# Improved H.264 frame-layer rate control for low bit rate video with high motion

**Changhyun Lee,[a] Seongjoo Lee,[b] Yunje Oh,[c] and Jaeseok Kim[a]**

[a]Yonsei University, Department of Electrical and Electronic Engineering, 134 Sinchon-Dong, Seodaemun-Gu, Seoul, 120-749, Korea
[b]Sejong University, Information and Communication Engineering, 98 Gunja-Dong, Gwangjin-Gu, Seoul, 143-747, Korea
[c]Samsung Electronics Co., Ltd., Multimedia Lab, 416 Maetan-3Dong, YoungTong-Gu Suwon, Gyeonggi-Do 442-742, Korea
E-mail: hyun2@yonsei.ac.kr

**Abstract.** In the existing frame-layer rate control for H.264, buffer status and content complexity are used improperly, causing quality fluctuations of high-motion video at low bit rates and high frame rates (under 19.2 kbps at 30 fps). We propose an improved H.264 frame-layer rate control scheme to obtain steady video quality and stable buffer management. Experimental results showed that the proposed scheme performed better than existing schemes. © *2006 Society of Photo-Optical Instrumentation Engineers.*
[DOI: 10.1117/1.2401154]

## 1 Introduction

Rate control is important for rate-constrained video coding. Good rate control results in high video quality, low fluctuation of video quality, stable buffer status, and low mismatch between the target bit rate and the encoded bit rate. H.264/AVC is the state-of-the-art video coding standard,[1] and its existing rate control is based on JVT.[2] According to Refs. 2 and 3, rate control generates a quantization parameter (QP) after three steps: estimation of the number of target bits (target-bit), computation of a QP, and adjustment of the QP. Because the JVT method does not consider the content complexity within a sequence as a target-bit is estimated, it cannot perform optimally for high-motion video. In Ref. 4, the quality fluctuations of high-motion video are improved by multiplying one part of the target-bit estimation by the relative content complexity (RCC). The RCC is measured by the ratio of the $i$'th predicted mean absolute difference (MAD) to the averaged actual MAD over all $P$ (forward predicted) frames encoded previously. However, that scheme still suffers from unstable buffer management and high mismatch between bit rates for two reasons. First, when the RCC is high, the target-bit is increased regardless of the buffer status. Second, a QP obtained by the modified target-bit may not appropriately maintain a stable ratio of buffer occupation. Therefore, both buffer status and RCC
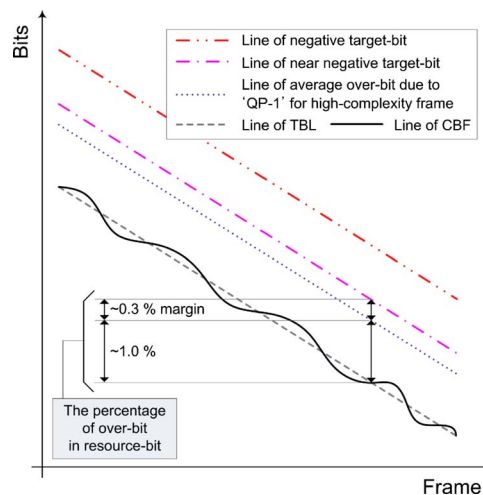

**Fig. 1** The effect of over-bit with respect to negative target-bit.

must be considered simultaneously, and it is reasonable to use them in the adjustment of the QP rather than in the estimation of a target-bit.

When the number of available bits per frame (avail-bit) is low (e.g., low target-bit rate and high target-frame rate) and the RCC is sequentially high, the probability of a target-bit dropping below zero is very high. In such cases, the QP for a current frame is forced to be larger than that of the previous frame by two,[2–4] producing poor video quality. When target-bits are frequently negative, video quality will fluctuate severely. Thus, it is important to prevent negative target-bits to avoid such fluctuations.

Based on the preceding observations, QPs were adjusted using buffer status and RCC to maintain positive target-bits for a low avail-bit application and high-motion video.

## 2 Proposed Scheme

It was assumed, without loss of generality, that the GOP structure was an $IPPP, \ldots, P$, where $I$ and $P$ denote an intracoded picture and a forward-predicted picture, respectively. The ultimate aim of rate control is to obtain an appropriate QP for high performance within permitted rates. Rate control is composed of three steps: estimation of a frame target-bit, computation of a QP, and adjustment of the QP.
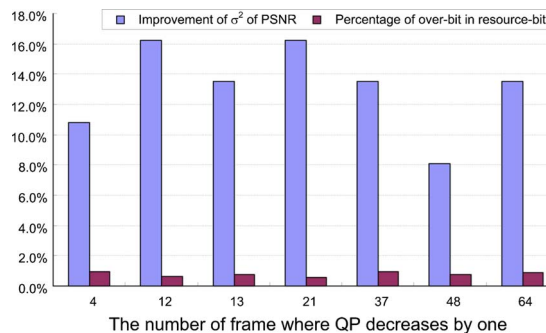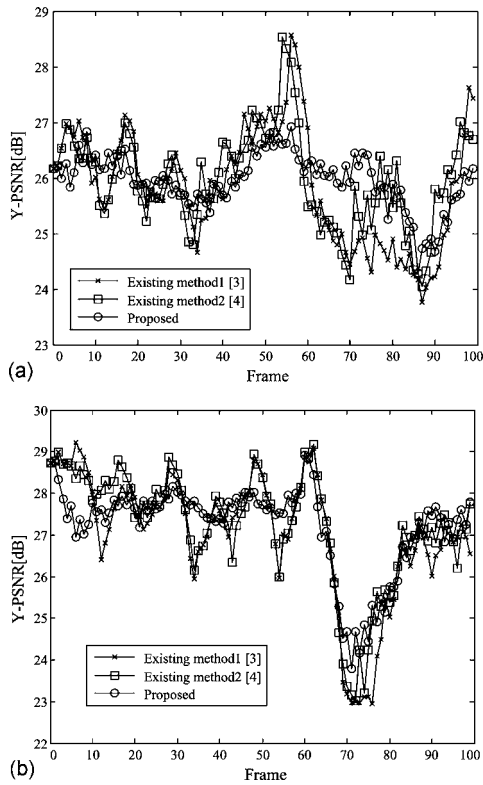

**Fig. 2** The percentage of over-bit in resource-bit and the improvement of fluctuation in video quality where the constant QP is 41 over 100 frames except for only one high-complexity frame that is quantized by 40.

**Fig. 3** PSNR curves for sequences (a) "Carphone" at 9.6 kbps/30 fps/1 reference picture, and (b) "Foreman" at 19.2 kbps/30 fps/1 reference picture.



**Fig. 4** Current buffer fullness curves for sequences (a) "Carphone" at 9.6 kbps/30 fps/1 reference picture, and (b) "Foreman" at 19.2 kbps/30 fps/1 reference picture.

To estimate a frame target-bit, the number of remaining bits is needed and both buffer fullness and the target buffer level are used to avoid overflow or underflow. According to Refs. 2 and 3, the target-bit $T_{b,i}$ is estimated before encoding the $i$'th frame:

$$T_{b,i} = \beta \frac{R_{b,i}}{N_{Pr,i}} + (1-\beta)\left[\frac{b_r}{f_r} - \Gamma(CBF_{i-1} - TBL_i)\right], \qquad (1)$$
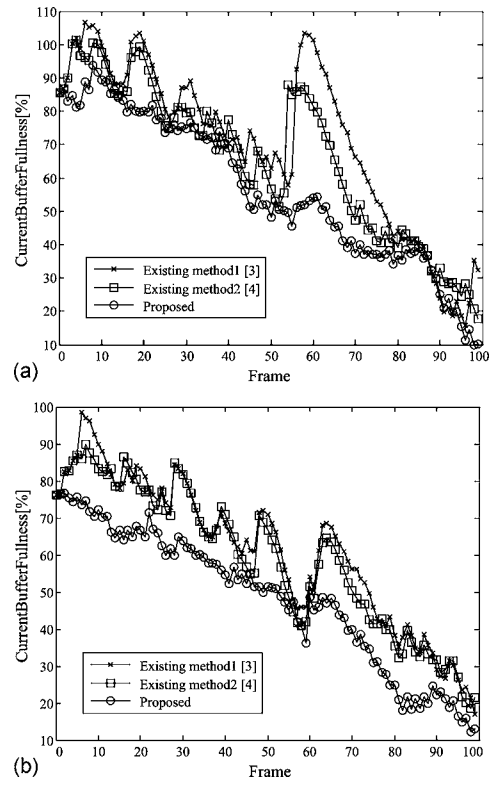
where $R_{b,i}$ and $N_{Pr,i}$ denote the total number of remaining bits for all noncoded $P$ frames and the number of noncoded $P$ frames for the $i$'th frame, respectively, and $b_r$ and $f_r$ are the target-bit rate and frame rate, respectively. $CBF_i$ and $TBL_i$ indicate the current buffer fullness and the target buffer level for the $i$'th frame. $\beta$ and $\Gamma$ are constants with typical values of 0.5 and 0.75, respectively. After estimating a target-bit, a QP was computed and adjusted in the two following cases.

## 3 Case 1: Positive Target-Bit

When the target-bit was positive, the QP was computed using the quadratic rate-distortion (R-D) model[5]:

$$\frac{T_{tb,i}}{PMAD_i} = \frac{x_1}{Q_{c,i}} + \frac{x_2}{Q_{c,i}^2}, \qquad (2)$$

where $T_{tb,i}$ is the texture-bit for the $i$'th frame. $T_{tb,i}$ shows the difference between the target-bit and the number of previously encoded header bits (header-bit). $PMAD_i$ and $Q_{c,i}$ denote the predicted MAD and the computed QP for the $i$'th frame, respectively, and $x_1$ and $x_2$ are the first- and

second-order coefficients, respectively. The texture-bit can fall below zero when the previously encoded header-bit is comparable to the size of the current target-bit in a low avail-bit application. In that case, the texture-bit was limited to one. Since the computed QP ($Q_{c,i}$) may oscillate noticeably for sequences with rapid changes in content complexity, changes in QP were limited to no more than ±2 units between pictures using

$$Q_{lm,i} = MAX[Q_{i-1} - 2, MIN(Q_{i-1} + 2, Q_{c,i})], \qquad (3)$$

where $Q_i$ is the final QP for encoding the $i$'th frame. The limited QP was adjusted using both buffer status and RCC measure. Then the final QP was obtained as

$$Q_i = \begin{cases} Q_{lm,i} - 1 & \begin{pmatrix}[(Q_{i-1} - Q_{lm,i}) < 2] \text{ and } [(CM_i > 1.09)] \\ \text{and } [(CBF_{i-1} - TBL_i) < b_r/(f_r \times \Gamma)]\end{pmatrix} \\ Q_{lm,i} + 1 & \begin{pmatrix}[(CM_i < 0.99)] \text{ and} \\ [(CBF_{i-1} - TBL_i) > b_r/(f_r \times \Gamma)]\end{pmatrix} \end{cases} \qquad (4)$$

where $CM_i$ indicates the RCC measure for the $i$'th frame. The MAD ratio was used as $CM_i$ for a fair comparison to the existing scheme.[4] $CM_i$ is the ratio of the $i$'th predicted MAD to the averaged actual MAD over all $P$ frames encoded previously. When the buffer status was low enough to keep the target-bit positive and the RCC measure was high, the QP decreased by one to reduce the fluctuation in video quality. In the experiments, when all frames were encoded by a constant QP (except for only one high-complexity frame that was encoded by one smaller QP), over-bit (the number of over bits due to the decrease of QP) accounted for less than 1% of resource-bit (the number of

**Table 1** Performance comparisons of the proposed scheme with Refs. 3 and 4 at 30 fps (#: number of reference pictures).

| Sequences | Target bit rate | # | Average Y-PSNR (dB) | | | Standard deviation | | | Encoded bit rate (kbps) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Ref. 3 | Ref. 4 | Proposed | Ref. 3 | Ref. 4 | Proposed | Ref. 3 | Ref. 4 | Proposed |
| Carphone | 9.60 kbps | 1 | 25.85 | 25.98 | 25.95 | 1.05 | 0.87 | 0.49 | 9.89 | 9.68 | 9.57 |
| | | 3 | 25.90 | 26.02 | 26.02 | 0.85 | 0.68 | 0.55 | 9.71 | 9.70 | 9.66 |
| Foreman | 19.20 kbps | 1 | 27.04 | 27.24 | 27.20 | 1.54 | 1.39 | 1.07 | 19.33 | 19.46 | 19.21 |
| | | 5 | 26.99 | 27.29 | 27.13 | 1.70 | 1.28 | 0.97 | 19.32 | 19.35 | 19.25 |

total given bits for encoding). This value did not cause even the right side of Eq. (1) to be negative. Figure 1 shows the effect of over-bit with respect to negative target-bit. The line of near-negative target-bit denotes that the right side of Eq. (1) is negative. There is enough margin to maintain positive target-bit. Figure 2 shows the percentage of over-bit in resource-bit and the decreased fluctuation in video quality. In brief, the test conditions were as follows: The sequence was "Foreman," and the constant QP was 41 over 100 frames except for only one high-complexity frame that was quantized by 40. In contrast, when the buffer status was high and the RCC measure was low, the QP increased by one to lessen the level of buffer fullness and prevent the target-bit of the next picture from being negative.

## 4 Case 2: Negative Target-Bit

As seen in Eq. (1), a negative target-bit occurs when $CBF_{i-1}$, to at least a certain extent, is larger than $TBL_i$. Generally, $CBF_{i-1}$ is large in a high-complexity picture and $TBL_i$ is small in a low avail-bit application. In these conditions, there is a high probability that the target-bit will be negative. Once a target-bit becomes negative, the level of the current buffer fullness must be reduced to make the next target-bit positive. To do so, the current QP was made larger than the previous one, as shown in Eq. (5). When the RCC measure was low, the increase in the current QP over the previous QP was greater, and the reduction in the buffer fullness level became more rapid. Hence, even if next picture was high RCC, the buffer had more room to avoid negative target-bit for the following picture:

$$Q_i = \begin{cases} Q_{i-1} + 2 & CM_i > 1.09 \\ Q_{i-1} + 3 & \text{otherwise} \end{cases}. \qquad (5)$$

After encoding a frame by a final QP, a linear regression method like Ref. 5 was used to update the parameters of the linear prediction model for MAD as well as $x_1$ and $x_2$ of the quadratic R-D model [Eq. (2)] for the next frame.

## 5 Experimental Results

Experiments were conducted on high-motion sequences ("Carphone" and "Foreman") under low avail-bit conditions (9.6 kbps and 19.2 kbps, 30 fps). The sequences were in QCIF 4:2:0 formats. H.264 reference software version JM6.1 was the test platform, and the proposed scheme was compared with the existing H.264 rate control schemes.[3,4]

For all tests, parameters were set as follows: RDO was enabled, the search range for motion estimation was 16, the number of reference frames was one for less computational load and three and five for better performance, a Hadamard transform was used, and the entropy coding method was CABAC. All other parameters were carefully selected to make the three schemes equivalent. The peak signal to noise ratio (PSNR) curves of the two sequences are plotted in Fig. 3. PSNR fluctuation of the proposed scheme was much less than that of Refs. 3 and 4 while maintaining a good PSNR. Detailed results are shown in Table 1. The proposed scheme considerably enhanced performance with respect to PSNR fluctuation and bit rate mismatch. The PSNR fluctuation of the proposed scheme was improved by 53.3% and 43.7% in "Carphone" and by 42.9% and 24.2% in "Foreman" compared to Refs. 3 and 4 respectively. Figure 4 shows the buffer fullness at each frame. The buffer size was set to half of the target-bit rate. According to Fig. 4, the proposed rate control managed a more stable buffer fullness level than Refs. 3 and 4.

## 6 Conclusion

An improved H.264 frame-layer rate control scheme for high-motion video at low bit rates and high frame rates was proposed. The scheme yielded much smaller fluctuations in video quality, lower mismatches between bit rates, and more stable buffer fullness levels compared with existing rate control schemes.

### References

1. T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. Circuits Syst. Video Technol.* **13**(7), 688–703 (2003).
2. Z. G. Li, F. Pan, K. P. Lim, G. Feng, X. Lin, and S. Rahardja, "Adaptive basic unit layer rate control for JVT," in *7th Meeting, Pattaya II*, JVT-G012-r1 (2003).
3. http://ftp3.itu.ch/av-arch/jvt-site/2003_03_Pattaya/JVT-G012r1_software.zip.
4. M. Jiang, X. Yi, and N. Ling, "Improved frame-layer rate control for H.264 using MAD ratio," in *IEEE Intl. Symp. Circuits Syst.*, Vol. **III**, pp. 813–816 (2004).
5. L. Hung-Ju, C. Tihao, and Z. Ya-Qin, "Scalable rate control for MPEG-4 video," *IEEE Trans. Circuits Syst. Video Technol.* **10**(6), 878–894 (2000).