HIGH QUALITY VIDEO CODING FOR MOTION JPEG 20000 BASED ON RATE CONTROL METHOD MAINTAINING CONSTANT MSE

Kiyoshi NISHIKAWA, Sho KUWABARA, Hitoshi KIYA

Dept of Information and Communications Systems Engineering
Tokyo Metropolitan University
6-6 Asahigaoka, Hino-shi, Tokyo, 191-0065 JAPAN
knishikawa@m.ieice.org, kuwabara-sho@sd.tmu.ac.jp, kiya@sd.tmu.ac.jp

ABSTRACT

In this paper, we propose high quality video coding method for JPEG 2000 based on a rate control method which maintains a constant MSE given as a target. In the proposed method, we can specify the target MSE, and additionally, a maximum bit rate simultaneously in order to maintain the desired image quality and file size. For compressing a video sequence using JPEG 2000, we can specify the target bit-rate at encoding time. In general, required bit-rate to maintain a specified MSE differs from frame to frame. Hence, if we need to maintain a constant MSE for a video sequence using the conventional one-pass compression methods, the target bit-rate should be set as the one for the frame whose compression ratio is the worst. In this case, a high target bit-rate that is unnecessary to achieve the required MSE for most frames is used so that the compression ratio become worsen. In the proposed method, we could specify the target MSE at the encoding time, so that we can encode each frame at the lowest bit-rate to maintain the desired MSE for the frame. Hence, by applying it to a video sequence, we can prevent to generate a frame whose MSE is lower than required.

Keywords: JPEG 2000, Image compression, PCRD-opt, Rate control, Image quality, Quantization

1. INTRODUCTION

In this paper, we propose a novel one-pass rate control method maintaining a image quality given as the target MSE (mean squared error) for JPEG 2000[1, 2] coded images. In most of the conventional rate control methods for JPEG2000, we specify a target bit-rate and the image quality will be determined dependent on the rate. On the other hand, in the proposed method, we can specify a target MSE, and additionally a maximum bit rate simultaneously in order to maintain the desired image quality and file size.

Recently, the need for processing a large amount of video sequences in digital format has been increasing, and the techniques for faster and higher quality compression are required. JPEG 2000 is employed as the standard compression method for DCI (Digital Cinema Initiatives)[3] so that it is applied to compress a large amount of high resolution images. For compressing a video sequence using JPEG 2000, we can specify the target bit-rate at encoding time[2, 4, 5]. In general, required bit-rate to maintain a specified MSE differs from frame to frame. Hence, if we need to maintain a constant MSE for a video sequence using the conventional one-pass compression methods, the target bit-rate should be set as the one for the frame whose compression ratio is the worst. In this case, a high target bit-rate that is unnecessary to achieve the required PSNR for most frames is used so that the compression ratio become worse.

In [6, 8], methods for controlling the image quality at the encoding process are proposed. The purpose of the method of [6] is to maintain a constant transfer bit rate when videos are transmitted via the Internet using UDP or similar protocols. At the first stage, the images are encoded by specifying a higher bit-rate than required and the bit stream will be truncated according to the amount of the available network bandwidth. Therefore, it is difficult to specify an specific MSE of the encoded images using the method of [6]. The method of [8] adjusts the quality of the encoded images according to the property of HVS (human visual system) and the images will be optimized under the perceptual distortion measure so that we could not specify the image quality in terms of the MSE.

In the proposed method, on the other hand, we could specify the target MSE, or similarly in the PSNR (peak signal to noise ratio) at the encoding time, so that we can encode each frame at the lowest bit-rate to maintain the desired MSE for the frame. Hence, by applying it to a video sequence, we can prevent to generate a frame whose MSE is lower than required. Note that bit-rate can be also controllable by specifying its upper limit and the target PSNR simultaneously. Also, by employing the proposed method with a conventional rate control method specifying a target bit-rate, it is possible to efficiently reduce the amount of coded data. Hence, we could have the following three ways to specify the target:

1. Only the target bit-rate (Identical to the conventional PCRD (Post-Compression Rate-Distortion) optimization)
2. Only the target MSE (Proposed)
3. Both of the target bit-rate and the MSE (Proposed)

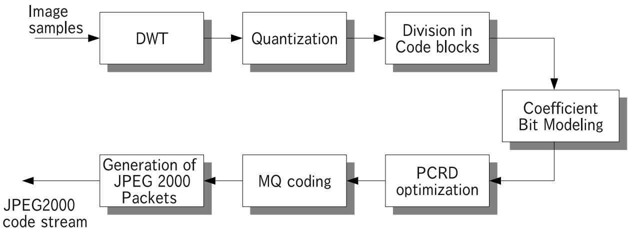


Fig. 1: Flow of JPEG 2000 coding

The proposed method could be used for the applications in which very high quality images are required. Although we could think of lossless compression to maintain the quality of coded images, in this case, it may result in very large file size because we can not specify the allowable maximum bit-rate to lossless encoders. On the other hand, the proposed method enables us to achieve the desired high image quality, for example by setting it as 45dB, with limiting the maximum bit-rate of the coded images within a specified upper limit. This is one of the main advantages of the proposed method over the conventional PCRD optimization methods.

2. JPEG2000 IMAGE CODING

In this section, we review the JPEG 2000 image coding standards, and the PCRD-Opt algorithm.

2.1 JPEG 200 image coding

Here, we summarize the standard procedure of JPEG2000 image encoding. In Fig. 1, the flowchart of the JPEG2000 encoding process.

An input image is divided into sub-bands using discrete wavelet transform (DWT). DWT coefficients in a sub-band are transformed in quantized coefficients by pre-quantization, and then, divided into sub-regions called code-blocks. Quantized coefficients in a code-block are processed as bit-planes, and each bit-plane is divided into three coding passes (SP, MR, CP). Then, rate control using the PCRD optimization is applied where, in each code block, the optimum truncation points are obtained. Once the truncation points are obtained, the bit stream are applied to the MQ encoder, and JPEG2000 packets are generated by adding header information. Finally, a JPEG 2000 code stream of the input image is outputted.

Here, we should note that iterative operation of MQ encoding and the following procedures with varying the truncation points are required when the target bit-rate is given for rate controlling although the method is categorized into one-pass rate control. Iterative processing is required to fix the length of the generated code stream.

2.2 PCRD optimization algorithm

In JPEG2000 standard, in each code block, truncation points are determined to minimize the total distortion given a target bit-rate. The procedure of the PCRD optimization is given as the following:

(2.1) In each code block, the slopes of the R-D curves at the truncation points are calculated, and so that the maximum and the minimum slops are obtained. Then, the maximum is used as the upper limit and assumed to be the optimum slope. The minimum, on the other hand, is treated as the lower limit.

(2.2) The slope at the optimum truncation point is assumed to be the middle value of the upper and the lower limits. In the following steps, this assumed slope $\hat{\lambda}$ is used as a threshold value.

(2.3) In each code block, a candidate of the optimum truncation point is obtained as the point where the slope at a truncation point is smallest value greater than the threshold $\hat{\lambda}$.

(2.4) By performing the MQ encoder and the following procedures in the encoding process at the candidate of the optimum truncation point of (2.3), the bit-rate $\hat{R}_C$ of the code stream will be calculated.

(2.5) The calculated bit-rate $\hat{R}_C$ is compared with the target bit-rate R_t. If $\hat{R}_C \sim R_t$, then PCRD optimization process quits by setting the optimum slope as $\hat{\lambda}$. If $\hat{R}_C > R_t$ then the lower limit is set as the assumed slope, or in the case that $\hat{R}_C < R_t$ then the upper limit and the optimum slope are set as the assumed slope $\hat{\lambda}$.

(2.6) If values of the upper and lower limits are close enough, then PCRD optimization would quit. In other cases, iterate the procedure from (2.2) above.

In Fig.2, we show the flowchart of the PCRD-opt method. It should be noted that, when the target bit-rate is specified, the PCRD optimization process iterates the procedure (2.4) through (2.6) although the rate control of JPEG 2000 is generally called one-pass rate control.

3. PROPOSED METHOD

Here, we describe the proposed method.

3.1 Estimation of the image quality

In the proposed method, we use the MSE in the spatial domain as the distortion D. When we truncate a coding pass $p_i^{(p,k)}$ ($k \in \{SP, MR, CP\}$) in p-th ($p = 0, 1, \ldots$) bit plane of i-th code-block, the distortion $\Delta D_i(p,k)$ is given using DWT coefficients as the following equation [4][9][10]:

$$\Delta D_i^{(p,k)} = G_{b_i} \times \Delta_i^2 \times 2^{2p} \times \left(\sum_{j \in P_i^{(p,k)}, v_i^{(p)}[j]=1} T_s(\tilde{v}_i^{(p+1)}[j]) + \sum_{j \in P_i^{(p,k)}, v_i^{(p)}[j]>1} T_m(\tilde{v}_i^{(p+1)}[j]) \right) \quad (1)$$

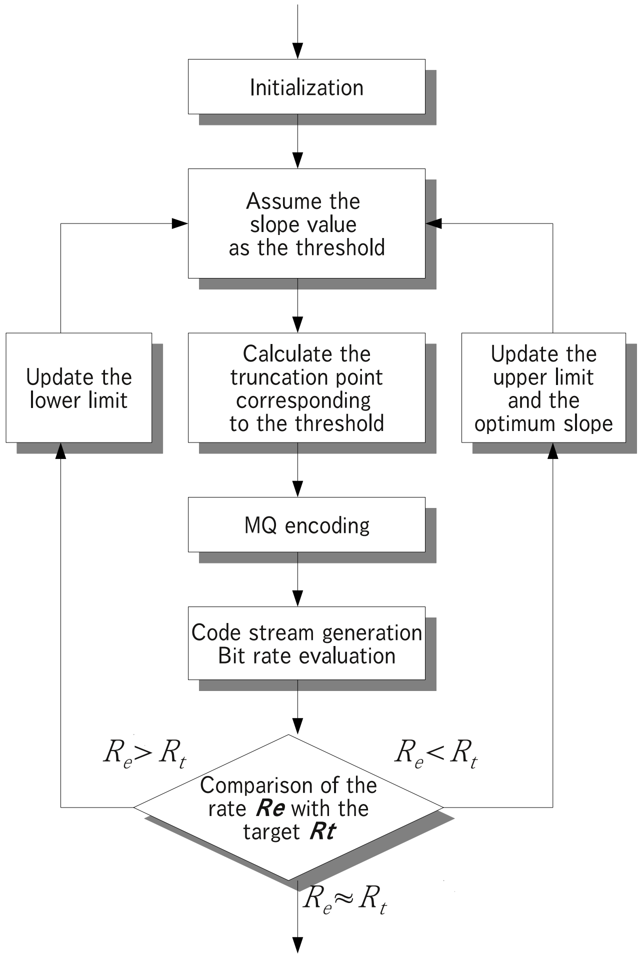


Fig. 2: Flowchart of the encoding process when the target bit-rate is specified.

where

$$T_s(\tilde{v}) = (2\tilde{v})^2 - (2\tilde{v} - 3/2)^2 \quad (2)$$

$$T_m(\tilde{v}) = (2\tilde{v} - 1)^2 - \left(2\tilde{v} - \lfloor 2\tilde{v}_i^{(p+1)}[j] \rfloor - 1/2\right)^2 \quad (3)$$

In this equation, Δ_i shows the quantization step-size of the i-th code block; $v_i^{(p)}[j]$ show the DWT coefficients after pre-quantization step; $\tilde{v}_i^{(p)}[j]$ shows the decimal part of the division of $v_i^{(p)}[j]$ by 2^p. Also, G_{b_i} shows the coding gain of inverse wavelet transformation in the subband where the code block b_i belongs. These values are calculated using the lowpass filter g_0 and the highpass filter g_1. We have the following expressions for 1-D and 2-D cases as

$$\begin{aligned} s_{L_1}[n] &= g_0[n], \quad s_{H_1}[n] = g_1[n] \\ s_{L_d}[n] &= \sum_k s_{L_{d-1}}[k] g_0[n-2k], d = 1,2,...,D \\ s_{H_d}[n] &= \sum_k s_{H_{d-1}}[k] g_0[n-2k], d = 1,2,...,D \end{aligned}$$

$$\begin{aligned} s_{LL_D}[n_1,n_2] = S_{L_D}[n_1] s_{L_D}[n_2] &\rightarrow G_{LL_D} = G_{L_D} \cdot G_{L_D} \\ s_{HL_d}[n_1,n_2] = S_{L_d}[n_1] s_{H_d}[n_2] &\rightarrow G_{HL_d} = G_{L_d} \cdot G_{H_d} \\ s_{LH_d}[n_1,n_2] = S_{H_d}[n_1] s_{L_d}[n_2] &\rightarrow G_{LH_d} = G_{H_d} \cdot G_{L_d} \\ s_{HH_d}[n_1,n_2] = S_{H_d}[n_1] s_{H_d}[n_2] &\rightarrow G_{HH_d} = G_{H_d} \cdot G_{H_d} \end{aligned}$$

where d ($d = 1, \ldots, D$) shows the level of DWT. Coding gain G_i of the code-block b_i is given as one of the above equations. These equations are derived under the assumption that the quantization noises are mutually uncorrelated.

3.2 PCRD optimization given a target MSE

Here, we describe the proposed procedure. In the proposed method, target is given either in MSE or PSNR. When the target is given in terms of PSNR, corresponding MSE value are calculated from the target. Then, the target MSE is added as a new constraint for the PCRD optimization process.

The procedure of the proposed method is given as follows:

(3.1) In each code block, the slopes of the R-D curve at effective truncation points are calculated and the maximum and the minimum slopes are obtained. Then, the maximum value is used as the upper limit. The minimum value is used as the lower limit and is assumed to be the optimum slope.

(3.2) We assumed that the slope at the optimum truncation point is given as the middle value of the upper and the lower limits. This slope $\hat{\lambda}$ will be treated as the threshold value in the following steps.

(3.3) In each code block, a candidate of the optimum truncation point under $\hat{\lambda}$ is obtained as z_i, where the slope becomes the smallest value greater than $\hat{\lambda}$. At the same time, the sum ΔD of the estimated MSE generated by the truncation is estimated by the following equation:

$$\Delta D = \sum_i \sum_{(p,k)<z_i} \Delta D_i^{(p,k)} \quad (4)$$

where $(p,k) < z_i$ shows the lower coding passes than the truncation point z_i.

(3.4) The bit-rate of the code stream is calculated by performing the MQ coding and the following procedures.

(3.5) The calculated bit-rate is compared with the target bit-rate. If they are almost equal then PCRD optimization quits by setting the optimum slope as the assumed slope. When the calculated bit-rate is greater than the target, we set the lower limit as the estimated slope $\hat{\lambda}$ and jump to the procedure (3.7) below.

(3.6) The MSE value calculated in (3.3) is compared with the target MSE. When the calculated MSE is greater than the target value, the upper limit is set as the assumed slope. Otherwise, the lower limit and the slope are set as the assumed slope.

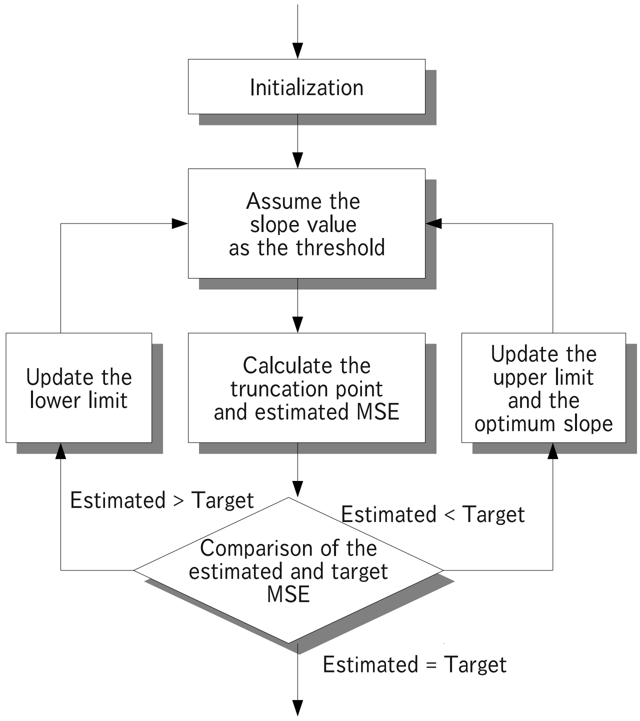


Fig. 3: Flowchart of the encoding process when the target MSE is specified.

(3.7) If the upper and the lower values are close enough then the PCRD optimization would quit. In other cases, repeat the procedure from (3.2).

Note that the above procedure is for the case in which both the target MSE and the maximum bit-rate are specified. We could omit the steps (3.4) and (3.5) when only the target MSE is given. The flowchart of the proposed method is shown in Fig.3. This chart shows the case where only the target MSE is given.

In the conventional optimization process of the previous section, the exact bit-rate is obtain in step (2.4) and it is compared with the target value in step (2.5). On the other hand, in the proposed method, an estimated MSE using equation (4) is used when comparing MSE values in (3.6) so that estimation error may be generated in this step. Especially, if all the pixels of the processed image have the same value then there exists small number of effective truncation points, and hence, noise components have non-zero cross-correlations. The correlated noise components affect the estimation of MSE. In this case, estimated MSE will not become smaller than the target MSE so that the optimum slope would never be updated.

In order to prevent this problem, we propose to initialize the optimum slope at the first step of the algorithm using the minimum value of the slopes calculated at the effective truncation points in all the code blocks. As a result of this, if the optimum slope is never updated then all the coding passes will be exist in the code stream, and hence, we could prevent the case that the obtained PSNR become much smaller than the target value.

4. SIMULATIONS

To verify the effectiveness of the proposed method, we evaluated MSE and file size of encoded images. For encoding, we implemented the proposed method into the JasPer encoder[11] and used it in the following simulations.

4.1 Accuracy of achieved MSE

We encoded still images with varying the value of the target MSE, and we compared the target values with the achieved MSE and PSNR of encoded images. The conditions of this simulation are shown in Table 1. In Fig.4(a)(b), we show the comparison of achieved MSE with the target. On the other hand, in Fig.4(c)(d) comparison of the achieved and the target values in terms of PSNR are shown. Note that the results of Fig.4 (b) and (c) are shown in the range from 30[dB] to 50[dB] in PSNR.

In Fig.4(a), it is shown that, in the region where the target values are relatively low, the PSNR of encoded image does not become lower than a certain value. This is due to the fact that many coding passes are truncated and almost all of the pixels of encoded image have a same value. Besides, from Fig.4(a), we can see that, as the target PSNR becomes higher, the achieved PSNRs of encoded image does not match with the target values. This disagreement is due to the errors generated in the DWT and quantization process. However, in Fig.4(c), it is shown that the MSE of the encoded image is very close to the target MSE in the region where the target value is small. Hence, we can say that the error is small in terms of the MSE in the region where the target PSNR is from 30[dB] to 50[dB].

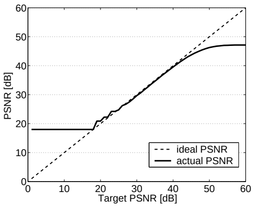
Fig.4(d) shows the comparison between the achieved and the target MSEs. From this figure, we can see that, when the target MSE is low, the MSE becomes large values, however when target MSE is over 30[dB], the error becomes small. We can say that, from these results, the proposed method enables us to achieve the image with the specified MSE with high degree of accuracy in practical PSNR (in 30[dB]to 50[dB]) region.

Table 1: Conditions of simulation 1.

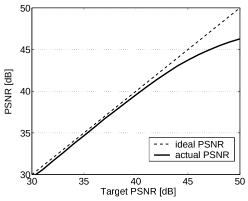
image	Lena
resolution	512(H) $\times$ 512(V) [pixels]
format	8 bit grayscale
DWT level	5 (9/7 filter)
codeblock size	64 $\times$ 64 [pixels]
target PSNR	1, 2, ..., 60 [dB]

4.2 Video coding specifying the target MSE and bit-rate simultaneously

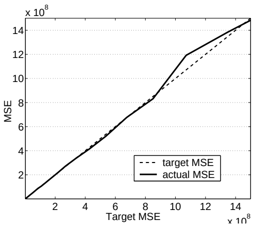
In this simulation, we encoded a video sequence with specifying the target MSE and the bit-rate simultaneously. The effectiveness of the proposed method is considered in terms of the estimated MSE, PSNR, file size, and encoding time. The conditions of simulation are summarized in Table 2.



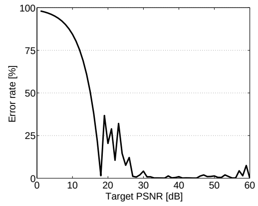
(a) accuracy of PSNR.



(b) closeup of (a).



(c) accuracy of MSE.



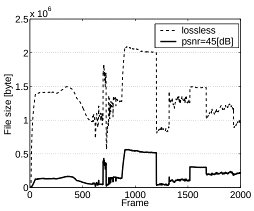
(d) error rate of MSE.

Fig. 4: the result of test 1.

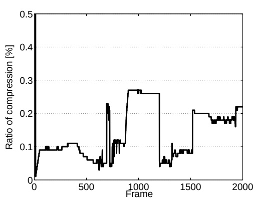
Table 2: Conditions of simulation 2.

condition	(a)	(b)	(c)
image	Elephants Dream		
resolution	1,920(H) × 1,080(V) [pixels]		
format	8 bit color		
number of frames	2,000 [frames]		
DWT level	5 (9/7filter)		
codeblock size	64 × 64 [pixels]		
target PSNR	45 [dB]	-	45 [dB]
target bit-rate	-	0.30 [bpp]	0.30 [bpp]

First, we encoded the video sequence with specifying MSE corresponding to 45[dB] in PSNR, and evaluated the file size of encoded data. The conditions are given as Table 2(a) and the results are shown in Fig.5. As a comparison, we show the file size of the video when compressed using JPEG2000 lossless mode in the figure. From the results in Fig.5 (a), we can confirm that the file sizes are significantly reduced by using the proposed method as compared with those by the lossless compression. Besides, in Fig.5 (b), we show the compression ratio of the file size of each frame, and it can be seen that the compression ratios are in the range from 5% to 28%, and the average is about 15%.



(a) Size of each frame

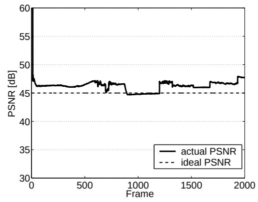


(b) Ratio of compression

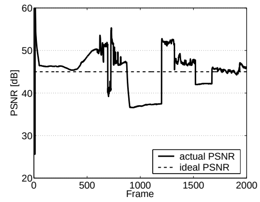
Fig. 5: The results of simulations when the target MSE (PSNR=45[dB]) was specified.

Next, in order to compare the proposed method with the standard PCRD-opt, we encoded a video sequence with specifying the average bit-rate, which was obtained in the simulation above, as the target bit-rate in the JasPar encoder. The conditions and results are shown in Table 2(b) and Fig.6. In Fig.6(a), the achieved PSNRs of each frame, when the target MSE is specified, are shown, and in Fig.6 (b), those of when the target bit-rate is specified are shown. Moreover, in Fig.6 (c), we showed the file sizes of each frame for both conditions.

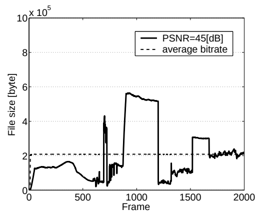
As shown in Fig.6(a), when specifying the target MSE, PSNR of frames are maintained as almost constant, however, the file sizes varies from frame to frame. On the other hand, as shown in Fig.6(b)(c), when specifying the target bit-rate, file size of frames become almost constant but PSNR varies from frame to frame.



(a) Proposed (specifying MSE)



(b) PCRD-opt (specifying bit-rate)



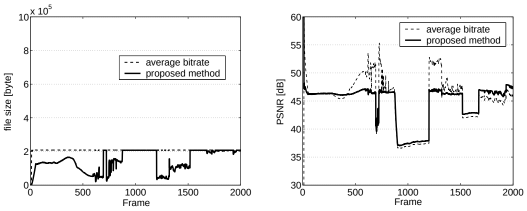
(c) File size of each frame

Fig. 6: Comparison of the results in the cases where the target MSE was specified and where the target bit-rate was specified.

Finally, we encoded the sequence with specifying both the target MSE and the target bit-rate simultaneously. The condition and results are shown in Table 2(c) and Fig.7 respectively. In Fig.7, for the comparison, the results using the PCRD-opt method are also shown.

In this simulation, as explained in Sec. 3.2, when the estimated MSE of a frame became lower than the target within the target bit-rate then we truncated following coding passes. In other cases, we selected the optimum truncation point of coding passes where the MSE becomes the smallest at the target bit-rate. Note that, in this case, the resultant MSE became lower than the target.

From the Fig.7 (c), we can confirm that there are no frame whose PSNR largely exceeded the target 45[dB] for those frames the proposed method can assign the required coding passes under the target bit-rate. Also, it is confirmed that the file sizes of frames are reduced from those obtained by the PCRD-opt method. Moreover, the PSNR of some frames obtained using the PCRD-opt reaches 50[dB] to 55[dB]. Consequently, the proposed method enables us to allocate smaller amount of bits to each frame while maintaining the desired quality.



(a) File size of each frame (b) PSNR of each frame

Fig. 7: Result of simulations when both the target MSE and the target bit-rate were specified in the proposed method.

Table 3: Encoding time.

	(a)	(b)	(c)
encoding time [sec]	1,927	2,061	2,047

We show the encoding time of each condition in Table 3. We note that we consider the proposed method under the condition (b) is identical to the standard PCRD-opt method. Hence, under this view point, encoding times of the proposed method are almost same as that of the standard PCRD optimization method. So we can conclude that increase of encoding time by the additional process of rate allocation is small.

5. CONCLUSION

In this paper, we proposed a novel one-pass rate control method for JPEG 2000. The target MSE, and additionally the maximum bit-rate if desired, can be specified in the proposed method. We described the proposed method as an extension to the PCRD optimization algorithm and noted that it requires no additional computation cost to be implemented.

Using the proposed method, we could maintain an almost constant MSE and could avoid the generation of low MSE images when we apply it to the compression of a video sequence. We confirmed the effectiveness of the proposed method through computer simulations and showed that the proposed method achieves the MSE (or the PSNR) very close to the specified target.

6. REFERENCES

[1] I.T.R.T. ISO/IEC 15444-1, “Information technology-JPEG 2000 image coding system-Part1: Core coding system,” ISO/IEC JTC 1/SC 29 WG1, Jan. 2001.

[2] Taubman, D., “High Performance Scalable Image Compression with EBCOT,” IEEE Trans. Image Processing, vol.9, no.7, pp.1151–1170, July 2000.

[3] Digital Cinema Initiatives, LLC Member Representatives Committee, “Digital cinema system specification V1.0,” Final Approval July 2005.

[4] D. Taubman and M. Marcellin, JPEG2000 Image Compression Fundamentals, Standards and Practice, Kluwer Academic Publishers, Boston, U.S.A., 2001.

[5] Taubman, D., Ordentlich, E., Weinberger, M. and Seroussi, G. (January 2002) ”Embedded block coding in JPEG 2000,” Signal Processing: Image Communication, volume 17 (number 1), pp. 49–72.

[6] J. C. Dagher, A. Bilgin, and M. W. Marcellin, “Resource-Constrained Rate Control for Motion JPEG2000,” IEEE Trans. Image Processing, vol. 12, no. 12, pp. 1522–1529 Dec. 2003.

[7] Z. Liu, L. J. Karam, A. B. Watson, “JPEG2000 Encoding With Perceptual Distortion Control,” IEEE Trans. Image Processing, vol. 15, no. 7, pp. 1763–1778, July 2006

[8] Y. M. Yeung, O. C. Au, “Efficient Rate Control for JPEG2000 Image Coding,” IEE Trans. on Circuits and Systems for Video Technology, vol. 15, no. 3, pp. 335–344, March 2005

[9] K.Nishikawa, S.Nagawara, and H.Kiya, “QoS Estimation Method for JPEG 2000 Coded Image at RTP Layer,” IEICE Trans. Fundamentals, vol.E89-A, no.8, pp.2119–2128, Aug. 2006.

[10] K.Nishikawa, K.Munadi, and H.Kiya, “No-Reference PSNR Estimation for Quality Monitoring of Motion JPEG2000 Video Over Lossy Packet Networks,” IEEE Trans. Multimedia, vol.10, Issue 4, pp.637–645, June 2008.

[11] M. D. Adams, JasPar Project Home Page, `http://www.ece.uvic.ca/mdadams/jasper2000`.

[12] ”Elephants dream,” `http://media.xiph.org/`.