EFFECTIVE RATE CONTROL METHOD FOR MINIMIZING TEMPORAL FLUCTUATIONS IN PICTURE QUALITY APPLICABLE FOR MPEG-4 AVC/H.264 ENCODING

Atsushi Matsumura, Sei Naito, Ryoichi Kawada and Atsushi Koike

KDDI R&D Laboratories Inc.

ABSTRACT

Appropriate rate control plays a very important role in encoding motion pictures under the constant bit-rate. One of the requirements for rate control is minimizing temporal fluctuations in picture quality, which may cause flicker artifacts. To satisfy the requirement, some methods for MPEG-1 and MPEG-2 have been proposed. However, these methods cannot be applied to the H.264 encoding algorithm based on R-D optimization since they require the DCT coefficients and the accurate output bit amount for each quantization parameter of a current picture in advance. To overcome the problem, we propose a rate control method to satisfy the requirement in the H.264 encoding algorithm by introducing approximations to estimate the bit amount and the distortion.

1. INTRODUCTION

In encoding motion pictures under the constant bit-rate, appropriate rate control plays a very important role. Therefore, several rate control methods that are applicable for H.264 have been proposed [1]-[4]. The goals of these methods are to reflect the encoding complexity of the current picture and to stabilize the occupancy of the coded picture buffer, called as CPB.

Meanwhile, there is a requirement for rate control to minimize temporal fluctuations in picture quality since it may cause flicker artifacts. However, the above methods cannot satisfy the requirement since they do not take quality differences of neighboring pictures into consideration. For the requirement, several methods for MPEG-1 and MPEG-2 have been proposed [5][6], and the effectiveness of the methods has been evaluated. Compared to the method proposed in [5] which needs to encode every picture in advance, the method proposed in [6] can be applied widely because only the information from a current picture and previous pictures is required to conduct the processes. However, since the method in [6] requires the DCT coefficients of the current picture in the processes to estimate the bit amount and the distortion, it cannot be applied to the general H.264 encoding algorithm based on R-D optimization.

Therefore, this paper proposes a method to satisfy the above requirement in the H.264 encoding algorithm by introducing approximations to estimate the bit amount and the distortion. The effectiveness of the proposed method is evaluated by coding experiments.

2. RATE CONTROL OF THE CONVENTIONAL METHOD AND ITS PROBLEM

As an example of the conventional rate control methods for H.264, the method adopted by JM9.2 [7], which is a reference software, is introduced below. In the method, the rate control is conducted by deciding an appropriate quantization parameter, called as QP.

In the P picture, the QP is updated by the basic-unit, constructed by one or several macroblocks. First, the bit amount for the basic-unit is decided from the following information.

- The target bit amount for the current picture.
- The used bit amount of the current picture.
- The complexities of the neighbor macroblocks.

Then the QP is decided to fulfill the bit amount. Therefore, the main purpose for deciding the QP for P pictures is fulfilling the target bit amounts.

Meanwhile, the QP for I pictures is decided from the average QP for the P pictures in the previous GOP, group of picture, and then clipping is conducted using the QP for the previous I picture. The QP for the B picture is decided by the QP for the previous reference picture, and then clipping is conducted using the QP for the previous B picture. Therefore, the QP for these pictures depends on the QP for the previous P pictures.

The method can fulfill the required encoding bit-rate. However, the method may cause temporal fluctuations in picture quality. As examples, PSNR results of Whale Show are shown in **Fig. 1**, where, the encoding bit-rate was defined as 1Mbps and initial QP was set as 25, which is the default value of JM9.2. From the result, it is shown that clipping I picture caused temporal fluctuations in picture quality

0-7803-9134-9/05/$20.00

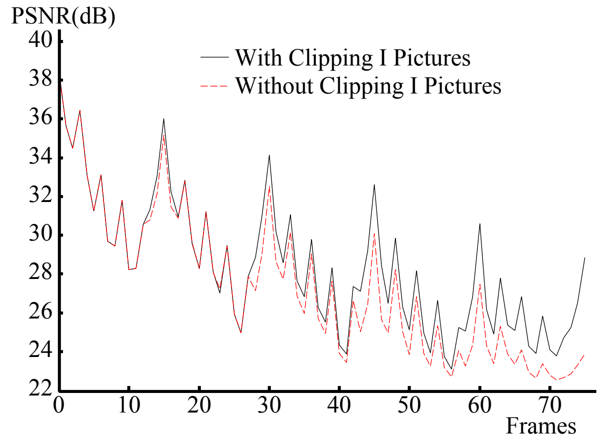

Fig. 1. Examples of temporal fluctuations of image quality.

between I pictures and the neighboring pictures. Moreover, a similar problem was also shown by the result without clipping I pictures since the QP for I pictures was decided by the average QP for P pictures in the previous GOP. To solve the problem, considering image quality between neighboring pictures is effective.

3. THE PROPOSED METHOD

To overcome the problem described in Section **2**, the method to minimize temporal fluctuations in picture quality for H.264 is proposed. The proposed method adopts the flow of the method proposed in [6], whose goal is to minimize the differences of distortions among GOP length pictures from the current picture. The flow is divided by the following four steps, initialization, definition of the target distortion, estimation of the bit amount, and decision of the QP for the current picture.

Since the method in [6] requires the DCT coefficients of the current picture to define the target distortion and estimate the bit amount, the method cannot be applied to the H.264 encoding algorithm based on R-D optimization. However, the problem can be overcome if these processes can be conducted without using the DCT coefficients. Therefore, the proposed method solves the problem by introducing the following approximations proposed in [8],

$$R(j, Q) = b_r(j) \times Q_{Step}(Q)^{a_r(j)}, \tag{1}$$

$$D(j, Q) = b_d(j) \times Q_{Step}(Q)^{a_d(j)}, \tag{2}$$

where, $R(j, Q)$ and $D(j, Q)$ indicate the bit amount of the DCT coefficients and the distortion of the j-th picture whose QP is Q respectively; $a_r(j)$, $b_r(j)$, $a_d(j)$ and $b_d(j)$ indicate the coefficients respectively; and $Q_{Step}(Q)$ indicates the Q-Step value calculated from Q defined by H.264.

Moreover, since the picture quality of non-reference pictures depends on those of reference pictures, increasing the picture quality of reference images relatively leads to increasing the picture quality of a whole sequence. Thus, the ratio of the distortions between reference pictures and non-reference pictures is defined and it is applied to decide the target distortion. Details of the proposed method are as follows.

3.1. Deciding the target bit amount

The target bit amount of N_G pictures from the current picture, defined as R_G, is decided based on the method proposed in [9], where, N_G indicates GOP length. In the first picture, R_G is calculated by the following equation,

$$R_G = N_G \times \frac{bit\_rate}{frame\_rate} + v_0 - v_T, \tag{3}$$

where, v_0 and v_T indicate CPB occupancy when picture data extraction commences and the target CPB occupancy, respectively. Then, R_G is updated after encoding the picture using the following equation,

$$R_G \leftarrow R_G + R(i) - S(i), \tag{4}$$

where, i, $\leftarrow$, $R(i)$ and $S(i)$ indicate the number of the current picture, substituting right values for left values, the target bit amount and the used bit amount in the current picture, respectively.

3.2. Deciding the QP for the current picture

The QP for reference pictures is defined as follows.

Step 1: *Initialization*

First, define $min(i)$ and $max(i)$, indicating the range of candidate QP for the i-th picture, using the following equations,

$$min(i) = Q''(i - N_G, F) - CL, \tag{5}$$

$$max(i) = Q''(i - N_G, F) + CL, \tag{6}$$

where, F, $Q''(i - N_G, F)$, and CL indicate, respectively, the number of the previous reference picture, the estimated QP whose picture number is $i - N_G$ and distortion is that of the F-th picture, and the constant value used for clipping. Secondly, encode the i-th picture using $min(i)$, $max(i)$ and $Q''(i - N_G, F)$ as the QP, respectively, and then calculate $a_r(i)$, $b_r(i)$, $a_d(i)$ and $b_d(i)$ by (1) and (2). Finally, define a candidate QP, Q_C, as $min(i)$.

Step 2: *Definition of the target distortion*
Define the target distortion of the reference pictures and the non-references pictures by (7),

$$D(k, Q_k) = \begin{cases} D(i, Q_C) & \text{if the } k\text{-th picture is a reference picture} \\ D(i, Q_C)/w_b & \text{otherwise} \end{cases}, \tag{7}$$

where, w_b indicates the ratio of distortions between reference pictures and non-reference pictures.

Step 3: *Estimation of the bit amount*
Estimate the bit amount of N_G pictures from the i-th picture using the following equation,

$$R'(i, Q_C) = R(i, Q_C) + \sum_{k=i+1}^{i+N_G-1} R(k, Q_k), \tag{8}$$

where, $R(k, Q_k)$ is calculated by (9), derived by (1) and (2), and the coefficients are defined by **Table 1**, where, $L(k)$ indicates the previous picture whose picture type is the same as the k-th picture,

$$R(k, Q_k) = b_r(k) \times \exp\left(\frac{\log\left(D(k, Q_k)/b_d(k)\right)}{a_d(k)}\right)^{a_r(k)}. \tag{9}$$

Step 4: *Decision of the QP for the current picture*
If $Q_C < max(i)$, increase Q_C and go back to *Step 2*, otherwise define the QP for the i-th picture, $Q(i)$, as Q_C which minimizes $|R_G - R'(i, Q_C)|$.

The QP for the non-reference picture is decided as Q_C that minimizes the following equation,

$$|D(i, Q_C) - P(F, i, F')/w_b|, \tag{10}$$

where, $P(F, i, F')$ is defined by the following equation,

$$P(F, i, F') = D(F, Q(F)) + \frac{D(F', Q(F')) - D(F, Q(F))}{F' - F} \times (i - F), \tag{11}$$

where, F' indicates the number of the following reference picture.

Table 1. The coefficients used to estimate the bit amount.

if $k > N_G$ and $D(k - N_G, min(k - N_G)) \leq D(k, Q_k) \leq D(k - N_G, max(k - N_G))$	otherwise
$a_r(k) = a_r(k - N_G)$, $b_r(k) = b_r(k - N_G)$, $a_d(k) = a_d(k - N_G)$, $b_d(k) = b_d(k - N_G)$.	$a_r(k) = a_r(L(k))$, $b_r(k) = b_r(L(k))$, $a_d(k) = a_d(L(k))$, $b_d(k) = b_d(L(k))$.

4. EXPERIMENTAL RESULTS

To evaluate the effectiveness of the proposed method, coding experiments were conducted. In the experiments, five materials in SIF standard sequences are used. The test conditions for encoding are given in **Table 2**. Every macroblock in a picture was encoded by the same QP in the proposed method.

The PSNR results are shown in **Fig. 2**. As a comparison, the results by JM9.2 are also shown. From the result, it was shown that the proposed method made the fluctuations of PSNR smaller compared to JM9.2 in both sequences.

Next, CPB occupancy results are shown in **Fig. 3**. From the result, an overflow occurred in case of Flower Garden by JM9.2. Overflows also occurred in cases of Cognac and Fruit, and Woman with Bird Cage by JM9.2. Meanwhile, neither overflows nor underflows occurred using the proposed method.

Finally, the variances of mean square error, called as MSE, are shown in **Table 4**. Large variances of MSE tend to cause the degradation of image quality. In addition, the PSNR is also shown. From the result, the proposed method decreased the variances of MSE while maintaining the similar or better PSNR compared to JM9.2 in all test sequences.

5. CONCLUSIONS

This paper proposed an effective rate control method for H.264 to minimize temporal fluctuations in picture quality. The proposed method achieved this in the H.264 encoding algorithm, which was difficult using conventional methods, by introducing approximations to estimate the bit amount and the distortion. The experimental results showed the proposed method made fluctuations in picture quality and MSE variances smaller compared to the conventional method.

ACKNOWLEDGEMENT

This study was conducted by the reseach project of NiCT (National Institute of Information and Communications Technology) regarding the software encoding technologies.

Table 2. The test conditions for the experiments.

Conditions	Values
CL	3
wb	0.9
Initial QP	39
Profile	Main Profile
Level IDC	3.2
Encoding bit-rate	1Mbps
GOP Structure	IBBPBBPBBPBBPBB

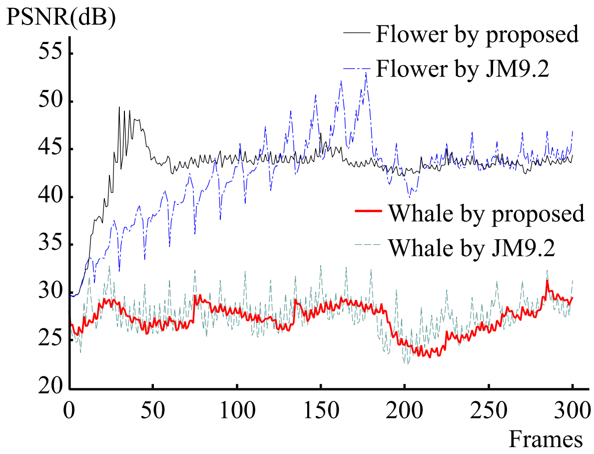

Fig. 2. The transitions of PSNR.

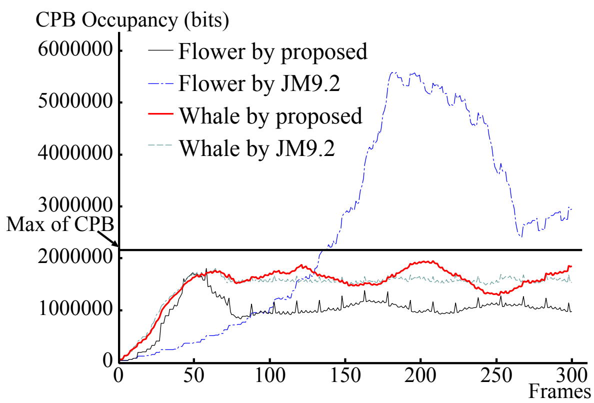

Fig. 3. The transitions of CPB occupancy.

Table 3. Variances of MSE and PSNR.

Materials		Variances of MSE	PSNR
Cognac and Fruit	JM9.2	94.87	39.84
	Proposed	82.13	41.07
Flower Basket	JM9.2	163.59	39.28
	Proposed	133.27	40.81
Woman with Bird Cage	JM9.2	67.13	38.80
	Proposed	66.14	39.34
Whale Show	JM9.2	3202.80	27.23
	Proposed	2609.14	26.99
Green Leaves	JM9.2	4609.03	24.75
	Proposed	1512.29	24.62

6. REFERENCES

[1] Minqiang Jiang, Xiaoquan Yi and Nam Ling, "Improved frame-layer rate control for H.264 using MAD ratio," IEEE International Symposium on Circuits and Systems, Vol. 3, pp. III-813–816. May 2004.

[2] Jianfeng Xu and Yun He, "A novel rate control for H.264," IEEE International Symposium on Circuits and Systems, Vol. 3, pp. III-809-12, May. 2003.

[3] Hongtao Yu, Feng Pan and Zhiping Lin "A new bit estimation scheme for H.264 rate control," IEEE International Symposium on Consumer Electronics, pp. 396 - 399 Sept. 2004.

[4] Xiaoquan Yi and Nam Ling, "Rate Control Using Enhanced Frame Complexity Measure for H.264 Video," IEEE Workshop on Signal Processing Systems, pp. 263–268. Oct. 2004.

[5] Dzung T. Hoang, Elliot L. Linzer and Jeffrey Scott Vitter, "Lexicographic Bit Allocation for MPEG Video," Journal of Visual Communication and Image Representation, Vol. 8, No. 4, pp. 384–404, Dec. 1997.

[6] Sung-Hoon Hong, Sang-Yo Yoo, Si-Woong Lee, Hyun-Soo Kang and Sung Yong Hong, "Rate Control of MPEG Video for Consistent Picture Quality," IEEE Trans. on Broadcasting, Vol. 49, No. 1, pp. 1–13, Mar. 2003.

[7] Karsten Sühring, H.264/AVC Software Coordination http://bs.hhi.de/ suehring/tml/download/

[8] Yucel Altunbasak and Nejat Kamaci, "An analysis of the DCT coefficient distribution with the H.264 video coder," IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004, vol.3, pp.iii-177–180, May 2004.

[9] Sei Naito, Atsushi Koike, Masahiro Wada, Shuichi Matsumoto, and Yoshinori Hatori, "Optimization of MPEG-2 Bit Assignment Scheme with Full Use of Distortion Minimization Oriented Rate Control Technology," IEICE Trans. D-II, Vol. J86-D-II, No. 11, pp.1565–1574, Jul. 2003.(in Japanese)