GOP Level Rate-Control for Real-Time Video Transmission

Sanghyun Park

Dept. of Multimedia Engineering, Sunchon National University

shark@scnu.ac.kr

Abstract

The growing demand for advanced mobile multimedia let the wireless mobile internet access be developed rapidly. The current mobile network is capable of providing flexible quality of service mechanisms including exact estimation of network bandwidth. In this paper, we propose a new GOP level rate control algorithm which can accurately estimate an initial QP according to the estimated network bandwidth. In H.264/AVC, the first frame of a GOP is encoded in intra mode which generates a large number of bits. The amount of bits for I-frame affects the qualities of the following frames of a GOP since they are encoded using the remaining bits of the bits allocated to the GOP. In addition, the first frame is used for the inter mode encoding of the following frames. Thus, the initial QP affects the following frames as well as the first frame. In the proposed algorithm, we observe the relation between the initial QP and the number of bits of I-frame in a GOP. Based on the observation, we propose a new method to estimate the optimal initial QP which maximizes the PSNR of a GOP. It is shown by experimental results that the proposed algorithm predicts the optimal initial QP accurately and thus achieves better PSNR performance than that of the existing algorithm.

Keywords: *H.264/AVC, Initial QP, Rate control, Video compression*

1. Introduction

One of the main trends in the mobile communications sector is the connected everywhere, anytime, anyhow philosophy. This philosophy is also denominated Always Best Connected (ABC) philosophy [1]. It is based on the facts that several Radio Access Technologies (RATs) can co-exist and each technology provides different capabilities. Therefore, to provide best service, a mobile station (MS) always has to analyze the current situation and select a proper RAT which can provide the best service [2-4]. In addition, there have been many researches on how to estimate network capabilities, so an MS can accurately estimate the bandwidth of the new network when it changes an RAT.

The H.264/AVC standard achieves remarkable higher compression performance than the previous MPEG and H.26X standards [5]. To improve video quality, H.264 standard adopts state-of-the-art technologies. Among them, rate control technology aims to achieve good perceptual quality given the transmission bit rate constraint. Usually, rate control regulates the coded bit stream by adjusting quantization parameter (QP) while optimizing the video presentation quality. Recently, many rate control algorithms have been proposed for H.264/AVC, but most of them only focus on P-frame coding [6]. Generally, the QP for an I-frame depends on the average QP of P-frames in the previous group of pictures (GOP). In addition, this QP is modified to be adaptive to the transmission bit rate. The potential problem of this scheme is that given bit budget when encoding the current I-frame, it is difficult to accurately estimate the optimal QP since the characteristics of the current GOP is not considered [7]. However, it is quite important to control the quality of the I-frame to a suitable level for a fixed target output bit rate. A high-quality I-frame usually consumes more bits of the bits allocated to a GOP, which degrades the video quality of P- and B-frames in the same GOP due to frame skip and buffer overflow. On the contrary, a low-quality I-frame certainly degrades the video quality because I-frame is used for encoding P- and B-frames.

In this Paper, an adaptive initial QP determination algorithm is proposed when an MS changes networks. The algorithm is capable of accurately estimating the optimal QP of I-frame according to the bandwidth of the new network. Experimental results show that the proposed algorithm outperforms the existing method for H.264/AVC rate control.

The rest of this paper is organized as follows. Section 2 presents the existing rate control algorithm for the initial QP in H.264 reference software. The development of the proposed method of the adaptive initial QP determination is discussed in Section 3. Section 4 demonstrates the experimental results for performance comparison. Finally, a conclusion is drawn in Section 5.

2. Existing GOP Level Rate Control

Rate control framework for H.264/AVC has been proposed in JVT-G012 [6] and recently modified in JVT-W057 [8]. The algorithm is used to create the stream satisfying the available bandwidth provided by a network and is also compliant to hypothetical reference decoder (HRD). It consists of three tightly consecutive components: the GOP level rate control, the frame level rate control, and the basic unit level rate control. Among them, the GOP level rate control includes the calculation of the total number of bits for a GOP and the determination of the initial QP for the GOP. This paper focuses on the determination of the initial QP of the GOP level rate control.

An initial QP is set for the IDR picture and the first stored picture of the ith GOP. For the first GOP, the initial QP is predefined based on the available channel bandwidth as follows:

$$QP_1(1) = \begin{cases} 40, & bpp \le l_1 \\ 30, & l_1 < bpp \le l_2 \\ 20, & l_2 < bpp \le l_3 \\ 10, & bpp > l_3 \end{cases}, \tag{1}$$

$$bpp = \frac{bit\_rate}{f \times N_{pixel}}.$$

where $QP_1(i)$ is the initial QP of the I-frame (first frame) in the ith GOP; bit_rate, f, and N_{pixel} are the target bit rate, the predefined frame rate, and the number of pixels in a frame, respectively. The values of l_1, l_2, and l_3 are recommended for QCIF/CIF in [8].

For other GOPs:

$$QP_1'(i) = \frac{SumPQP(i-1)}{N_p(i-1)} - \min\left\{2, \frac{N_p(i-1)}{15}\right\}, \tag{2}$$

$$QP_1(i) = \max\{QP_1(i-1)-2, \quad \min\{QP_1(i-1)+2, QP_1'(i)\}\}$$

where $N_p(i)$ and $SumPQP(i)$ are the total number of P-frames and the sum of average QPs for all P-frames in the ith GOP, respectively. Eq. (2) shows that the QP for an I-frame depends on the average QP's of P-frames in the previous GOP. This initialization scheme is simple and adaptive to the available channel bandwidth, but it does not consider the visual characteristics of each video sequence [7].

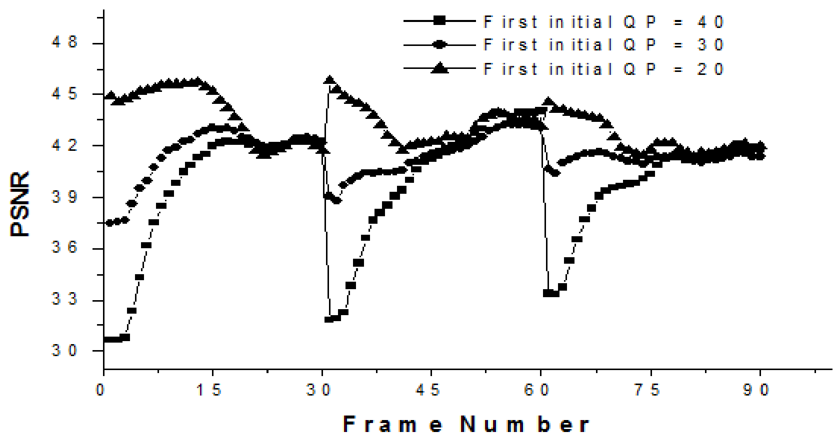


Figure 1. PSNR Comparison versus Frame Number for Akiyo Sequence when Initial QP's are 20, 30, and 40

Figure 1 shows PSNR results of QCIF Akiyo sequence when the GOP size is 30, the frame rate is 30fps, and the bit rate is 100kbps. JVT algorithm determines the first initial QP according to (1), so the first initial QP is set to 40. For comparison, PSNR results are added when the first initial QP's are 20 and 30. In the case of Akiyo sequence, the first initial QP of 40 is too big, so the quality of I-frame is not good. The bad quality of I-frame of the first GOP degrades the qualities of the following GOP's as well as that of the first GOP. On the other hand, when the first initial QP is 20, the quality of I-frame is much higher and the overall qualities of GOP's are also better than those of JVT algorithm.

The selection of the initial QP based on (1) and (2) has been adopted for implementation of the H.264/AVC reference model. However, in order to enhance the H.264 overall performance, a more efficient rate control scheme is needed. The details of the proposed rate control scheme which improves the existing method are described in the next section.

3. Proposed GOP Level Rate Control

This paper focuses on the initial QP determination of the GOP level rate control which can be applicable to the real-time application. Thus, it is assumed that the frame structure is "IPPP…" without B frame.

In JVT rate control scheme, the QP for an I-frame depends on the average QP of P-frames in the previous GOP as shown in Eq. (2). This initialization scheme is simple, but it is roughly adaptive to the available channel. Therefore the initial QP's of QCIF video sequences are always 40 when the bandwidth is less than about 114kbps. Figure 1 shows that the initial QP of 40 is not proper for Akiyo sequence. Also, it does not consider the characteristics of each video sequence.

Efficient rate control scheme has to find the optimal value more quickly. In addition, it has to take into consideration the properties of each video sequences such as frame complexities and motion characteristics. However, the algorithm becomes more complex as the number of parameters is increased, and the complicated algorithm cannot be used for real-time applications. The proposed algorithm uses only PSNR properties of a GOP, so it is simple and can be applicable to real-time applications.

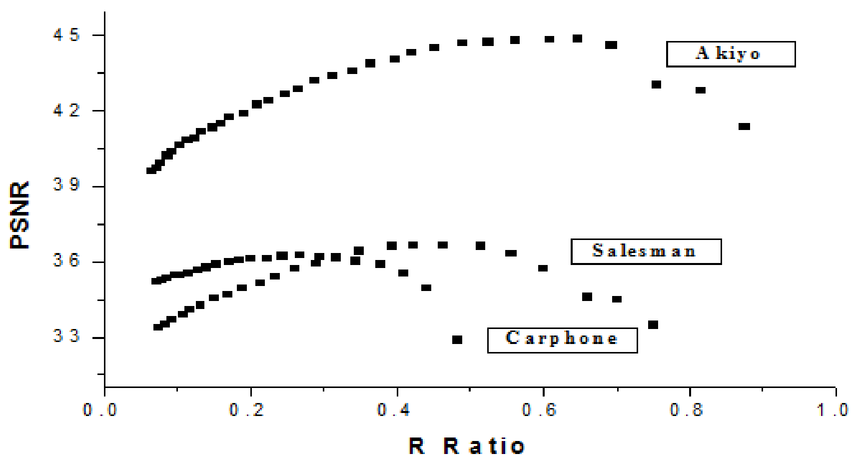
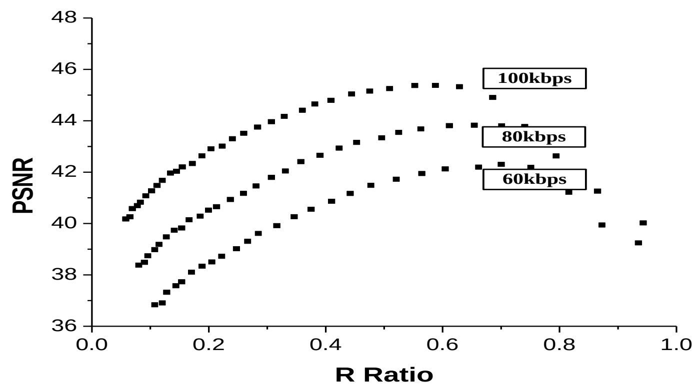
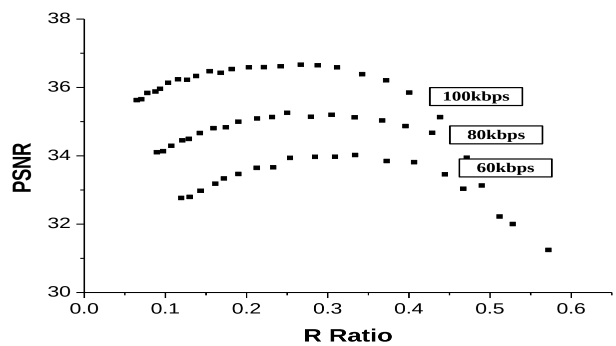


Figure 2. PSNR Comparison versus R Ratio for Akiyo, Carphone, and Salesman Sequences when the Bandwidth is 100kbps

Various test sequences have been encoded using different initial QP's in the H.264/AVC, and PSNR characteristics of GOP's have been studied. As the initial QP decreases, the PSNR of I-frame becomes better but that of P-frame is more degraded. It is because I-frame consumes so much bits that there are not enough bits left for P-frames which are encoded using the remaining bits. On the contrary, as the initial QP increases, the PSNR of I-frame becomes worse and it degrades the qualities of P-frames. To find the optimal initial QP which maximizes the PSNR of a GOP, we observe the relations between the PSNR of a GOP and the number of bits of the I-frame. Let R Ratio denote the number of bits of the I-frame divided by the number of bits allocated to a GOP.



(a) Akiyo sequence



(b) Carphone sequence

Figure 3. PSNR Comparison versus R Ratio when the Bandwidth is 60kbps, 80kbps, and 100kbps

Figure 2 shows the relations between the PSNR of a GOP and the R Ratio for Akiyo, Carphone, and Salesman sequences with the target bit rate of 100 kbps. As it can be seen from the figure, the relation of each sequence is unique. But it is common that the PSNR of a GOP increases as the R Ratio increase. The increase of R Ratio means the decrease of initial QP. After reaching the maximum PSNR value, the PSNR decreases as the R Ratio increases. That is, there exists an optimal R Ratio which maximizes the PSNR of a GOP and the optimal R Ratio is unique to each sequence.

We also observe the characteristics of R Ratio according to the network bandwidth using various test sequences. Figure 3 shows that the shape of PSNR results for each bandwidth is very similar to each other. That is, the results of 60kbps become very similar to those of 100kbps when a constant of 2.5 is added to the results of 60kbps. It means that the optimal R Ratio is similar regardless of the network bandwidth.

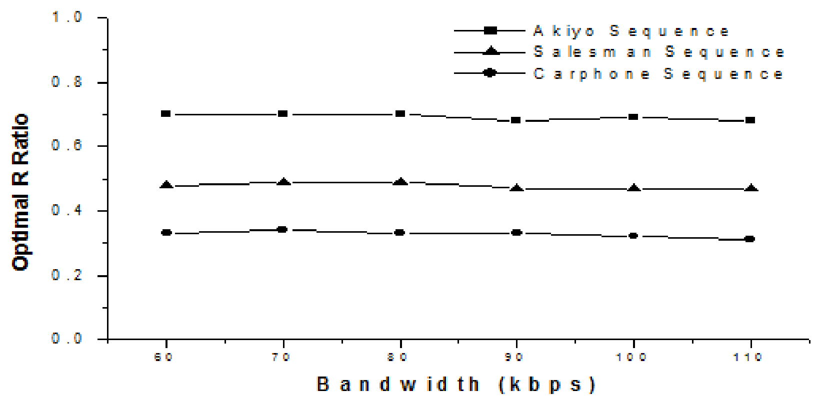


Figure 4. Optimal R Ratio Comparison versus Network Bandwidth for Akiyo, Carphone, and Salesman Sequences

We investigate the optimal R Ratio according to the network bandwidth. Let R_{op} denote the optimal R Ratio. Figure 4 shows that R_{op}'s of each sequence is similar regardless of the network bandwidth.

The proposed rate control algorithm uses the property that R_{op} of a video sequence maintains similar values as the target bandwidth varies. When the MS changes the network, details of the proposed rate control algorithms is as follows:

Step 1: Determining R_{op}

In the proposed scheme, the first GOP of a sequence is encoded by the existing method. That is, the first initial QP is set to 40 and the initial QP is adapted gradually to the optimal value by (2). To find R_{op}, the MS updates R_{op} with the new R Ratio if the PSNR of the GOP is greater than the maximum PSNR value after encoding a GOP. R_{op} usually converges to a value after encoding several GOP's. We assume that the MS knows the R_{op} of the current video before changing the network connection.

Step 2: Estimating the Bandwidth of New Network

We assume that an MS can estimate the bandwidth of the new network. There are many methods for an MS to estimate the bandwidth. For instance, QoS-aware scheduling algorithms for LTE OFDMA was proposed to deliver real-time video, where an MS negotiates the bandwidth with the network [2, 3, 10]. That is, an MS knows the bandwidth before establishing new connection.

Step 3: Calculating the Optimal Initial QP

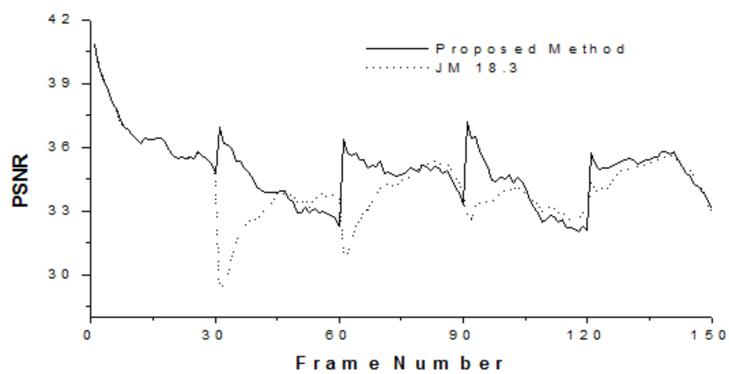
In this step, we have to calculate the QP which is used to encode the I-frame and generate the number of bits corresponding to the R_{op}. There are several methods for H.264/AVC which estimate the number of bits of the I-frame given the QP value. To calculate the optimal initial QP, we use Kamaci's method [9]. Kamaci's method uses two parameters. To determine parameters, we first encode the I-frame with the QP of 10 and then determine model parameters and calculate the optimal initial QP.

4. Experimental Results

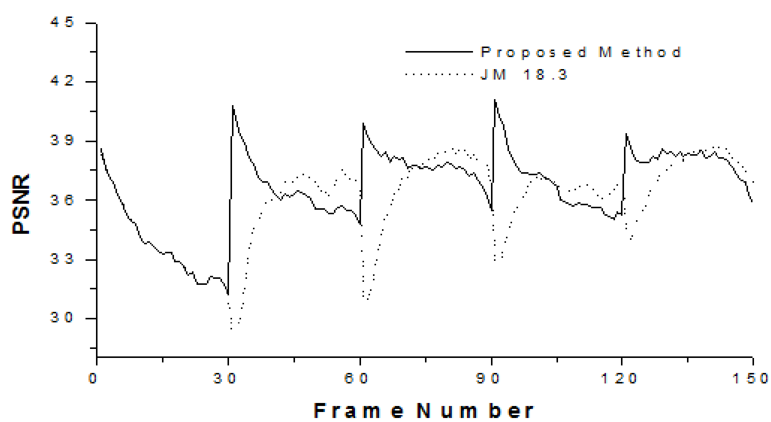
Numerous experiments have been conducted to evaluate the performance of the proposed rate control algorithm, which has been implemented with the latest version of the JVT reference software, JM18.3 using baseline profile. The results achieved here are compared with those achieved using JVT-W057 rate control algorithm adopted by JM18.3.

The same encoding parameters are used for both algorithms in order to ensure that the comparison is fair. For the experiments, the following test conditions are used: an “IPPPP…” GOP structure with a GOP size of 30 is used, the motion vector search range and the number of multiple reference frames for motion estimation are set to 16 and 2, respectively, and fast full search motion estimation and rate-distortion optimization are enabled.

The simulation was conducted with the first 150 frames of four QCIF test sequences of Akiyo, Carphone, Foreman, Salesman. In order to ensure the equivalence of the rate control parameters, the sizes of the basic units for the basic unit-level rate control are fixed at 1 macroblock. The network bandwidth varies at the second GOP from 60kbps to 110kbps and from 110kbps to 60kbps.



(a) bandwidth varies from 110kbps to 60kbps at the 31st frame



(b) bandwidth varies from 60kbps to 110kbps at the 31st frame

Figure 5. PSNR Results using GOP Level Rate Control Algorithms of JVT-W057 and the proposed Algorithm for Carphone Sequence

Table 1. PSNR Comparison with Video Sequences

Video sequence	Bandwidth (kbps)	JVT[8]			Proposed		
		2nd	3rd	4th	2nd	3rd	4th
Akiyo	60 →110	40.09	40.71	40.09	45.36	43.63	43.18
	110→ 60	36.78	37.24	37.02	42.05	40.73	40.05
Carphone	60 →110	35.64	36.74	36.16	36.56	37.85	36.83
	110→ 60	32.79	34.02	33.31	34.01	34.96	33.82
Foreman	60 →110	34.41	34.42	33.58	35.71	35.22	34.28
	110→ 60	31.86	31.87	31.02	33.11	32.67	31.27
Salesman	60 →110	33.93	35.17	36.02	37.25	38.41	38.42
	110→ 60	30.92	31.99	32.76	33.46	34.34	34.85

Since the major issue for video coding is the quality of the video at the given target bit rate, the average PSNR value of each GOP is calculated and listed in Table 1 in order to provide an objective evaluation of the video quality. The proposed scheme uses JVT algorithm for the first GOP, so the PSNR results of the first GOP's are not included in Table 1. The proposed scheme shows better video quality than the rate control of JVT algorithm in terms of the average PSNR value. The frame-to-frame PSNR results of Carphone sequence are shown in Figure 5, where the network bandwidth varies from 110kbps to 60kbps and from 60kbps to 110kbps at the 31st frame. It is shown that better results are obtained by the proposed scheme than JVT algorithm.

5. Conclusion

In this paper, an adaptive initial QP determination algorithm for H.264/AVC is proposed. The proposed algorithm uses the property that the optimal R Ratio's of each sequence maintain similar values regardless of the network bandwidth. The proposed algorithm also takes the characteristics of each video sequence into consideration. Thus it can precisely estimate the optimal initial QP compared with the existing method. Experimental results show that the proposed scheme achieves better video quality than that of JVT-W057. In case of Akiyo sequence, the proposed algorithm improves the average PSNR of GOPs more than 3dB.

References

[1] M. Tsagkaropoulous, I. Politis, T. Dagiuklas and S. Kotsopoulos, "Enhanced vertical handover based on 802.21 framework for real-time video streaming", Proceedings of ProMobiMedia '09, London, United Kingdom, **(2009),** pp. 62-67.

[2] J. Y. Pyun, D. Y. Choi, S. K. Noh and G. R. Kwon, "OFDMA Channel Aware Error-Resilient Mobile Streaming Video Service over Mobile WiMAX", IEEE Trans. on Comsumer Electron., vol. 57, no. 1, **(2011)**.

[3] A. A. Lenskiy and S. Seol, "The Analysis of R/S Estimation Algorithm with Applications to WiMAX Network Traffic", International Journal of Multimedia and Ubiquitous Engineering, vol. 7, no. 3, **(2012)**.

[4] J. Choi, S. Lee, I. Kim and Y. Woo, "Live-Video Service using Multicast in Wireless Network", International Journal of Multimedia and Ubiquitous Engineering, vol. 7, no. 4, **(2013)**.

[5] S. Ghorbani and F. ZArgari, "A Unified Architecture for Implementation of the Entire Transforms in the H.264/AVC Encoder", International Journal of Multimedia and Ubiquitous Engineering, vol. 8, no. 1, **(2013)**.

[6] B. Yan and M. Wang, "Adaptive Distortion-Based Intra-Rate Estimation for H.264/AVC Rate Control", IEEE Signal Process. Lett., vol. 16, no. 3, **(2009)**.

[7] X. Jing, L. Chau and W. Siu, "Frame Complexity-Based Rate-Quantization Model for H.264/AVC Intraframe Rate Control", IEEE Signal Process. Lett., vol. 15, no. 1, **(2008)**.
[8] Z. Li, F. Pan and K. Lim, "Text Description of Joint Model Reference Encoding Methods and Decoding Concealment Methods", JVT-W057, 23rd Meeting, San Jose, USA, **(2007)**.
[9] N. Kamaci, Y. Altunbasak and R. M. Mersereau, "Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortin models", IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 8, **(2005)**.
[10] H. Luo, S. Ci, D. Wu, J. Wu and H. Tang, "Quality-Driven Cross-Layer Optimized Video Delivery over LTE", IEEE Comm. Mag., vol. 48, no. 2, **(2010)**.

Authors

Sanghyun Park received his B.S., M.S., and Ph.D. degrees in electronics engineering from Korea University, Seoul, Korea in 1995, 1997, and 2002, respectively. In 2004, he joined the Department of Multimedia Engineering at Sunchon National University, Sunchon, Korea, where he is currently an associate professor. His research interests include image processing, pattern recognition, and multimedia communication.