One Pass Quality Control and Low Complexity RDO in A Quadtree Based Scalable Image Coder

Yi Liu, Olivier Deforges, François Pasteau, Khouloud Samrouth

One Pass Quality Control and Low Complexity RDO in A Quadtree Based Scalable Image Coder

Yi Liu, Olivier Deforges, François Pasteau, Khouloud Samrouth
UMR CNRS 6164 IETR Group Image
INSA de Rennes
Rennes, France
{yiliu, olivier.deforges, fpasteau, khouloud.samrouth1}@insa-rennes.fr

***Abstract* — This paper presents a joint quality control (QC) and rate distortion optimization (RDO) algorithm applied to a still image codec called Locally Adaptive Resolution (LAR). LAR supports scalability in resolution for both lossy and lossless coding and has low complexity. This algorithm is based on the study of the relationship between compression efficiency and relative parameters. The RDO model is proposed firstly to find suitable parameters. Relying on this optimization, relationships between the distortion of reconstructed image and quantization parameter can be described with a new linear model. This model is used for parametric configuration to control compression distortion. Experimental results show that this algorithm provides an effective solution for an efficient one pass codec with automatic parameters selection and accurate QC. This algorithm could be extended to codecs with similar functions, such as High Efficiency Video Coding (HEVC).**

Keywords — quantization; quality control; rate distortion optimization; image coding

I. INTRODUCTION

Digital images can be found in most of today's computer media. To satisfy the increasing demands of media data exchange, image compression techniques are widely used in various image/video coding systems. Besides this traditional need for image/video coding systems, new services and applications ask for more capacities and functionalities such as scalability, data security, error robustness, Rate Control (RC) and unique algorithm for lossy and lossless coding.

When it comes to existing still image coding standards, the well-known JPEG codec is still commonly used despite its limited functionalities. The main reason is that it still provides a good compromise between compression efficiency and complexity. Conversely, JPEG2000 [1-2] is currently the best standard solution outperforming JPEG up to 2db in PSNR and provides most of the previously mentioned functionalities (scalability in both resolution and quality, efficient lossy and lossless solutions, etc). The main limitation of JPEG2000 is its complexity which can explain why it has not superseded and replaced JPEG. JPEG XR [3] was recently accepted as a new standard by the JPEG committee. JPEG XR can be considered as an intermediate solution between JPEG and JPEG2000 in terms of performances, functionalities and complexity.

In our previous work [4], we proposed an efficient content based image coding called LAR (Locally Adaptive Resolution). It has a global coding framework providing a number of functionalities such as lossy/lossless compression, resolution and quality scalability, partial/full encryption, and Region of Interest coding. The complexity of the LAR is similar to that of JPEG XR. Based on a multi-resolution representation of the image, the LAR provides a natural scalability in resolution. The major drawback is that the encoding process requires setting different parameters in the multi-layer structure. In order to solve this problem automatically and efficiently, a Rate Distortion Optimization (RDO) technique adapted to the LAR is proposed in this paper. The well-known Lagrange-based framework used in JPEG2000 provides significant improvements, but at the expense of computation cost [5]. The solution proposed in this paper is similar with the method proposed on Rate Control based on the ρ-domain for H.264/MPEG AVC [6] for a low complexity. It mainly consists in building a global model to estimate the rate/distortion behavior of the codec. Instead of setting the target rate, we set the target quality (expressed in MSE), and conduct the RDO jointly with a Quality Control (QC). With this model, the coding also provides a scalable complexity.

This paper is organized as follows. A brief introduction to LAR is given in Section II. In Section III, the rate-distortion optimization model based on compression efficiency is described. The quality control scheme is presented in Section IV. Experimental results are shown in Section V. Conclusions are provided in Section VI.

II. LOW BIT-RATE COLOR LAR CODEC

In this section, a framework of a 2D still image compression method called Locally Adaptive Resolution (LAR) is introduced. LAR is based on the idea that an image can be considered to be made up of a global image information (low resolution component) and a texture information (detail component). For this reason, it is a two-layer codec: a Flat coder for the global image information and a Texture coder for detail information (Fig. 1). Both the Flat coder and the Texture coder support the multi-level pyramidal structure [7] that builds a variable resolution image. This structure relies on a quadtree partitioning encoded in the Flat coding stage. The Flat coding scheme is given in Fig. 2. It starts with a pyramidal partition of the original image, followed by pixel value predictions in the use of Wu prediction algorithm [8]. The prediction error is then

sent to be quantized. Bi-linear filter is used to smooth block effects for the post-processing at low bitrates.

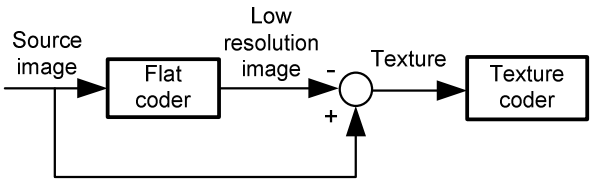


Fig. 1. General scheme of two-layer LAR coder

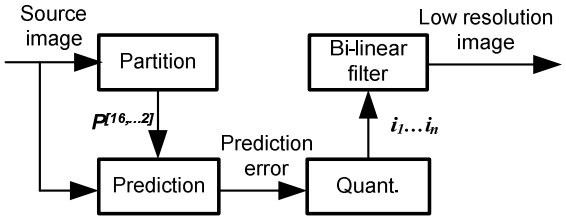


Fig. 2. Flat coder schemes

Combined with the Texture coder, LAR codec supports scalability in both resolution and PSNR. However, this structure increases codec complexity and requires more bit-resources. Two parameters: the threshold (*Thr*) for the quadtree partitioning and *quqp* for the quantization process in (2), which are introduced in Section III, enable the Flat codec to obtain decoded images with different qualities: lossless coding with *Thr* = 0 and *quqp* = 1; lossy coding with other parametric combinations. Without the Texture coder, LAR loses the scalability function in PSNR, but has lower complexity. This paper will study the relationship between the performance and parameters (*Thr* and *quqp*) used in the Flat coder to improve the compression efficiency of the codec. The objective is to find the suitable set of parameters for a quality target. With this method, the complexity of LAR is directly dependent on the number of blocks of quadtree partitioning and is approximately linear with the bitrate. Therefore, at low bitrates, fewer blocks and less time are required for coding.

III. Parameter Optimization

In this section, a parameter optimization method is introduced based on the performance of the LAR codec. In the lossy coding mode of LAR, compression errors are introduced in two parts: quadtree partitioning and the quantization process. Parameters correlated with these two parts are considered together in models to maintain high coding quality.

A. Quadtree Partitioning

The quadtree decomposition relies on a homogeneity criterion. Each square block $P^{[Nmax...Nmin]}$ in the quadtree has a side length equaling to a power of two, where N_{max} and N_{min} represent the maximum and minimum allowed block sizes. Let $I(x, y)$ represent the pixel of coordinates (x, y) in image I of size $N_x \times N_y$, $I(b^N(i, j))$ be the block $b^N(i, j)$ of size $N \times N$, described as follows:

$$b^N(i,j) = \{(x,y) \in N_x \times N_y, \\ N \times i \le x < N \times (i+1), N \times j \le y < N \times (j+1)\}. \quad (1)$$

The quadtree partition is used for the detection of local activity. The contour of an object can be marked by blocks and the size of a block is determined by the difference between the maximum and minimum luminance values. The difference is compared with a threshold to decide whether a block should be separated or not. This threshold is represented by *Thr* mentioned in section II and determined by the proposed model. For color images with three components Y:Cr:Cb, a single threshold mainly calculated from the Y component is chosen.

B. Quantization Process

Prediction errors are uniformly quantized. When an image is processed in different resolution levels, the quantization factor Q is controlled by the global quantization parameter *quqp*, which can be set at the beginning of coding.

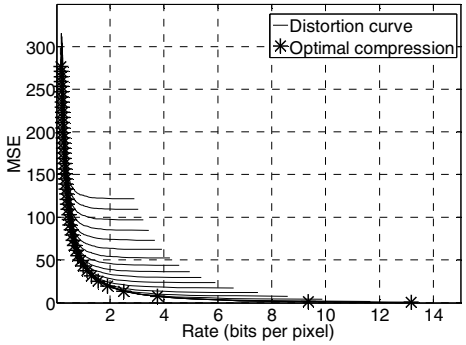
$$Q_{Li} = quqp \cdot F_{Li}, \quad for\ level\ Li. \quad (2)$$

Factor F_{Li} is fixed in the codec. Since predictions are based on intra and inter-level data, quantization performed at a given level will impact all the following levels. Therefore, the F_{Li} factors reflect the distribution of the quantization (distortion) among the pyramid levels.

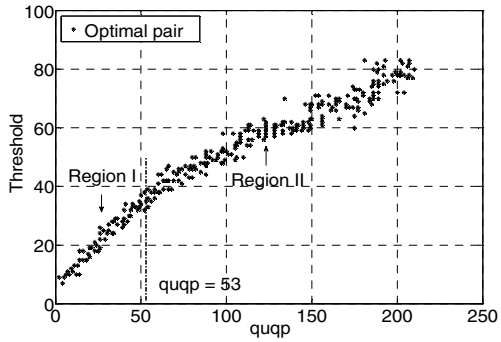
C. Proposed quqp-threshold Model

To optimize the configuration, it is impractical to change parameters repeatedly until suitable choices are found. On the other hand, based on features of the compression efficiency, it is possible to find a way to optimize the efficiency at a lower cost. In this part, the relationships between *quqp* and *Thr* are studied and an optimization scheme is proposed.

An example of the compression distortion for the image "bike_crop" (1280×1600) is shown in Fig. 3(a). To each distortion curve, the *Thr* is set and *quqp* increases. Thus, each curve reflects the distortion trend caused by the quantization at a specific threshold. The optimal compression distortion should be the lowest in both mean squared error (MSE) and bitrate. Pairs of *quqp* and *Thr*, corresponding to the optimal dots in Fig. 3(a), are connected to draw Fig. 3(b). It can be seen that optimal combinations of *quqp* and *Thr* are given by the curve with an inflexion point approximately at *quqp* = 53. Similar situations can be found from other images, and Fig. 4 shows optimal combination curves for four images that have different texture complexities from low to high respectively.



(a) Measured curve



(b) Measured combination

Fig. 3. Measured figures

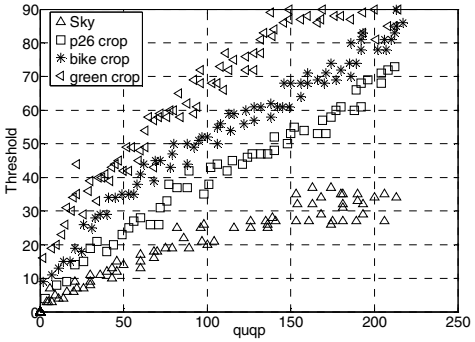


Fig. 4. Optimal combination curves of images

To describe the curve trend clearly, the curve is divided into two regions as seen in Fig. 3(b). Linear models are used for each region and correlative factors are presented here. Firstly, H_G is introduced to describe the variety of adjacent pixels. An image is separated into 2×2 blocks. The difference between the maximum and minimum luminance values in each block is named the gradient of this block. Based on the probabilities of gradients, an entropy form is adopted to represent the variety as shown in (3).

$$H_G = -\sum_i p_i(g)\cdot \log_2 p_i(g), \quad where\ g = \left|\max\left[\left(I(b^2(x,y))\right)\right] - \min\left[\left(I(b^2(x,y))\right)\right]\right| \tag{3}$$

$p(g)$ is the probability of the gradient. H_G values of example images in Fig. 4 are shown in Table I and they generally correspond to the distributions of curves.

TABLE I. ENTROPIES OF THE GRADIENTS

	Image examples			
	Sky	***p26 crop***	***bike crop***	***green crop***
H_G	***2.335***	***4.495***	***5.498***	***5.892***

H_G does not concern values of gradients. However, the amplitudes of gradients can show the contrast of local pixels. Therefore, another factor $r(i)$ is used here. It is a cumulative probability distribution function, as given in (4).

$$r(i) = \sum_{j=0}^{i} p(g = j) \tag{4}$$

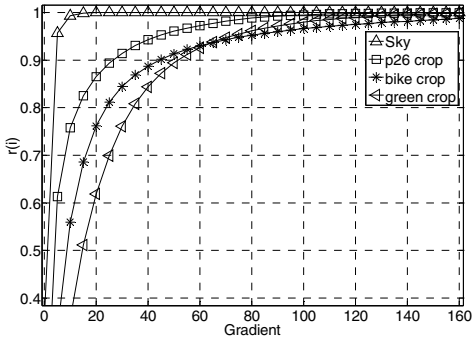


Fig. 5. $r(i)$ curves of images

Fig. 5 presents four $r(i)$ curves of example images. If an image has large portions of the same color and moderate transitions, most gradients have small values and the $r(i)$ curve increases more quickly. To reflect the speed of this trend, a difference between $r(45)$ and $r(7)$ is used expectedly. After tests on training images in high resolution, this difference reflects the main growth rate of the $r(i)$ curve. Generally, small values of the difference correspond to high $r(i)$ curves. It means $r(i)$ raises quickly in small gradient parts.

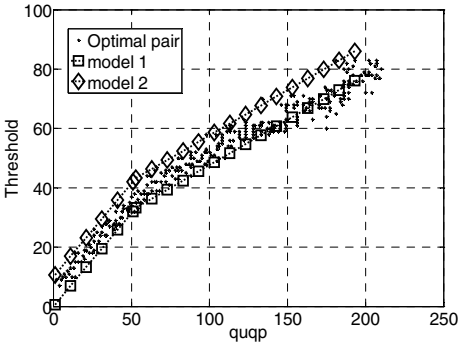
$$\Delta = r(45) - r(7) \tag{5}$$

Based on the studies above, the proposed RDO model is expressed in (6). In each region shown in Fig. 3(b), two linear models are used to simulate the boundaries of the belt. Thr_1 is the result of model 1 and Thr_2 is the model 2. α and β are two parameters obtained by linear regression in an training database. This database includes 8 cropped images from the JPEG test set ISO 12640, and other 12 free images in high resolution. To each training image, α and β are optimized by curve fitting. $\alpha = 17.93$ and $\beta = 121.07$ are average values of α and β of the training images, and selected under the consideration of the overall performance.

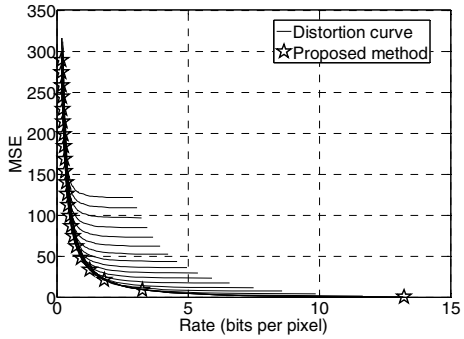
$$\begin{cases} \begin{cases} Thr_1 = \dfrac{H_G}{\alpha}\left(quqp + \Delta\cdot\beta\right) \\ Thr_2 = \dfrac{H_G}{\alpha}\left(quqp + \Delta\cdot\beta\right) + 10 \end{cases} & , if\ quqp \geq 53 \\ \begin{cases} Thr_1 = \dfrac{H_G}{53\cdot\alpha} quqp\left(53 + \Delta\cdot\beta\right) \\ Thr_2 = \dfrac{H_G}{53\cdot\alpha} quqp\left(53 + \Delta\cdot\beta\right) + 10 \end{cases} & , if\ quqp < 53 \end{cases} \tag{6}$$

Fig. 6 (a) shows the simulation result to "bike_crop" . It can be seen that most optimal pairs locate in the region drawn by models. In practical use, the value of Thr for coding can be the average of model 1 and 2.

$$Thr_{use} = (Thr_1 + Thr_2)/2 \tag{7}$$



(a) Simulated combination



(b) Simulated curve

Fig. 6. Simulated figures

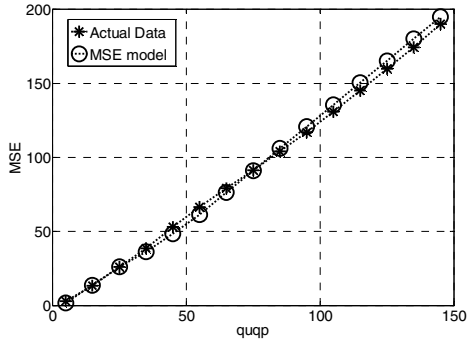
Fig. 6 (b) shows the compression efficiency of this proposed method to "bike_crop". The performance is exactly or close to the optimal choice.

IV. Quality Constraint

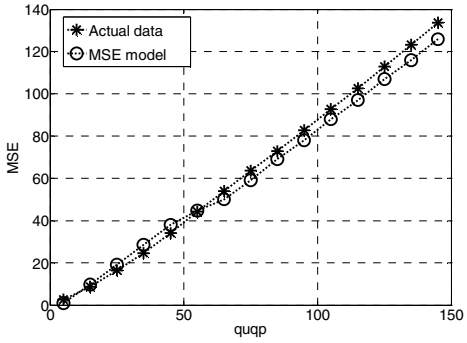
In this section, a quality constraint method is proposed based on the parameter optimization in section III.

A. MSE Determination Model

Based on the observations of image sets encoded by LAR after the parameter optimization, a linear relationship between the quality distortion in MSE and *quqp* is found. This linear relation can be expressed approximately in (8). Similarly, coefficients are determined by the linear regression applied in Section III. MSE_{est} is the estimated MSE value based on parameters. For illustration, the fitting accuracy of the linear MSE determination model is shown in Fig. 7 for two images.



(a) bike crop



(b) p26 crop

Fig. 7. Fitting accuracy of the MSE model

$$\begin{cases} MSE_{est} = (0.058H_G^2 - 0.9\frac{H_G}{\alpha})quqp \\ \qquad\qquad - 0.9\frac{H_G}{\alpha}\Delta - 4.5 & , quqp \geq 53 \\ MSE_{est} = \left[0.058H_G^2 - \frac{0.9}{\alpha}H_G(1-\frac{\Delta}{53}\beta)\right]quqp \\ \qquad\qquad - 4.5 & , quqp < 53 \end{cases} \tag{8}$$

B. MSE Setting Algorithm

Using the MSE Determination Model, it is also possible to control the compression distortion in the opposite way. As shown in equation (9). MSE_{set} is the expected MSE value that is set before encoding. $MSE_{boundary}$ is obtained by equation (8) using *quqp* = 53. After the determination of *quqp* in equation (9), threshold for quadtree can be calculated with the use of equation (6).

$$\begin{cases} quqp = \dfrac{MSE_{set} + 0.9\Delta\frac{H_G}{\alpha} + 4.5}{(0.058H_G^2 - 0.9\frac{H_G}{\alpha})} & , MSE_{set} \geq MSE_{boundary} \\ quqp = \dfrac{MSE_{set} + 4.5}{0.058H_G^2 - \frac{0.9}{\alpha}H_G(1-\frac{\Delta}{53}\beta)} & , MSE_{set} < MSE_{boundary} \end{cases} \tag{9}$$

Then the total encoding starts under optimal selected parameters. The step-by-step description of the algorithm is shown below.

Step 1. *Analyze and compute required parameters such as H_G and r_i of the image to be encoded.*

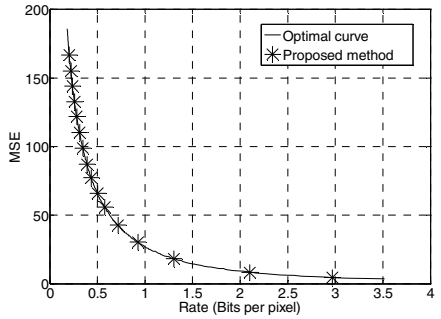
Step 2. *Compute $MSE_{boundary}$ according to the equation (8) using quqp = 53.*

Step 3. *Judged by $MSE_{boundary}$, MSE_{set} is used to calculate the suitable $quqp_{exp}$ with equation (9).*

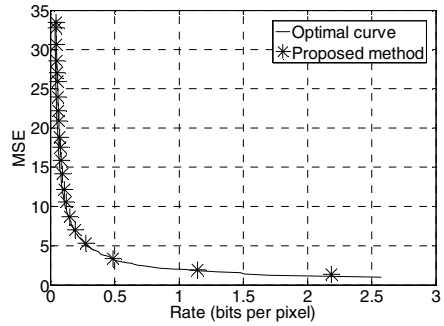
Step 4. *Use $quqp_{exp}$ to obtain the corresponding threshold Thr_{exp} with equation (6) and (7), finish the optimal configuration of the codec.*

V. Experimental Results

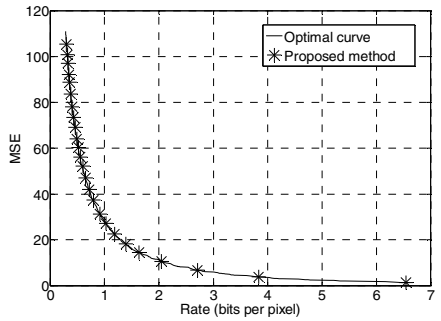
The experimental results are shown in three parts. The first part checks the performance of RDO models, and the second part is dedicated to quality constraint. Finally, a comparison of compression efficiency is made. In experiments, images "p06", "flower_foveon", "leaves_iso_200" and "louvre", which are



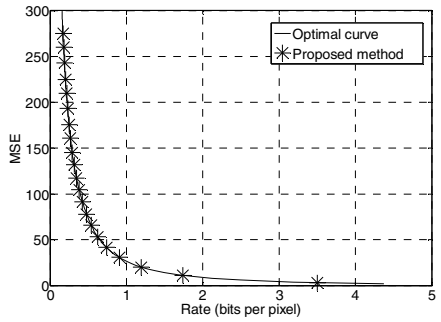
(a) p06



(b) flower_foveon

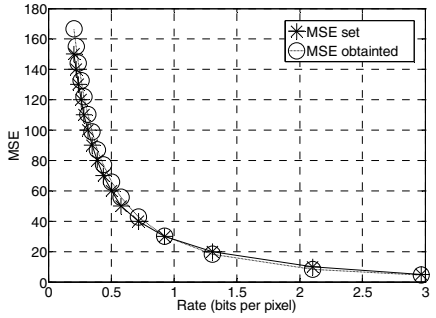


(c) leaves_iso_200

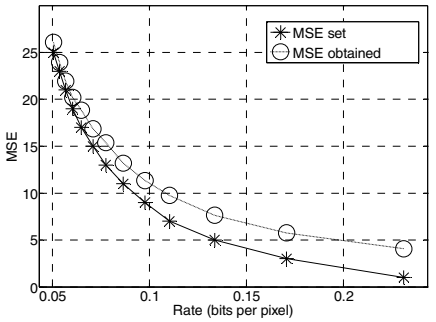


(d) louvre

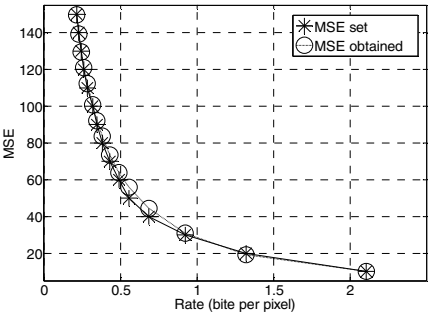
Fig. 8. Compression efficiency results



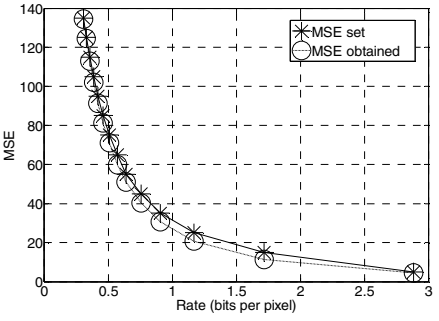
(a) p06



(b) flower_foveon



(c) leaves_iso_200

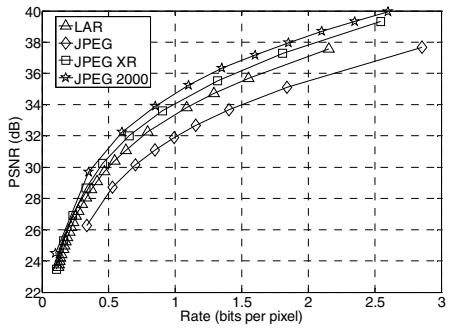


(d) louvre

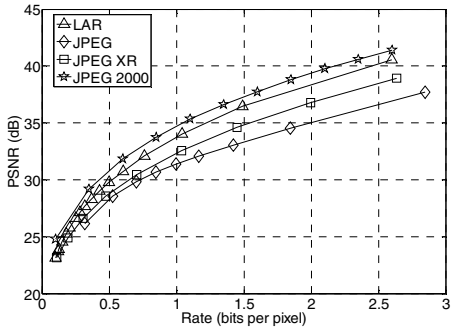
Fig. 9. Comparison between predefined MSE and obtained MSE

outside the training set used in Section III, are used as examples to evaluate the performance. The compression efficiency of the proposed method is shown in Fig. 8. Optimal curves are the best results after total searching with the possible combinations of *quqp* and *Thr*. As shown in Fig. 6 (b), most results of the proposed method locat in the optimal curve and others are close to it.

Fig. 9 shows the performance of the quality constraint method. For each MSE_{set}, the encoder calculates the proper configuration parameters automatically as introduced in Section IV. And then the parameters are used to encode and decode the test image. The decoded image is compared with the original one in order to compute the real obtained MSE. From Fig. 9, it can be seen that the distortion of reconstructed



(a) PSNR performances on bike



(b) PSNR performances on p06

Fig.10. Comparison with other algorithms

image is close to the predefined distortion and confirms the effectiveness of the QC algorithm. Certainly, quality constraint should have been equivalently expressed in PSNR.

To further check the performance, the compression efficiency of the proposed method is also compared with that of other algorithms, JPEG, JPEG XR and JPEG2000. Results of JPEG series can be provided by the JPEG Online Test Facility [9]. Two comparison results for example images, "bike" (2048×2560) and "p06" (4064×2704), of the proposed algorithm are shown in Fig.10. Although LAR does not achieve PSNR as well as JPEG2000, its performances remain close to the state-of-the-art at low bitrates. In addition, it has a better performance than JPEG XR in some cases as "p06" in Fig. 10(b). Considering the useful functionalities of LAR, it is hopeful to offer an ideal choice for image coding.

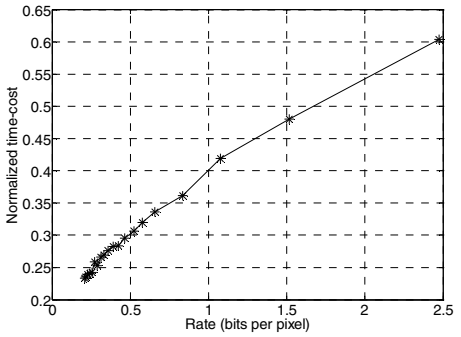


Fig. 11 LAR scalable complexity for "bike"

In LAR codec, the operation of coding depends on the quadtree partitioning. The reduction of the number of blocks can decrease the operations required. This proposed RDO model increases the threshold of quadtree when bitrate declines. This procedure produces less blocks to code at low bitrates. Fig. 11 shows the time-cost for encoding at different bitrates to image "bike" (2048×2560). The time-cost is divided by the one of lossless coding for normalization. Fig. 11 shows that the time-cost is approximately linear with the bitrate and raises with the increase of the bitrate. So this model brings a complexity-scalable relationship with the bitrate.

VI. CONCLUSION

In this paper, a quality constraint algorithm is presented for a predictive image coder. Linear models are first proposed to describe the distribution of the optimal combination of parameters. Based on the optimization, a linear model is found to simulate compression distortion. Depending on a prior work, a quality constraint algorithm focusing on MSE and quantization is introduced. Experimental results show that this algorithm reaches an ideal performance that is close to the optimal coding based on the compression efficiency of the codec.

Instead of exhaustive searches, the proposed technique helps LAR select suitable parameters to realize rate-distortion optimization. The method could be extended to other coders based on similar schemes, such as HEVC that also employs quadtree partitioning and intra-picture prediction [10].

REFERENCES

[1] A.Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Process.*, vol. 18, no. 5, pp. 36-58, Sept. 2001.

[2] David S. Taubman, Michael W. Marcellin, *JPEG2000 Image Compression Fundamentals, Standard and Practice*, Kluwer Academic Publishers, p. 402, 2002.

[3] T. Richter, "Visual quality improvement technique of HD Photo/JPEG-XR," *Proc. IEEE ICIP*, pp. 2888-2891. Oct. 2008.

[4] O. Deforges, M. Babel, L. Bedat and J. Ronsin, "Color LAR Codec: a color image representation and compression scheme based on local resolution adjustment and self-extracting region representation," *IEEE Trans. on Circuits and Systems for Video Technol.*, vol 17, pp. 974-987, Aug. 2007.

[5] David S. Taubman, "High performance scalable image compression with EBCOT," *IEEE Trans. on Image Process.*, vol. 9, no. 7, pp. 1158-1170, Jul. 2000.

[6] I.H. Shin, Y.L. Lee and H.W. Park, "Rate control using linear rate-ρ model for H.264," *ELSEVIER Signal Process on Image Comm.,* vol. 19, pp. 341-352, Apr. 2004.

[7] M. Babel, O. Deforges and J. Ronsin, "Interleaved S+P pyramidal decomposition with refined prediction model," *IEEE ICIP*, vol. 2, pp. 750-753, Sept. 2005.

[8] X. Wu and N. Memon, "Context-based, adaptive, lossless image coding," *IEEE Trans. on Comm.*, vol. 45, no. 4, pp. 437-444, Apr. 1997.

[9] University of Stuttgart (2012) JPEG Online Test Facility [Online]. Available: http://jpegonline.rus.uni-stuttgart.de/index.py.

[10] Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, Thomas. Wiegand, "Overview of the High Efficiency Video Coding (HEVC)," *IEEE Trans. on Circuits and Systems for Video Technol.*, vol 22, pp. 1649-1668, Sept. 2012.