Novel Pre-Compression Rate-Distortion Optimization Algorithm for JPEG 2000

Yu-Wei Chang, Hung-Chi Fang, Chung-Jr Lian, and Liang-Gee Chen

DSP/IC Design Laboratory, Graduate Institute of Electronics Engineering
National Taiwan University
1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan

ABSTRACT

In this paper, a novel pre-compression rate-distortion optimization algorithm is proposed, which can reduce computation power and memory requirement of JPEG 2000 encoder. It can reduce the wasted computational power of the entropy coder (EBCOT Tier-1) and unnecessary memory requirement for the code-stream. Distortion and rate of coding passes are calculated and estimated before coding, and therefore truncation point is selected before coding. Experimental results show that the computation time of EBCOT Tier-1 and memory requirement for the code-stream can be greatly reduced, especially at high compression ratio. The quality of the proposed algorithm is slightly lower than that of post-compression rate-distortion optimization algorithm.

Keywords: JPEG 2000, rate control, EBCOT, low power, rate distortion optimizatoin.

1. INTRODUCTION

JPEG 2000[1] is the new still imaging coding standard in next generation. The performance of JPEG 2000 is superior to JPEG at all bit-rate.[2] However, the computational complexity of JPEG 2000 is much higher than that of JPEG. There are two major parts in JPEG 2000: Discrete Wavelet Transform (DWT), and Embedded Block Coding with Optimized Truncation (EBCOT).[3] In general, quantization is not used to control the rate of code-stream in JPEG 2000 encoder. It is applied to adjust weights of different frequency bands based on the filter bank and decomposition level of DWT, and no quantization is used at reversible wavelet transform mode. After DWT and quantization, the coefficients are partitioned into code-blocks, which are encoded by EBCOT. The most complex part in JPEG 2000 is EBCOT. EBCOT is a two-tiered algorithm. Tier-1 is embedded block coder, which utilizes context-based arithmetic coding to encode each code-block into independent embedded bit-stream. Tier-2 is post-compression rate-distortion optimization algorithm. EBCOT Tier-1 is the most complex part of JPEG 2000, which consumes more than 50% of total computation power [4].[5] Reducing its computation time can significantly decrease the total run time of JPEG 2000 encoder.

Most lossy still image coding standard, including JPEG, use quantization scheme to achieve rate control. However, this scheme cannot provide best quality at a given bit-rate and get precise rate in one iteration. Instead of using quantization scheme to perform the rate control, JPEG 2000 uses a better scheme to control the rate by EBCOT Tier-2 processing. It uses lagrange optimization to precisely control the bit-rate and guarantees the best quality at specific bit-rate. However, in the rate-distortion optimization, all transformed coefficients must be processed by EBCOT Tier-1 to get the rate and distortion information. In most cases, most compressed bit-streams generated by EBCOT Tier-1 will be discarded through the procedure of EBCOT Tier-2. The memory spent to store the discarded bit-stream and computations used are all wasted. Some previous works, Chang's,[6] Yeung's[7],[8] focus on memory and power reduction for PCRD. Chang et al.,[6] use EBCOT Tier-2 feedback control to terminate redundant computation of EBCOT Tier-1. Computation time of EBCOT Tier-1 can be reduced to 40% and 20% at medium to high compression rate. Yeung [78] proposed a scheme based on priority scanning. It is to encode the truncation points in a different order by priority information and terminate block coding adequately. The computational cost and memory requirement can be reduced by 52% and 71% respectively in the case of 0.25 bpp.

In this paper, a new idea of quality control is proposed. Instead of controlling the bit-rate, it minimizes the bit-rate at given image quality, that is, distortion constraint. Quality control has an excellent feature that it is a pre-compression rate-distortion optimization algorithm. By selecting truncation point before coding, wasted computational power and unnecessary memory requirement is eliminated. The quality slightly degrades comparing with the post-compression rate-distortion algorithm.

Visual Communications and Image Processing 2004, edited by Sethuraman Panchanathan,
Bhaskaran Vasudev, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 5308

This rest of this paper is organized as follows. Rate-distortion optimization algorithm is reviewed in Section 2. Section 3 and 4 will explain quality control algorithm in detail. Experiment results are shown in Section 5. Finally conclusion is drawn in Section 6.

2. RATE-DISTORTION OPTIMIZATION

In JPEG 2000, original image is decomposed as several subbands by DWT and each subband is partitioned into code-blocks. EBCOT Tier-1 compresses a code-block as an independent embedded bit-stream which is composed of several coding passes. An embedded bit-stream of code-block *i* may truncated at some feasible truncation point, n_i, to yield rates, $R_i^{n_i}$. The total rate *R* of the code-stream is

$$R = \sum_i R_i^{n_i}. \tag{1}$$

The distortion corresponding to n_i is denoted by $D_i^{n_i}$, which is measured by the Mean Square Error (MSE). The total distortion of the code-stream is

$$D = \sum_i D_i^{n_i}. \tag{2}$$

For a Rate-Distortion (R-D) optimized code-stream with rate *R* and distortion *D*, it is impossible to lower *D* without increase of *R* or vice versa. Taubman[3] has solved the R-D optimization problem for rate control. For rate control, the goal is to minimize the distortion while keeping the rate of code-stream smaller than a target bit-rate, R_T. The problem has been mapped into Lagrange optimization problem as the minimization of

$$D + \lambda R = \sum_i (D_i^{n_i} + \lambda R_i^{n_i}). \tag{3}$$

An R-D slope for a code-block at each candidate truncation point *n*, which is defined as

$$S_i^n = \frac{\Delta D_i^n}{\Delta R_i^n} = \frac{D_i^{n-1} - D_i^n}{R_i^n - R_i^{n-1}}, \tag{4}$$

is used to find the optimal λ^* and the optimal truncation point set z^*. With increase of *n*, the candidate truncation point will be from the most bit-plane to least bit-plane. Therefore, the slope value must decrease monotonically, that is, the value of S_i^n must be larger than S_i^{n+1}. λ^* and z^* is optimal if the following conditions is satisfied

$$\begin{cases} S_i^n \geq \lambda^*, & n \leq n_i \\ S_i^n < \lambda^*, & n > n_i \end{cases} \tag{5}$$

$\forall n_i \in z^*$ and for all code-blocks.

In additional to rate control, one can minimize the total rate, *R*, at the target distortion, D_T. This is called "quality control". As in rate control, the R-D optimization of quality control can be achieved by minimizing

$$R + \lambda' D = \sum_i (R_i^{n_i} + \lambda' D_i^{n_i}) \tag{6}$$

where λ' is the Lagrange multiplier for quality control. It can be readily proved that (3) and (6) are equivalent by replacing $1/\lambda'$ by λ,

$$\begin{aligned} min(R + \lambda' D) &= min(\lambda'(\frac{1}{\lambda'} R + D)) \\ &= \lambda' min(D + \lambda R). \end{aligned} \tag{7}$$

The optimal truncation point set of rate control problem that resulting the distortion *D* is the same as the quality control problem with $D_T = D$. Furthermore, the resulting rate *R* of quality control will be the same as the rate constraint R_T of the rate control problem.

The concept of quality control is contrary to that of rate control. Rate control minimizes distortion of the code-stream at given rate constraint while quality control minimizes rate of the code-stream at given distortion constraint. It is believed that the quality constraint is more meaningful than rate constraint for image compression in general. Besides, quality control has an advantage that it can be done at DWT domain, i.e. before the EBCOT Tier-1, without significant quality loss.

3. RATE-DISTORTION CALCULATION IN DWT DOMAIN

The distortion of a code-block is measured as the Mean Square Error (MSE) of the reconstructed pixels to the original ones. A sample coefficient must be checked whether or not it belongs to magnitude refinement pass (P_2) since the reconstructing approach of P_2 is different from that of significant propagation pass (P_1) and cleanup pass (P_3). The checking of whether a sample coefficient belongs to P_2 can be done prior to EBCOT. Let μ^k be the sample bit at k-th bit-plane of a j-th coefficient μ, and n be the candidate truncation point. Let ϕ^k denotes whether μ^k belongs to P_2 or not, which can be found as

$$\phi^k = \begin{cases} 1, & 2^{k+1} \le \mu \\ 0, & 2^{k+1} > \mu \end{cases} . \tag{8}$$

Let t^k be the value of μ lower than k-th bit-plane, which is the value of the truncated part. The delta distortion for coefficient belonging to candidate truncation point $n \in P_2$ at bit-plane k is

$$\Delta d_j^n = \begin{cases} (t^{k+1} - 2^k)^2 - (t^k - \tilde{t^k})^2, & \phi^k = 1 \\ 0, & \phi^k = 0 \end{cases} \tag{9}$$

The distortion for coefficient of candidate truncation point $n \in (P_1 or P_3)$ at bit-plane k is

$$\Delta d_j^n = \begin{cases} 0, & \phi^k = 1 \\ (t^{k+1})^2 - (t^k - \tilde{t^k})^2, & \phi^k = 0 \cap u^k = 1 \\ (t^k)^2, & \phi^k = 0 \cap u^k = 0 \end{cases} \tag{10}$$

where $\tilde{t^k}$ is the reconstructed value of t^k and is given by

$$\tilde{t^k} = \begin{cases} 2^{k-1}, & k > 0 \\ 0, & k = 0 \end{cases} . \tag{11}$$

and the the delta distortion, $\Delta D_i^{n_i}$, can be calculated as

$$\Delta D_i^n = \sum_{j=0}^{N^2-1} \Delta d_j^n \tag{12}$$

where N is code-block size.

The rate, on the other hand, must be estimated since the exactly rate can only be known after EBCOT. In order to select the optimized truncation point before EBCOT, the slope of each truncation point must be estimated. The randomness property of sample coefficients in P_2 is utilized to estimate the rate. Figure 1 shows the randomness property of P_2. Each point in Fig. 1 represents one P_2 bit-stream and the x-axis and y-axis is the number of sample coefficient and the resulting rate of it. It is obvious that the compression ratio of P_2 is almost constant regardless of different image types, subbands, code-block sizes and which bit-plane it belongs to. Therefore, the rate of P_2 can be accurately estimated. In order to increase the number of feasible truncation points, the propagation property of P_1 is used. In the lowest two bit-planes of an code-block, most of coefficient has been significant, and therefore almost all the non-significant sample coefficients is propagated in P_1 as shown in Fig. 2. As can be seen, almost all the non-significant sample coefficients are propagated in P_1 for the lowest two bit-planes. The delta rate of the truncation point at bit-plane k for $n \in P_2$ is estimated as

$$\Delta r_j^n = \begin{cases} 1, & \phi^k = 1 \\ 0, & \phi^k = 0 \end{cases} . \tag{13}$$

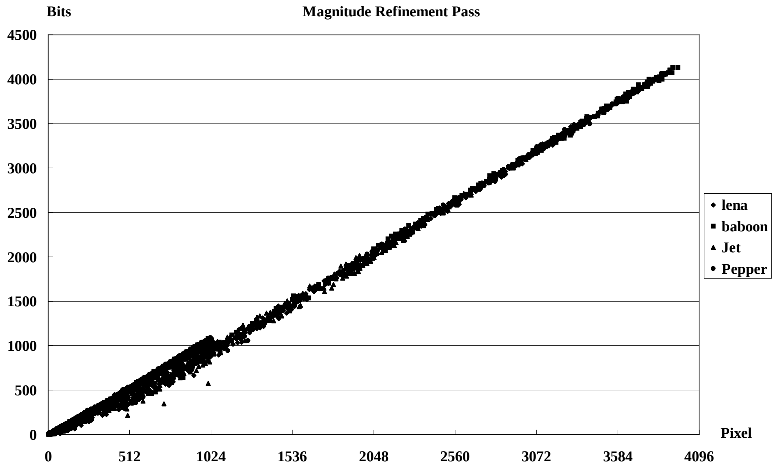


Figure 1: Randomness property of the magnitude refinement pass.

The delta rate of the truncation point at bit-plane $k < 2$ for $n \in P_1$ is estimated as

$$\Delta r_j^n = \begin{cases} 1, & \phi^k = 0 \\ 0, & \phi^k = 1 \end{cases} . \tag{14}$$

The total delta rate, ΔR_i^n, can be estimated by

$$\Delta R_i^n = \begin{cases} \omega_{P_2} \times BC_i^n, & n \in P_2 \\ \omega_{P_1} \times BC_i^n, & n \in P_1 \cap k < 2 \\ \infty, & otherwise \end{cases} \tag{15}$$

and

$$BC_i^n = \sum_j^{N^2} \Delta r_j^n \tag{16}$$

where ω_{P_2} and ω_{P_1} are the empirically decided compression ratio of P_2 and P_1 of lowest two bit-planes, and BC_i^n is bit-counts of truncation point n.

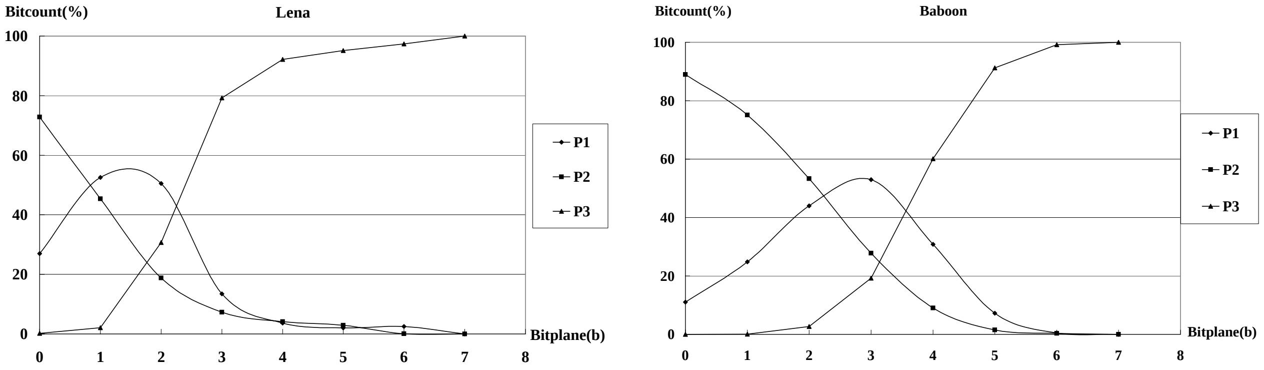
4. QUALITY CONTROL

The flow chart of proposed quality control algorithm is shown in Fig. 3. It has two stages: accumulation stage and decision stage. Rate and distortion are calculated and accumulated in the accumulation state. Optimal truncation point set is selected in the decision stage. Let the the depth of magnitude bit of DWT coefficient is H. P_1 and P_2 are truncation points for bit-plane $k < 2$ and only P_2 is truncation point for $2 \leq k \leq H-1$. Therefore,the number of truncation points are $H+2$. Let the value of n between 0 and $H+1$ $(0 \leq n \leq H+1)$ denotes the candidate truncation point from higher bit-plane to lower bit-plane and two sets, A_{P_1} and A_{P_2}, represent the collection of candidate truncation points P_1 and P_2 respectively, i.e.

$$n \in \begin{cases} A_{P_1}, & n = H+1, H-1 \\ A_{P_2}, & 0 \leq n \leq H-2, n = H \end{cases} . \tag{17}$$

4.1. Accumulation stage

This stage including Pass Detection, R-D Calculation and R-D Accumulation. Pass Detection step detects whether or not the bit belongs to P_2. R-D Calculation step estimates rate and computes distortion for each bit. R-D Accumulation step accumulates rate and distortion to obtain ΔD_i^n and ΔR_i^n for each truncation point. Accumulation stage is repeated until coefficients of a tile are finished.



(a) (b)

Figure 2: Propagation property of P_1 for (a) lena and (b) baboon.

4.1.1. Pass Detection

When one DWT coefficient μ at position j in code-block i is fetched in, pass of each bit at each bit-plane k is detected by using equation (8). Then the ϕ^k $(0 \leq k \leq H-1)$ information are obtained.

4.1.2. R-D Calculation

The ϕ^k information are exploited to calculate distortion and estimate rate. The delta rate and distortion for each truncation point n is calculated according to the equation (9) and (13) if $n \in A_2$ and (10) and (14) if $n \in A_1$. Then Δd_j^n and Δr_j^n for $0 \leq n \leq H+1$ are obtained.

4.1.3. R-D Accumulation

ΔD_i^n and ΔBC_i^n are set to zero initially for $j = 0$. To get ΔR_i^n and ΔD_i^n, Δd_j^n and Δc_j^n are accumulated with

$$\begin{aligned} \Delta D_i^n &= \Delta D_i^n + \Delta d_j^n \\ \Delta BC_i^n &= \Delta BC_i^n + \Delta r_j^n \\ j &= j+1 \end{aligned} \tag{18}$$

for n=0 to n=H+1. If $j = N^2 - 1$ at the end of one code-block, the ΔBC_i^n must be multiplied with ω_{P_1} or ω_{P_2} to get ΔR_i^n with the equation (15), and ΔR_i^n and ΔD_i^n for code-block i are obtained.

4.2. Decision stage

This stage including Calculate Slope step and Truncation Points Selection step. The Calculate Slope step computes the slope value for all code-block i with the information ΔR_i^n and ΔD_i^n generated by stage 1 by using equation (4).

The target distortion of image are averagely distributed into every tile and the distortion of every tile is D_T. The Truncation Points Selection step adjusts threshold value λ^* with several iterations to satisfy the condition (5) without total distortion of all code-block exceeding D_T. Then the optimal truncation points n_i for all every code-block is found. These information are exploited by EBCOT Tier-1 to compresses code-blocks without processing truncated parts.

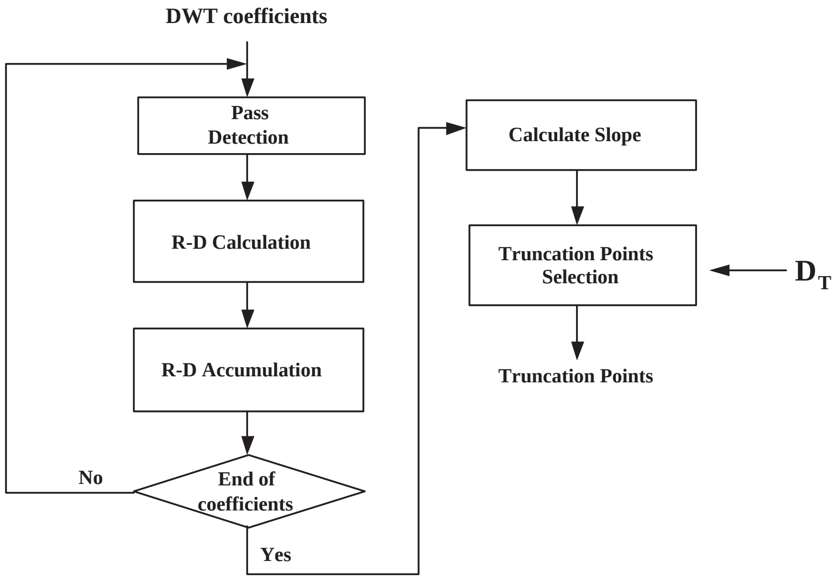
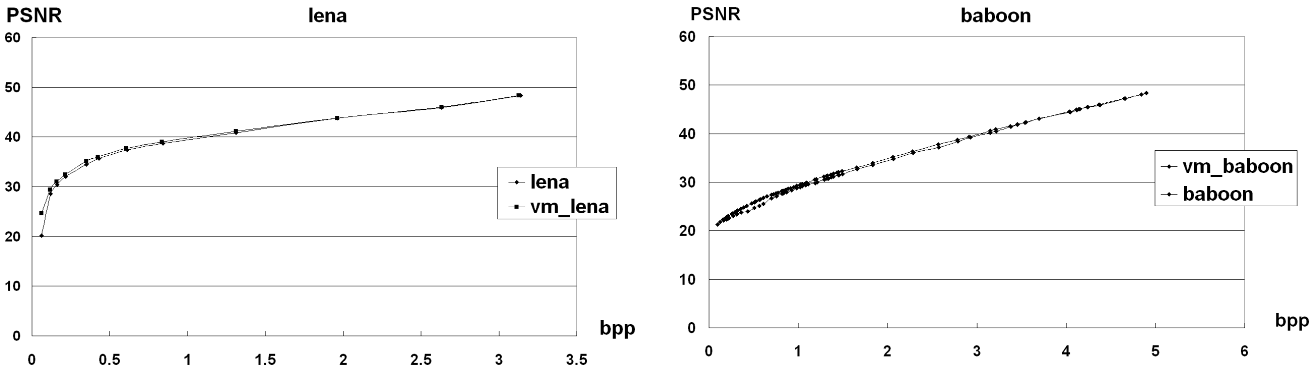
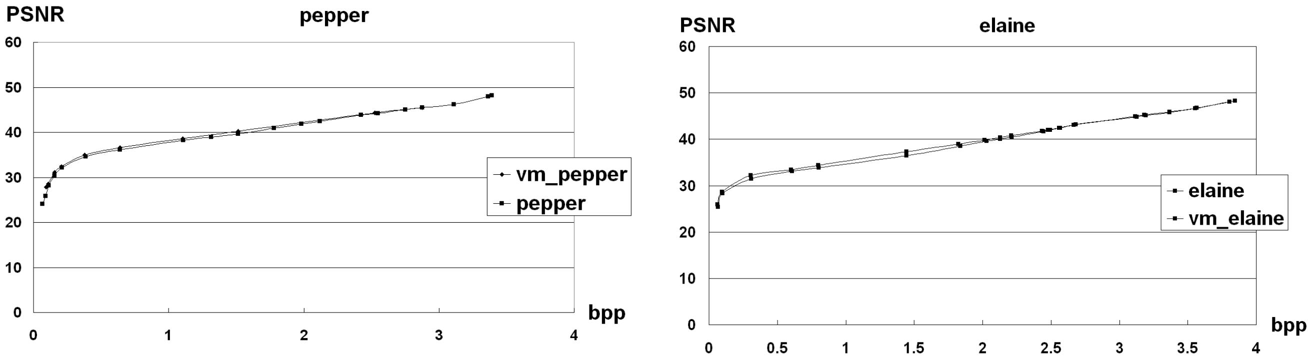


Figure 3: Flow chart of proposed quality control.



(a) lena

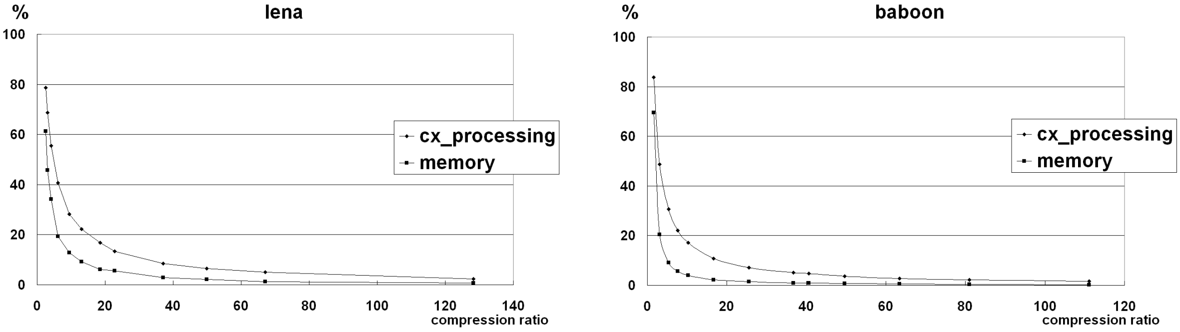
(b) baboon



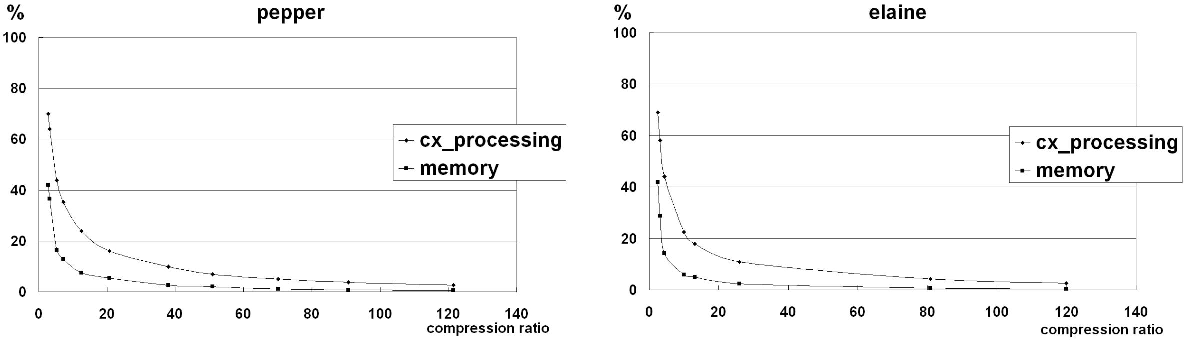
(c) pepper

(d) elaine

Figure 4: PSNR comparison between proposed method and VM9.0. (Test image: 512 × 512, 1 tile, 1 layer, 2 decomposition level)



(a) lena (b) baboon



(c) pepper (d) elaine

Figure 5: Normalized computation power and memory usage. (Test image: 512x512, 1 tile, 1 layer, 2 decomposition level)

5. EXPERIMENT RESULT

5.1. Experimental results

The PSNR value of proposed quality control algorithm, which is compared with VM9.0, is shown in Fig. 4 and the detail is listed in Table 1. For these four test images, the value of PSNR difference between proposed method and VM9.0 is almost the same at median and high bit-rate. At low bit-rate, the average PSNR value degradation below VM9.0 is about 0.3∽0.7 dB. The worst case may be over 1 dB at very low bit-rate. There are two reasons about PSNR degradation. First, the rate of each truncation point is estimated. Estimated rate may not be the same as the real rate. Second, truncation points of proposed algorithm are only P_2 at all bit-plane and both P_1 and P_2 at lower two bit-planes. The truncation points selected by proposed algorithm may not be optimal set, because some truncation points can not be chosen for the sake of lacking slope information.

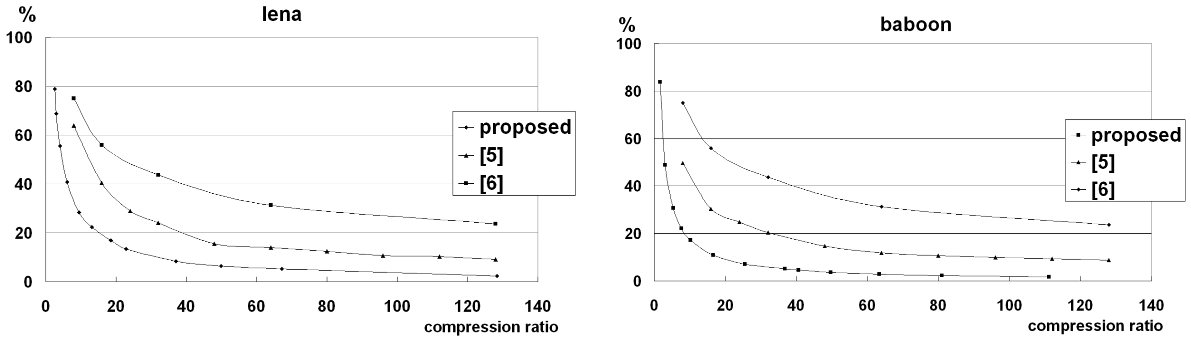
The normalized computation power and memory usage for buffering bit-stream is shown in Fig. 5. The computation power is measured with the number of contexts processed by EBCOT Tier-1 since it is highly correlated to processing time of EBCOT Tier-1. The memory usage is measured with the number of bytes buffered in memory. Compared with conventional post-compression rate-distortion optimization, the proposed algorithm can reduce considerable computation power and memory requirement.

5.2. Comparison

The normalized computation power and memory usage compared with Chang's work [6] and Yeung's work[7] is shown in Fig. 6 and Fig. 7 respectively. In Fig. 6, computation power is measured with the number of contexts for our work and Change's.[6] For Yeung's[7] approach, it is measured with the number of passes needed to be processed. In Fig. 7, Normalized memory usage is measured similarly with bytes stored in memory Compared with Chang's [6] work and Yeung's[7] work, our proposed algorithm has better computation reduction and lower memory requirement. The reason why Chang's [6] work

Table 1: PSNR value compared with VM9.0 for four test image.

Image	Compression Ratio	VM9.0	Proposed
Lena	3	46.03	45.92
	13	37.66	37.36
	37	32.45	32.01
	68	29.40	28.58
Baboon	2	44.58	44.44
	7.5	29.88	29.54
	22	24.52	23.73
	50	22.29	22.09
Pepper	3.10	44.28	44.24
	12.5	36.57	36.14
	37	32.41	32.15
	70	28.53	28.17
Elaine	3	43.06	43.17
	10	34.39	33.84
	32	32.19	31.48
	80	28.69	28.30



(a) lena

(b) baboon

Figure 6: Comparison of normalized computation power.

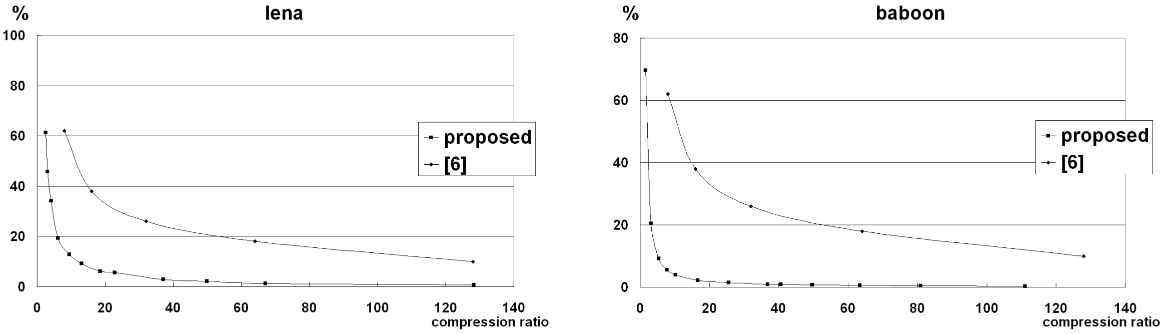
and Yeung's[7] work have less power and memory reduction is that these two works focus on rate control. Although there are many passes can be skipped during encoding process, some passes, which will be discarded finally for the rate control issue, still need to be encoded by EBCOT Tier-1.

6. CONCLUSION

A pre-compression rate-distortion optimization scheme is proposed in this paper. A novel idea, quality control, for the coder control of JPEG 2000 is also proposed. The quality control exploits the randomness property of P_2 and the propagation property of P_1 to estimate the rate in DWT domain. Distortion calculation in DWT domain is also presented. Experimental results show that the proposed algorithm can greatly reduce the computation time for EBCOT and the memory requirement for bit-stream. The PSNR degrades by only 0.3 dB in average comparing to the reference software.

REFERENCES

1. ISO/IEC JTC1/SC29/WG1 N1855, *JPEG 2000 Part I: Final Draft International Standard (ISO/IEC FDIS15444-1)*, Aug. 2000.



(a) lena (b) baboon

Figure 7: Comparison of normalized memory requirement.

2. A. Skodras, C. Christopoulos, and T. Ebrahimi, “The JPEG 2000 still image compression standard,” *IEEE Signal Processing Mag.* **18**, pp. 36–58, Sept. 2001.
3. D. Taubman, “High performance scalable image compression with EBCOT,” *IEEE Trans. Image Processing* **9**, pp. 1158–1170, July 2000.
4. K. F. Chen, C. J. Lian, H. H. Chen, and L. G. Chen, “Analysis and architecture design of EBCOT for JPEG-2000,” in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS 2001)*, pp. 765–768, 2001.
5. M. D. Adams and F. Kossentini, “Jasper: A software-based JPEG 2000 codec implementation,” in *Proc. IEEE Int. Conf. Image Processing (ICIP 2000)*, pp. 53–56, 2000.
6. T. H. Chang, L. L. Chen, C. J. Lian, H. H. Chen, and L. G. Chen, “Computation reduction technique for lossy jpeg2000 encoding through ebcot tier-2 feedback processing,” in *Proc. IEEE Int. Conf. Image Processing (ICIP 2002)*,
7. Y. M. Yeung, O. C. Au, and A. Chang, “An efficient optimal rate control technique for JPEG2000 image coding using priority scanning,” in *Proc. IEEE Int. Conf. Multimedia Expo. (ICME 2003)*, pp. 277–280, 2003.
8. Y. M. Yeung, O. C. Au, and A. Chang, “An efficient optimal rate control scheme for JPEG2000 image coding,” in *Proc. IEEE Int. Conf. Image Processing (ICIP 2003)*, 2003.