RDO techniques have to be dependent of the coding technique. One other example that can be somewhat considered as RDO, is the estimation of optimal quantization tables for the DCT coefficient in JPEG.

The objective of RC is to enable more or less precisely encoding at a given bit-rate. RC techniques are also coding technique dependent, with various original possibility of the coder. For example, Jpeg2000 provides an embedded stream enabling a fine RC[1], [2]. Recently, Rho-domain has been introduced as an efficient RC technique for H.264 [3], but that cannot be extrapolated to different coding approaches. More generally, a coding technique providing scalability can support some RC by “cutting” sub-streams after the given bit-rate, but it is not certain that these cuts give the optimal choice.

The general problem of fixing the bit-rate can still be an issue for video compression domain, but for still image one, there is no more applications for which it is necessary to precisely respect this type of constraint: use-cases for images are more concerned by the image quality. The capacity to encode at a given quality constraint can be called Quality Control (QC). State-of-the-art proposes very few contributions in this domain.

The paper is organized as follows. Section 2 introduces the LAR modeling based on results in [4]. In section 3, we describe the quality allocation scheme and the rate-distortion optimization technique derived from the LAR modeling. In section 4, we illustrate our experimental results and discuss the performance of the proposed rate-distortion optimization technique. We conclude our work in section 5.

2. LAR MODELING

Locally Adaptive Resolution (LAR) is an efficient content-based image coder offering advanced scalability at different semantic levels. A local analysis of image activity leads to a non-uniform block representation supporting two layers of image description. The first layer provides global information enabling low bit rate while preserving contours. The second layer holds local texture information. Self-extraction of regions of interest is also possible [4]. Furthermore, subjective quality image evaluation shows that the low resolution LAR picture is better than the same bit-rate JPEG2000 coded picture [5].

LAR coder has 3 profiles so that each profile corresponds to different functionalities and different complexities: baseline profile, pyramidal profile and extended profile. Our work concerns the pyramidal profile, a multi-resolution coding scheme, figure 1 [6]. It consists of a dyadic pyramidal decomposition on L levels of the original image, followed by pixel value predictions between levels using Wu prediction algorithm [7]. The error between the predicted pixel value and the real pixel value is the prediction error. This prediction error is the information that has to be quantized and then coded by an entropy coder.

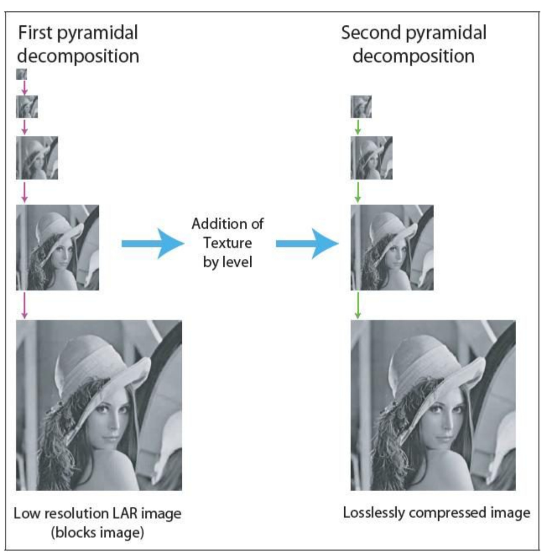


Figure 1: Block diagram of LAR Pyramidal profile

To perform quality allocation and rate-distortion optimization, a statistical study of the predictive codec is needed. This study, done in an early work [2], should conduct to understand the behavior of the predictive coder and provide mathematical tools which would enable us to simulate the coder behavior. Therefore it enables a better understanding of the relationships of the different parameters with the obtained quality and rate. Being a multi-resolution framework, LAR has different levels of quality. Therefore, it is difficult to effectively set the quantization parameters between different resolution levels [4].

2.1 Characterization of prediction errors

Being based on a predictive scheme, the distribution model of the prediction errors of LAR codec may follow the Laplacian model [8], [9]. To find the best fitted probability distribution model, practical probability distributions of prediction errors have been observed in a previous work [4]. From these observations, two empirical conclusions have been drawn and will serve as hypothesis for the rest of the study. First, as shown in figure 2, a Laplace law function can represent a well fitted probability distribution model. Secondly, the mean value of the practical probability distribution has been empirically determined to be 0. Such hypothesis can be insured by adaptively removing the bias in case of a shifted distribution.

Laplace law can be expressed by (1), where $\tilde{P}_Q(x)$ represents the probability of the prediction error x, g_0 is the number of errors equal to zero produced by the prediction process and b is function of g_0 [4]. Q is the quantization parameter.

$$\tilde{P}_Q(x) = \alpha_Q e^{-\frac{|Qx|}{b}} \tag{1}$$

Where
$$\alpha_Q = \frac{1-\mathrm{e}^{-\frac{Q}{b}}}{1+\mathrm{e}^{-\frac{Q}{b}}}, \; b = \frac{1}{2g_0}.$$

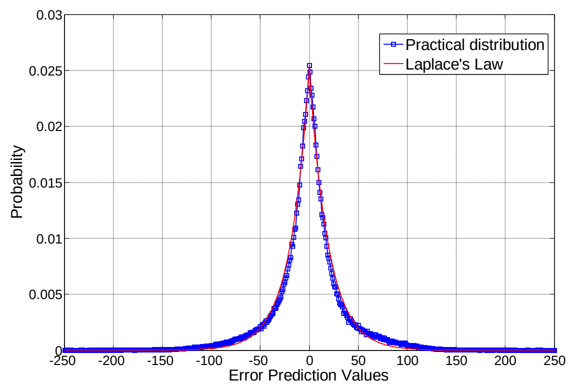


Figure 2: Comparison between practical error prediction distribution and estimated distribution (Laplace Law) [4]

Therefore, the behavior of the information generated at the output of the LAR encoder (quantized error prediction) can be modeled by a Laplace law that is characterized by only one parameter *b*. This parameter can be determined online during the prediction process by monitoring the number of errors equal to zero produced by the prediction process.

2.2 Entropy Estimation

The entropy gives the minimum amount of information required by a source of information for a lossless coding. In our previous statistical study, entropy is expressed in bits per pixel of coded image. So the entropy evaluates the rate at the encoder output. From the probability distribution model, entropy can be computed [4]. This leads us to obtain an entropy estimator *H* with a simple expression defined as follows:

$$\mathrm{H} = \log_2 (2.\mathrm{e}.\frac{b}{Q}). \tag{2}$$

where *Q* is the quantization factor and *b* is the Laplace law parameter dependent on the image. We tested this expression on six 1280*1600 test images including “Lena”, “Bike”, “Green”, “Baboon”, “Pepper” and “Woman”, using different

values of quantization factor [4]. Figure 4 presents an example for entropy estimation from equation 2 as well as entropy experimentally found for bike image.
Figure 4 shows that the entropy estimator works well until we reach large quantization factor, where it estimates negative entropy values.

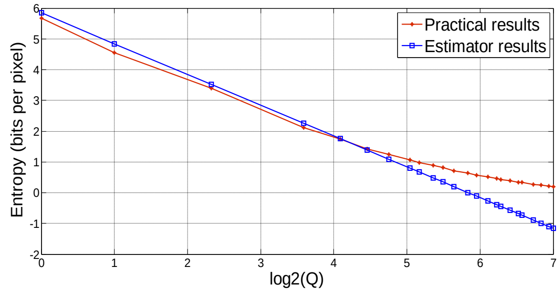


Figure 3: Comparison between practical and estimated values of Entropy

The reason is that this model consists of a linear relationship between entropy and log-based 2 of the quantization; so for high value of Q, we have $\frac{b}{Q} \approx 0$, therefore $\log_2(2.e.\frac{b}{Q})$ becomes negative. But as the entropy tends to zero for large quantization factor, a very simple solution is to reset the entropy once the estimator gives a negative value.
As we could estimate the rate from the quantization factor, we should see how we can estimate the distortion level of the reconstructed image from the quantization factor.

2.3 Quality Estimation

In this part we search to estimate the distortion created by the quantization process. The level of distortion can be expressed mathematically in mean square error MSE. So, in our previous work, we estimated the MSE from the probability distribution model and the quantization factor Q using expression 3. Details of MSE equation can be found in [4].

$$\mathrm{MSE} \approx 4\alpha_1 b^3 + \frac{4\alpha_1 e^{-\frac{\frac{\lfloor Q \rfloor}{2}}{b}}}{1-e^{-\frac{Q}{b}}} \left(\begin{array}{l} \left(e^{-\frac{1}{b}}-1\right)b^3 - \left(\frac{\lfloor Q-1 \rfloor}{2} e^{-\frac{1}{b}} + \frac{\lfloor Q \rfloor}{2}\right) b^2 \\ + \frac{1}{2}\left(\left\lfloor \frac{Q-1}{2} \right\rfloor^2 e^{-\frac{1}{b}} - \left\lfloor \frac{Q}{2} \right\rfloor^2\right) b \end{array} \right) \quad (3)$$

where Q is the quantization factor, b is a parameter dependent on the image and α1 is αQ for Q =1. Figure 4 presents MSE estimation from equation 3 as well as MSE experimentally found for bike-crop image with different quantization steps.
Tests of this expression on typical color images using different values of quantization factor shows that the MSE estimator works well. Figure 5 show the comparison between practical and estimated values of MSE of woman image.
More details can be found in [4].

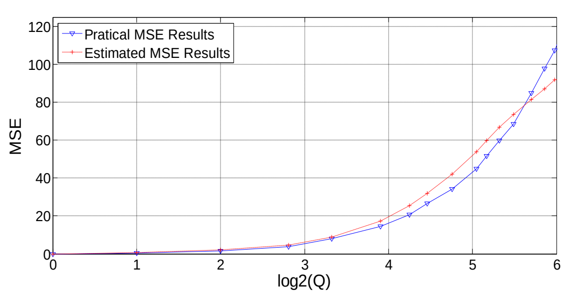


Figure 4: Comparison between practical and estimated values of MSE

Now we have the mathematical tools to estimate the rate at the encoder output, and the distortion level of the reconstructed image, deduced from our previous work, in next section we will build our proposed joint quality allocation/rate-distortion optimization scheme and we will show the interest of LAR modeling based approach.

3. QUALITY CONSTRAINT & RATE-DISTORTION OPTIMIZATION

Joint quality constraint/rate-distortion optimization techniques are generally used in order to find the best configuration of the coder and able to minimize the rate for a given quality constraint. For a distortion level set D^{set} expressed in Mean Square Error (MSE) of the reconstructed image, our rate-distortion optimization technique consists of finding the quantization parameters Q minimizing the rate R in bit per point of the coded image. The optimization problem can be formulated as follows:

$$Q = \arg\min R(Q),$$
$$Q = \{ Q_{Li} \},\ Li=0, 1, \ldots, L_{max}.$$

Under the constraint that $$D(Q) \leq D^{set},$$

where D is the distortion of the reconstructed image.

Our joint quality constraint/rate-distortion optimization scheme contains two modules. First, as we apply our technique on LAR pyramidal profile, a MSE allocation policy is used to distribute the distortion set between pyramid levels. Then, reaching the optimized rate can be accomplished by acting on the quantization factor Q. So it is turned from a global optimization problem into a local one. Therefore, the goal of this technique is to determine the optimal quantization parameters Q_{Li} for coding each level (Li) of the pyramid, such that the output rate is minimized while the reconstructed image distortion set is respected.

3.1 Distortion Allocation

In our scheme, the distortion level set is distributed between the L levels of LAR pyramid. Therefore, a distortion level set D_{Li}^{set} is specific for each level of the pyramid:

$$D_{Li}^{set} = Factor[Li] * D^{set}$$

The optimal allocation is shown in table 1. This allocation is obtained after many tests on a large set of 1280*1600 color images.

Table 1. Optimal repartition of Distortion set between L levels of LAR pyramid

Level	5	4	3	2	1	0
Factor	0	0.028	0.046	0.23	0.67	1.6

3.2 Q computation

Once the distortion allocation between levels is performed, the optimal value of Q_{Li} is computed for each level of the pyramid.

$$Q_{Li} = \arg\min R(Q_{Li}),$$

Under the constraint that

$$D_{Li}(Q_{Li}) \leq D_{Li}^{set}$$

$$Q = \{ Q_{Li} \},\ Li=0, 1, \ldots L_{max}$$

It consists of choosing the quantization parameters of all levels at once from the top level of the pyramid. After coding the top level without quantification (Q_{Lmax} = 1), we have the actual number of errors equal to zero produced by the prediction process α_{Lmax}. Using the estimator tool *alphaprediction* [4], we estimate alpha of all sub levels $\tilde{\alpha}_{Li}$.

At a given level *Li*, it is possible to estimate the D_{Li}^{set} from α_{Li} and the quantization parameter Q_{Li}, using expression (3). So in our technique we just have to do the opposite approach. Indeed, after calculating the distortion *D*, in terms of MSE, from a few values of Q_{Li}, we select the quantization parameter $\hat{Q}_{Li}$ that produces the closest *D* to the D_{Li}^{set} . We call this process *Qcomputing*. We redo the same procedure of choice for the quantization factors for all lower levels of the pyramid. Once all the quantization factors are determined, we code all different levels.

Algorithm: Quantizatrion parameters determination
Input: α_{Lmax}, D_{Li}^{set}
Output: Q_{Li}
1. **if** $i = L_{max}$ **then**
2.   **for all** i **do**
3. $\tilde{\alpha}_{Li}$ = *alphaprediction* (α_{Lmax}, *i*)
4. $\hat{Q}_{Li}$ = *Qcomputing* (D_{Li}^{set} , $\tilde{\alpha}_{Li}$)
5.   {return the quantization parameter}
6.   return $\hat{Q}_{Li}$
7.   end for
8. end if

These quality allocation scheme and rate-distortion optimization techniques were designed to be as simple as possible. It can still be improved by including some more advanced features such as a distortion repartition update after each level coding.

4. EXPERIMENTAL RESULTS

We tested the rate-distortion allocation scheme proposed in section 3 on multiple test images mentioned above and for different PSNR set. To evaluate the performance of the scheme, we first compare the quality set to the practical quality obtained after decoding. Then we see if the encoder reached the minimum rate. After applying rate-distortion allocation technique to bike-crop image, quality results are shown in table 2.

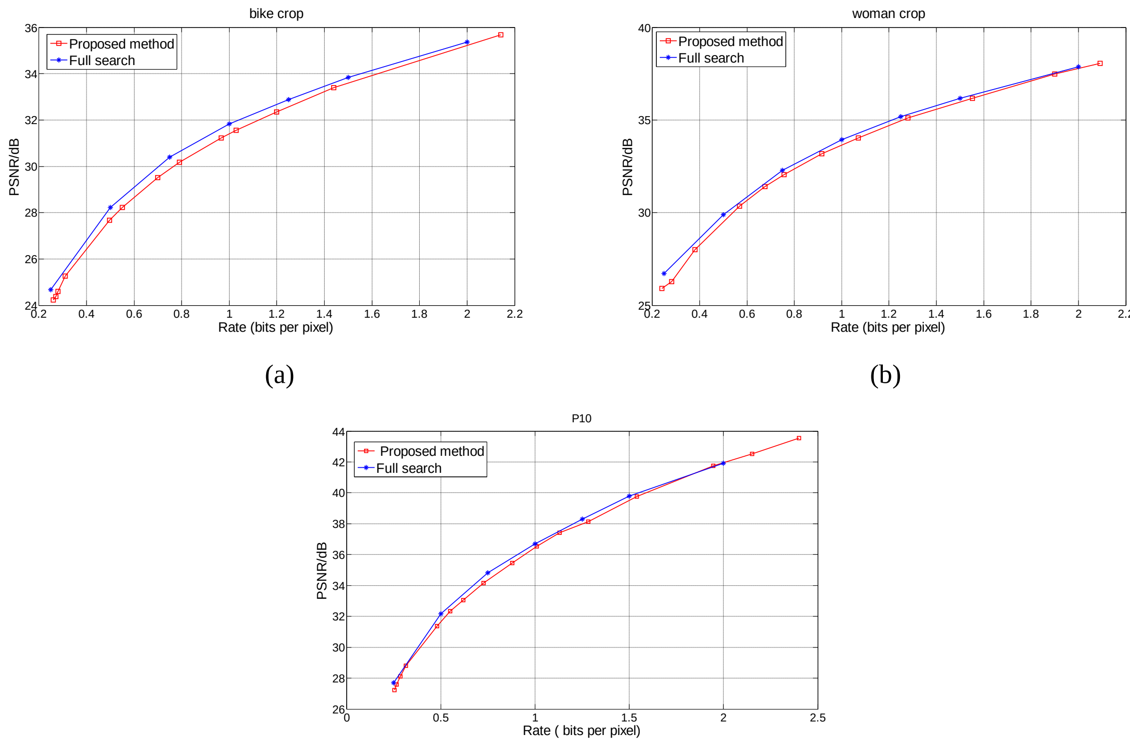
Table 2. MSE set results

MSE set	***PSNR set dB***	***MSE obtained***	***PSNR obtained dB***
2.05	**45**	3.26	**43**
6.50	**40**	8.89	**38.63**
20.56	**35**	21.19	**34.86**
65.02	**30**	96.16	**28.30**

Table 2 shows that, after coding with our technique, the quality of reconstructed images are practically very close to the quality set. This validates the global functioning of the algorithm.

By applying our quality constraint scheme and rate-distortion optimization technique, calculating the coded images rate and evaluating the distortion level of reconstructed images, we obtain the rate-distortion curves below (Figure 5). On each chart we have two curves. On the first one, we find the ideal quality (in terms of PSNR) of reconstructed image for a certain rate of coded image. This curve is obtained after doing an exhaustive search to determine the optimal quantization step at each level of LAR pyramid. The second curve represents the results obtained after applying the proposed rate-distortion technique to the LAR. For example, by applying the LAR coder with the rate-distortion allocation technique to bike-crop image, we obtain at a rate of 0.75 bpp a $\mathrm{PSNR} \approx 29.9\ \mathrm{dB}$, which is very close to the ideal $\mathrm{PSNR} \approx 30\ \mathrm{dB}$ (obtained after an exhaustive search).

These curves show the satisfactory behavior of the proposed algorithm. Indeed, the curves show that the gap between the results achieved and the ideal behavior of the LAR, found by an exhaustive search, is between 0dB and 0.3dB. Hence the great interest of our proposed technique. It allows you to follow a set amount of distortion while optimizing the flow generated without resorting to an exhaustive search of quantization factor.

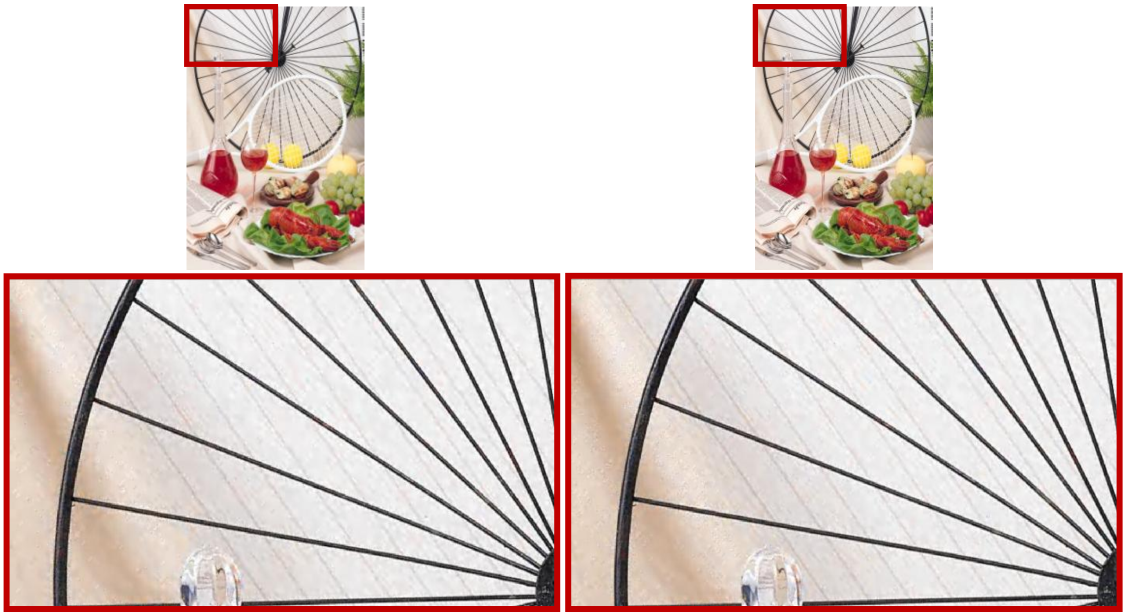


(a) (b)

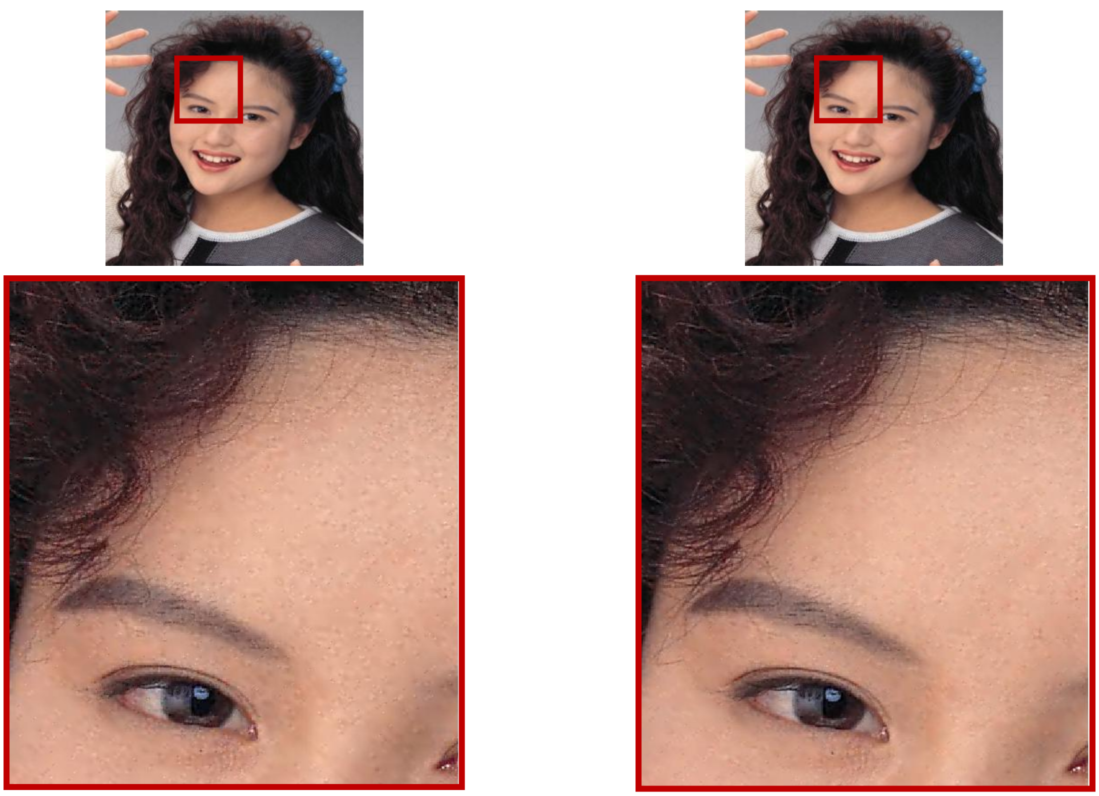
(c)

Figure 5: Rate-distortion curve for (a) Bike image, (b) Woman image, (c) P10 image

Figures 6 shows that even for the same rate, our proposed rate/distortion optimization technique leads to a better visual quality.
The aim of our work was to add a new internal functionality to LAR codec, so we only compared our results to the LAR classic results. Comparison with the state of the art was studied in [6].



(a)



(b)

Figure 6: Visual quality enhancement for (a)Bike image: 0.75 bpp coded image; (b) Woman image: 1bpp coded image. To the left: image coded with LAR without rate-distortion allocation technique. To the right: image coded with LAR and with proposed rate-distortion allocation technique.

5. CONCLUSION

In this paper, we have presented a quality allocation and a rate-distortion optimization technique for predictive coders. Statistical analysis were performed in order to characterize the prediction errors at the output of the predictive coder by Laplace law, and to find tools for estimating rate of coded images and quality of reconstructed images. These tools have been useful for computing optimized quantization parameters without resorting to an exhaustive search and by online parameters determination. Experimental results show that the proposed algorithms can achieve almost the same performance of a full-search parameter optimization.

The technique we proposed for the quality allocation and rate-distortion optimization is of great interest. The encoder complies with a set level of distortion while optimizing the generated rate, instead of an exhaustive search to reach it.

Furthermore, it is possible to imagine some modifications in distortion allocation scheme. Indeed, it is possible to update the MSE set of level Li with the error on the MSE practically obtained in level Li+1, in that way we compensate the error occurred on the MSE of higher levels.

This is a new feature added to the LAR. Furthermore, it could be extended to many other coders based on predictive schemes such as HEVC.

6. ACKNOWLEGMENT

This work is done as part of 4EVER, a French national project with support from Europe (FEDER), French Ministry of Industry, French Regions of Brittany, Ile-de-France and Provence-Alpes-Cote-d’Azur, Competitivity clusters “Images&Reseaux” (Brittany), “Cap Digital” (Ile-de-France) and “Solutions Communicantes Sécurisées” (Provence-Alpes-Cote-d’Azur).

REFERENCES

[1] C.-H. Chou, K.-C. Liu, and P.-H. Chung, “Perceptually Optimized Rate Control for JPEG2000 Coding of Color Images” , in Congress on Image and Signal Processing (CISP), vol. 2, p. 80 -84 (2008).

[2] F. Zhang, Q. Wang, and J. Duan, “Algorithm for JPEG2000 rate control based on number of coding passes”, in International Conference on Multimedia Technology (ICMT), p. 137 -140 (2011).

[3] Y. Lin, Y.-M. Lee, and C.-D. Wu, “Efficient Algorithm for H.264/AVC Intra Frame Video Coding”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 10, p. 1367 -1372 (2010).

[4] F. Pasteau, Statistical Study of a Predictive Codec: a LAR-Based Robust and Flexible Framework, PHD Thesis, available on HAL database, INSA Rennes, (2011).

[5] C. Strauss, F. Pasteau, F. Autrusseau, M. Babel, L. Bedat, and O. Deforges, “Subjective and objective quality evaluation of lar coded art images”, in IEEE International Conference on Multimedia and Expo (ICME), p. 674-677 (2009).

[6] M. Babel and O. Deforges, “Lossless and lossy minimal redundancy pyramidal decomposition for scalable image compression technique”, in Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP), vol. 3, p. 249-52. (2003).

[7] X. Wu and N. Memon, “Context-based, adaptive, lossless image coding”, IEEE Transactions on Communications, vol. 45, no. 4, p. 437-444, (1997).

[8] Y. Kuroki, Y. Ueshige, and T. Ohta, “New statistical models of the JPEG lossless mode subject to the super high definition images”, in Proceedings International Conference on Image Processing,. (ICIP), vol. 1, p. 448-452 (1999).

[9] Y. Kuroki, Y. Ueshige, and T. Ohta, “Redesigning of JPEG statistical model in the lossy mode fitting distribu.tion of DCT coefficients”, in Proceedings International Conference on Image Processing, vol. 3, p. 825-828 (2000).