Target Bit Matching for MPEG-2 Video Rate Control

Jeong-Woo Lee and Yo-Sung Ho

Kwangju Institute of Science and Technology
1 Oryong-Dong Puk-Gu, Kwangju, 500-712, Korea
E-mail: jwlee@gogh.kjist.ac.kr

ABSTRACT
The MPEG-2 Test Model 5 (TM5) algorithm is widely used for bit rate control. In TM5, however, the target number of bits and the number of actual coding bits for each picture do not match well. Therefore, degradation of picture quality and the buffer overflow may occur at the end of the GOP. In this paper, we propose a new bit rate control algorithm based on accurate bit estimation. The main idea of the proposed algorithm is to determine quantization parameters that generate the number of actual coding bits close to the target number of bits for each picture, while maintaining uniform picture quality. The proposed algorithm uses the relationship between the number actual coding bits and the number of estimated bits of the previous macroblock.

1. INTRODUCTION
For the bit rate control in MPEG-2, we estimate a reference quantization parameter before we adjust the actual quantization step size during macroblock (MB) quantization. However, if we examine the reference and the actual quantization step sizes for each MB in the MPEG-2 Test Model 5 (TM5)[1], two parameter values are often quite different because they are not correlated closely to each other in terms of the bit rate. In other words, the target number of bits and the number of actual coding bits of the picture do not match well. Consequently, if more coding bits than the allocated target number are exhausted, the remaining bit resource in the GOP is getting smaller. In such a case, insufficient coding bits can be allocated to the pictures at the end of the GOP, which may result in severe degradation of picture quality or buffer overflow.

If the number of actual coding bits is similar to the target number of bits for each picture, we can prevent buffer overflow and buffer underflow. In order to obtain uniform picture quality within each picture, we should select an appropriate coding parameter for each MB. Based on the properties of the MB, we should set a strategy to generate the number of actual coding bits close to the target number of bits for the picture.

This paper proposes a new algorithm for finding a good reference quantization parameter that can produce uniform picture quality by minimizing the difference between the target number of bits and the number of actual coding bits based on accurate bit estimation.

2. TM5 RATE CONTROL ALGORITHM
The MPEG-2 TM5 rate control scheme works in the following three steps.

2.1 Bit Allocation
After encoding each picture, a global complexity measure, $X_i$, $X_p$ or $X_b$, is updated as follows:

$$X_{(i,p,b)} = S_{(i,p,b)} \times Q_{(i,p,b)}$$

(1)

where $S_i$, $S_p$ and $S_b$ are the number of bits generated by encoding a picture type I, P and B, respectively, and $Q_i$, $Q_p$ and $Q_b$ are their average quantization parameters over all the macroblocks in the picture. A target number of bits, $T_i$, $T_p$ or $T_b$, is then assigned to the next picture of the same type in the GOP as

$$T_i = \max \left\{ \frac{R}{1 + \frac{N_i X_p}{X_i K_p} + \frac{N_b X_b}{X_b K_b}}, 8 \times \text{picture rate} \right\}$$

(2)

$$T_p = \max \left\{ \frac{R}{N_p + \frac{N_b K_p X_b}{K_p X_p}}, 8 \times \text{picture rate} \right\}$$

(3)

$$T_b = \max \left\{ \frac{R}{N_b + \frac{N_b K_p X_b}{K_p X_b}}, 8 \times \text{picture rate} \right\}$$

(4)

where $K_p$ and $K_b$ are constant parameters, $R$ is the remaining number of bits assigned to the current GOP, $N_p$ and $N_b$ are the number of P and B pictures remaining in the current GOP, respectively. $\text{bit rate}$ is the rate at which the coded bitstream is delivered from encoder to decoder, and $\text{picture rate}$ is the frame rate at which pictures are reconstructed from the decoding process. Before encoding each GOP, the total bit resource $R$ is updated as

$$R = R + \frac{N \times \text{bit rate}}{\text{picture rate}}$$

(5)

where $N$ is the number of pictures in the GOP.

2.2 Rate Control
A reference quantization step size $Q$, which is used for the bit rate control, is computed for each macroblock based on the status of buffer fullness.
Before encoding each macroblock \( j \), the fullness of the appropriate virtual buffer of the picture type, I, P, or B, is first calculated:

\[
d_j^{i, p, b} = d_0^{i, p, b} + B_{j-1} - \frac{T_{i, p, b} \times (j-1)}{MB_{cnt}}
\]

(6)

where \( d_0^{i, p, b} \) and \( d_b^{i, p, b} \) are the initial buffer fullness for each corresponding picture type, \( B_{j-1} \) is the number of bits generated by encoding all the macroblocks in the picture up to \( j-1 \), and \( MB_{cnt} \) is the total number of macroblocks in the picture. The buffer fullness \( d_j^{i, p, b} \) is then fed back and used to adjust the quantization step size for macroblock \( j \) as:

\[
Q_j = \frac{d_j \times 31 \times \text{picture rate}}{2 \times \text{bit rate}}
\]

(7)

3 Adaptive Quantization

The spatial activity measure for macroblock \( j \) is first calculated from the four luminance frame-based subblocks and the four luminance field-based subblocks using the original pixel values:

\[
act_j = 1 + \min_{n=1,8} (\text{var}_{\text{blk}})
\]

(8)

\[
\text{var}_{\text{blk}} = \frac{1}{64} \sum_{i=1}^{64} (P_k - P_{\text{mean}})^2
\]

(9)

\[
P_{\text{mean}} = \frac{1}{64} \sum_{i=1}^{64} P_k
\]

(10)

where \( P_k \) are the pixel values in the original 8x8 block.

The normalized macroblock local activity \( N_{\text{act}} \) is defined as:

\[
N_{\text{act}} = \frac{2 \times act_j + \text{avg}_{\text{act}}}{act_j + 2 \times \text{avg}_{\text{act}}}
\]

(11)

where \( \text{avg}_{\text{act}} \) is the average spatial activity measure for the last coded picture. The quantization parameter for macroblock \( j \) is then modulated as:

\[
M_{\text{quant}} = Q_j \times N_{\text{act}}
\]

(12)

3.3 A PROPOSED ALGORITHM

3.1 Target Bit Allocation

The target number of bits for the picture \( T \) is computed according to the method described in the TMS [Eq. (2), Eq. (3), or Eq. (4)].

3.2 Reference Quantization Parameter

For estimating the reference quantization parameter for each macroblock, we define the following equation based on the rate distortion theory [2]:

\[
q_i = 2^C \times \gamma_i
\]

(13)

where \( C \) is a parameter that controls the bit rate, and \( \gamma_i \) is a scaling factor which characterizes the properties of the current macroblock. In order to calculate the reference quantization parameter, we should determine the value of \( C \) and \( \gamma_i \).

We may use \( N_{\text{act}} \) in Eq. (11) as the scaling factor \( \gamma_i \) for macroblock \( i \). However, since a good measure of the human visual sensitivity is the power of AC coefficients normalized by the DC value [3], we can define the scaling factor \( \gamma_i \) as:

\[
\gamma_i = \sqrt{\frac{\sum_{j=0}^{3} \sum_{k=0}^{3} \max(\text{DC}, \text{DC}_{\min}) \times \text{dct}_{j,k}}{232}}
\]

(14)

In order to determine the value of \( C \) for each frame from Eq. (13), we need to pick a scaling factor \( \gamma \) and a quantization parameter \( q \). We first calculate the average value of the scaling factors.

\[
\Gamma = \sum_{i=1}^{MB_{\text{cnt}}} \gamma_i
\]

(15)

\[
\gamma_{\text{avg}} = \frac{\Gamma}{MB_{\text{cnt}}}
\]

(16)

We then select a reference macroblock that has the average scaling factor \( \gamma_{\text{avg}} \). Since the reference macroblock should characterize the coded picture, we choose a \( \text{MB}_{\text{INTRA}, \text{MB}_{\text{FORWARD}}, \text{MB}_{\text{FORWARD}}, \text{MB}_{\text{BACKWARD}}} \) coded macroblock for I, P, and B pictures, respectively.

While encoding the reference macroblock, we adjust the initial quantization parameter \( Q_{\text{init}} \) such that the number of actual coding bits is close to the average number of coding bits \( B_{\text{avg}} \) for the macroblock.

\[
B_{\text{avg}} = \frac{T}{MB_{\text{cnt}}}
\]

(17)

The value of \( C \) is then calculated from Eq. (13), that is,

\[
C = \log_2 \frac{Q_{\text{init}}}{\gamma_{\text{avg}}}
\]

(18)

3.3 Target Bit Matching for Adaptive Quantization

In order to generate the number of actual coding bits close to the target number of bits for a picture, we can first distribute the target number of bits for the picture into macroblocks, and then generate the number of actual coding bits close to the target number of bits for each macroblock.

For this operation, we can exploit the relationship between the number of actual coding bits, \( \text{BIT}_{\text{actual}} \), and the number of estimated coding bits, \( \text{BIT}_{\text{estimated}} \), of the previous macroblock. \( \text{BIT}_{\text{actual}} \) and \( \text{BIT}_{\text{estimated}} \) represent the number of actual coding bits and the number of allocated bits for the previous macroblock, respectively.

If the activity of the current macroblock is large, it implies that the corresponding region is fairly complex. Therefore, if the quantization step size is fixed for all macroblocks, more coding bits will be generated in the macroblock of more activity. Here, we can assume that the number of the generated bits is linearly proportional to the activity of the coded macroblock.
In P and B pictures, the macroblocks coded in the predictive mode would generate a smaller number of coding bits than the number of bits estimated only by the macroblock activity. For a more accurate estimation of coding bits for the current macroblock, we examine the number of actual coding bits and the number of estimated bits for the previously encoded macroblocks that have both the same coding mode and the same activity as the current macroblock.

If we denote $\alpha$ as the ratio between the sum of $BIT_{estimated}$ and the sum of $BIT_{actual}$ of all the previous macroblocks that have the same coding mode and the same activity, we calculate the number of estimated bits $BIT_{estimated}$ of the current macroblock as

$$BIT_{estimated} = \alpha \cdot \beta \cdot \gamma \cdot \frac{T}{\Gamma}$$  \hspace{1cm} (19)$$

where $\beta$ is a constant which characterizes the coding mode of the previous macroblock. If the coding mode of the macroblock is MB_INTRA, $\beta$ equals to 1; we set $\beta = 0.8$ for the predictive coded macroblocks.

Fig. 1 explains a procedure of determining the quantization step size adaptively.

![Figure 1: Adaptive Mechanism for Quantization Step Size](image)

In Fig. 1, we define a difference $D$ as

$$D = BIT_{actual} - BIT_{estimated}$$  \hspace{1cm} (20)$$

If $D$ is positive, we increase the reference quantization parameter of the current macroblock since we have generated more bits than the number of allocated bits. Furthermore, we should determine how much quantization parameter should be adjusted in the current macroblock.

The quantization parameter $Q_i$ of macroblock $i$ is calculated as

$$Q_i = Q_{ref} + \Delta$$  \hspace{1cm} (21)$$

where $Q_{ref}$ is the reference quantization parameter of the current macroblock, and $\Delta$ is the amount of quantization step size to be adjusted. If the value of $\Delta$ is large, the quantization parameter is changed abruptly and the reconstructed macroblock may not have uniform quality. Therefore, the quantization parameter should be changed over several macroblocks in this situation.

As shown in Fig. 1, the value of $\Delta$ is determined based on the difference $D$. The rate control is performed as follows [5].

$$B = \frac{M}{quantization \ step \ size}$$  \hspace{1cm} (22)$$

where $B$ is the number of coding bits generated in the picture, and $M$ is a constant that is a function of picture complexity. Therefore, we can define a constant $M$ as

$$M = B_{avg} \cdot Q_{min} = B_{ref} \cdot Q_{ref}$$  \hspace{1cm} (23)$$

where $B_{ref}$ is the target number of bits for a given reference quantization parameter $Q_{ref}$. The number of coding bits allocated for a given difference is

$$B = B_{ref} - D_i$$  \hspace{1cm} (24)$$

where $D_i$ is the difference $D$ of macroblock $i$.

If we define

$$D_i = \eta \cdot B_{ref}$$  \hspace{1cm} (25)$$

Then,

$$B = B_{ref} - D_i = \frac{M}{Q_{ref} + \Delta}$$  \hspace{1cm} (26)$$

We obtain

$$\Delta = \frac{\eta \cdot Q_{ref}}{1 - \eta}$$  \hspace{1cm} (27)$$

where $\eta_{min} \leq \eta < 1$.

4. EXPERIMENTAL RESULTS

We have tested the performance of our proposed algorithm with two sample sequences, BUS of 60 frames and BALLET of 38 frames. The former contains fairly complex regions and fast moving objects. The latter includes only simple regions and slowly moving objects. They are coded at 6 Mbits/s using the proposed algorithm and the TM5 algorithm. The number of pictures in the GOP is 12 and the distance between I and P picture is 3.

Fig. 2(a) and Fig. 3(a) plot the differences between the target number of bits and the number of actual coding bits of BUS and BALLET sequences, respectively. From these figures, we can see that the actual coding bits are well matched to the target bits in our proposed algorithm, while there are large fluctuations in the TM5 algorithm. Therefore, we have not experienced any buffer overflow or underflow with the proposed algorithm.

Fig. 2(b) and Fig. 3(b) show their corresponding PSNR values of the two reconstructed sequences. From those figures, we can observe that the new algorithm performs better than the TM5 algorithm for the complex and fast-moving sequence, while both algorithms yield similar performances for the simple and slowly-moving sequence.

In order to obtain uniform picture quality, we may employ an iterative rate control algorithm. However, iterative methods generally have a problem of high computational complexity. The computational requirement of the proposed algorithm is low because the parameter $C$ is determined only from one reference macroblock at the beginning of each picture. We observe that the computational complexity of the proposed algorithm is similar to that of TM5 because we just add a calculation of $C$ on TM5.
5. CONCLUSIONS

In this paper, we propose a new bit rate control algorithm for MPEG-2 video coding based on accurate bit estimation. The key idea of the proposed algorithm is to determine a quantization parameter that can generate the number of actual coding bits close to the target number of bits, while maintaining uniform picture quality. When we estimate the number of coding bits for the current macroblock, we consider both the activity and the coding mode of the previous macroblock. Since the target number of bits and the number of actual coding bits for a picture match well to each other in our proposed algorithm, we have not experienced neither buffer overflow nor severe degradation of the picture quality at the end of GOP. With the proposed algorithm, we have obtained higher PSNR values for complex and fast-moving sequence. The proposed algorithm works fairly well even in the non-uniform quantization mode. Although the proposed algorithm includes an iterative procedure, the computational complexity of the proposed algorithm is similar to that of TM5.

ACKNOWLEDGMENTS

This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Ultra-Fast Fiber-Optic Networks (UFON) Research Center at Kwangju Institute of Science and Technology (K-JIST).

REFERENCES