Improved Initial QP Prediction Method in H.264/AVC

Soon-young Kwon
Ksy2020@dgist.ac.kr

Sang-heon Lee
pobbylee@dgist.ac.kr

Dong-ha Lee
dhlee@dgist.ac.kr

Department of IT Intelligent Embedded Software Team
Daegu Gyeongbuk Institute of Science & Technology (DGIST)

ABSTRACT
To improve video quality and coding efficiency, H.264/AVC adopted an adaptive rate control. But this method has a problem as it cannot predict an accurate quantization parameter (QP) for the first frame. The first QP is decided among four constant values by using encoder input parameters. It does not consider encoding bits, results in significant fluctuation of the image quality and decreases the average quality of the whole coded sequence. In this paper, we propose a new algorithm for the first frame QP decision in the H.264/AVC encoder. The QP is decided by the existing algorithm and the first frame is encoded. According to the encoded bits, the new initial QP is decided. We can predict optimal value because there is a linear relationship between encoded bits and the new initial QP. Next, we re-encode the first frame using the new initial QP. Experimental results show that the proposed algorithm not only achieves better quality than the state of the art algorithm, but also adopts a rate control for the sequence that was impossible with the existing algorithm. By reducing fluctuation, subjective quality also improved.

Keywords
H.264/AVC encoder, initial QP, rate control, and a linear QP prediction model.

1. INTRODUCTION
H.264/AVC was approved by ISO/IEC and by ITU-T in 2003 [1]. To improve coding efficiency, it adopted intra prediction and variable block size motion estimation with multiple reference frames at quarter-pixel accuracy [2, 3]. In contrast with the MPEG-4 advanced simple profile, the H.264/AVC coder further reduces the bit-rate by up to 50% [4]. Accordingly, it is expected to be in wide use and to replace other standards in the near future.

Recently, as multimedia data transmission on the wireless network increase and various devices receive it, the rate control in H.264/AVC becomes an important issue. The rate control is used to compute QP for the current frame and the number of skipped frames. Specifically, the rate control in H.264/AVC is more important than other components because QP is used in both the RDO and rate control. To achieve the rate control in the recommended algorithm, the encoder fixes an initial QP to a constant value so that the encoder needs a few seconds to adapt the input sequence. During this time, it happens to overflow, underflow, fluctuate or reduce in quality.

In this paper, we propose a linear QP prediction model that is a QP decision algorithm for the first frame to reduce fluctuation and adaptation time. The first frame is encoded by QP by using a existing algorithm. Then, we predict optimal QP according to encoded bits of the first frame. Encoded bits of the first frame are in inverse proportion to QP. In other words, if QP improves, the number of encoded bits may be reduced. The linear QP prediction model is proposed to predict optimal initial QP by the relation and model values obtained from actual encoding. In this experiment, we have implemented our proposed rate control scheme by enhancing the jm12.2 test model software and confirmed average PSNR (Peek Signal to Noise Rate) and variation PSNR under Korea's TDMB (Terrestrial Digital Multimedia Broadcasting) standard.

This paper is organized as follows. In section 2 we describe the rate control and method of QP decision for the first frame in H.264/AVC. We present problems with the state of the art algorithm and the proposed algorithm in section 3 and show experimental results in Section 4. We close the paper with concluding remarks in Section 5.

2. RELATED WORKS
2.1 Rate Control in H.264/AVC
Rate control is not a part of the H.264/AVC standard, but the standards group has issued non-normative guidance to aid implementation. Since QPs are involved in both rate control and RDO (Rate Distortion Optimization) in H.264/AVC, there exists a dilemma when the rate control is implemented. To perform RDO for a MB (Macroblock), a QP should first be determined for the MB by using the MAD (Mean Absolute Different) of the MB and the number of header bits. However, the MAD of the current MB and the number of header bits are only available after performing the RDO. This is a typical chicken and egg dilemma [5, 6]. Figure 1 shows a block diagram of this coding process. To solve the problem mentioned above, we should efficiently estimate MAD.

MOBIMEDIA 2007, August 27-29, Nafpaktos, Greece

DOI 10.4108/ICST.MOBIMEDIA2007.1703

This work was supported by the DGIST Basic Research Program of the MOST.

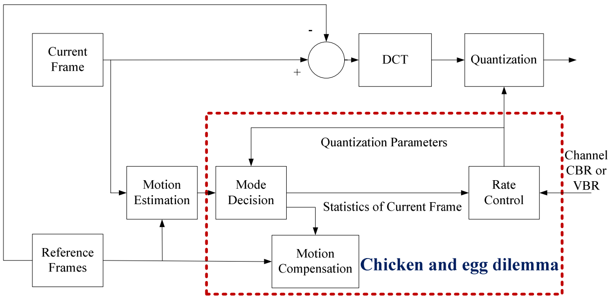

Figure 1. The coding process related to the rate control.

In the current standard, the chicken and egg dilemma can be solved by a linear model

$$MAD_{cb} = X_1 \times MAD_{pb} + X_2 \tag{1}$$

$$MAD = \sum |residual| = \sum |source - prediction| \ , \tag{2}$$

where MAD_{cb} is the predicted MAD of the basic unit in the current frame, MAD_{pb} is the actual MAD of the basic unit in the previous frame, and X_1 and X_2 are model variables. The rate control scheme works step by step as follows:

Step 1. Allocate the target bits for the current frame by using the fluid traffic model.

Step 2. Predict the MADs of the remaining basic units in the current frame by using linear model (2).

Step 3. Allocate the remaining bits to the remaining basic units in the current frame according to the predicted MADs.

Step 4. Calculate QP by using a quadratic R-D model

$$R = \frac{C_1 \times MAD}{Q_{step}} + \frac{C_2 \times MAD}{Q_{step}^2} \ , \tag{3}$$

where R is the target bits for the current frame, C_1 and C_2 are model variables, and Q_{step} is the quantization step.

Step 5. Perform RDO for each MB in the current basic unit by the QP derived from Step 4 and update model variables.

Figure 2 is a block diagram that shows important elements of the H.264/AVC rate controller. The encoder system must be damped to guarantee stability and to minimize perceptible variations in quality, so a QP-Limiter block is applied which typically limits changes in QP to no more than ±2 units between pictures. Because the rate controller can't directly change QP, it needs a few seconds to estimate reasonable QP. This is the reason that incorrect QP decisions in the previous frame negatively influence successive frames. A rate quantization model block and complexity estimation block associates Step 4 and Step 2, respectively. A QP initializer block focused by this paper is described in the next section. Refer to [5-7] for a more detailed explanation about the rest of the elements. In this paper, we define QP_0 , both QP in the first frame of the first GOP and initial QP, and QP_0' as new optimal QP_0 obtained by using the proposed algorithm.

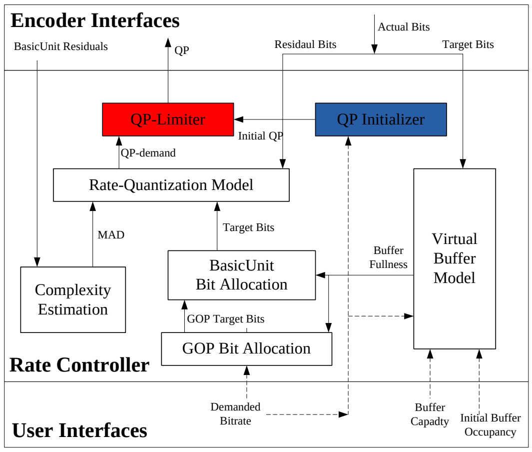

Figure 2. Elements of the H.264/AVC rate controller

2.2 QP_0 decision algorithm in H.264/AVC

The first frame of the first GOP is encoded intra mode because there are no reference frames to reduce temporal redundancy. There is no information to estimate QP. Therefore, the encoder selects a constant value among four values which are different according to bpp as obtained below by (4), instead of a linear MAD prediction model in the current H.264/AVC.

$$bpp = \frac{T_{bitrate}}{f \times N_{pixel}} \tag{4}$$

$$QP_o = \begin{cases} 35 & bpp \le l1 \\ 25 & l1 < bpp \le l2 \\ 20 & l2 < bpp \le l3 \\ 10 & l3 < bpp \end{cases} , \tag{5}$$

where bpp is bits per pixel, $T_{bitrate}$ is a target bitrate, and N_{pixel} is the number of pixels in a picture. The candidate of QP_0 and $l1$, $l2$, and $l3$ values changed after the standard was committed; $l1$ =0.1/02/0.6, $l2$ =0.3/0.6/1.4 and $l3$ =0.6/1.2/2.4 are recommended for QCIF, CIF and other resolutions, respectively, in jm12.2 [8].

3. A LINEAR QP_0 PREDICTION MODEL

3.1 Problems with the QP_0 Decision Algorithm

If the encoder does not consider features of input sequence when encoding the first frame, it will not perform an efficient rate control. For example, for any sequence, QP_0 is fixed as "10" under 30 frames/s, 544 kbps, and QCIF size (176 × 144). In case of sequences with high activity, the first frame should be encoded using a bigger QP than 10 but it will be encoded by using a constant value 10. The next frame is also encoded by using a QP smaller than optimal QP on account of the QP-limiter. As a result, the quality of a sequence drops and fluctuation increases.

Figure 3 shows the PSNR (Peek Signal to Noise Rate) and bitrate of the frames ranging from 1 to 200 for the Football and the Carphone QCIF sequences under 30 frames/s, and 544 kbps. In case of a sequence of Football which has high activity and many objects, 4 GOPs are bigger than the target bitrate and 7 GOPs are smaller than the target bitrate among 13 GOPs. As a result, the rate control for that sequence has failed for that sequence. In addition, the graph shows radical PSNR changes during the first part of the Carphone sequence, about 30 frames. We can show radical PSNR changes in the next GOP because QP change between GOPs is also limited in H.264/AVC.

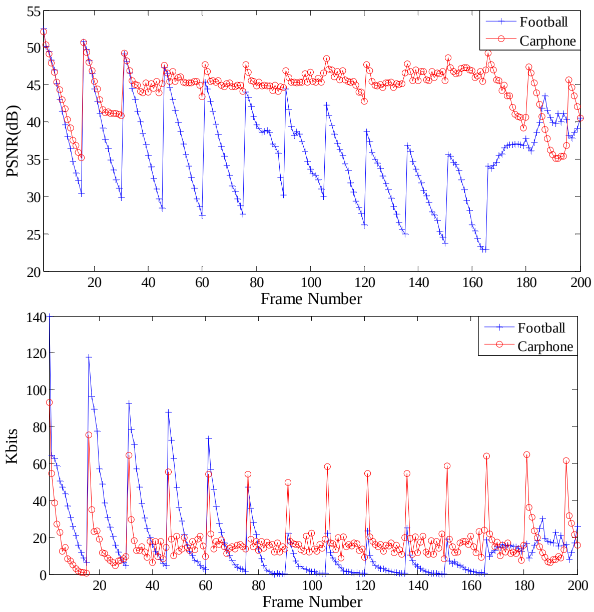

Figure 3. The PSNR performance for the sequence

3.2 A Linear QP_0 Prediction Model

In this section, we propose a linear QP_0 prediction model which is an algorithm derived from the current standard used to decide QP. Bitrate is in inverse proportion to QP. If a QP value increases by 12%, then the bitrate decreases by 12% statistically in H.264/AVC [7]. Considering the relationship, we can predict QP_0' by using E_0 , bits obtained from encoding the first frame. If the result bits E_0 are more than target bits, the encoder has to use a bigger QP value. On the other hand, if they less than the target bits, the encoder should use a smaller QP value. Thus, E_0 are directly proportional to optimal QP_0. We can estimate the feature of input sequence according to E_0 and then predict QP_0' by using E_0. As a result, there is a linear relationship between E_0 and QP_0'. The relationship is formalized as

$$QP_0' = Max\,(1, Min\,(\left[\alpha \times E_0 + \beta\right], 51)) \quad , \qquad (6)$$

where E_0 is bits produced from the first frame, and α , β are linear relation model variables which change in accordance with the conditions of encoding, so that they are obtained heuristically. Table 1 shows the value of linear model variables for QCIF and CIF usually used in Korea's TDMB. We confirm that most sequences keep the linear relation. The smaller the resolution is, the bigger the slope of equation (6). The reason is that when QP values vary in a fixed size under the same bitrate, smaller QP values will produce a much bigger difference in bits.

For example, the existing algorithm selects 10 for QP_0 under the QCIF sequence of Football, 30 frames/s, and 544 kbps. However, the proposed scheme decides QP_0' as below. The first frame is encoded by using 10 derived from the existing algorithm and then encoder produce "143272" bits. Since we can predict overflow through this value, we calculate QP_0' , 27, by using equation (6) as

$$\left[2.17 \times 10^{-4} \times 143272 - 4\right] = 27 \qquad (7)$$

Finally, the first frame is re-encoded by using QP_0' .

One more encoding time for the first frame is added to the proposed algorithm, but the additional time can be negligible when compared to the entire sequence encoding time. Also, encoding time for the Intra frame is only 1/3 ~ 1/4 of that for the Inter frame.

Table 1. Linear QP_0 prediction model variables

	QCIF	CIF
α	2.17×10^{-4}	1.04×10^{-4}
β	-4	24

4. EXPERIMENTAL RESULTS

In this section, we compare the performance of our proposed scheme with that of H.264/AVC. We have implemented this scheme by modifying the jm12.2 [9]. In Korea's TDMB standard, encoding conditions are less than 352×288 resolution and have 30 frames/s. In this experiment, test video sequences are constrained by a GOP structure of (N=15, M=1) and by 544kbps, at a constant frame rate of 30. Also, it involved encoding the first 60 frames because we focus on encoding the first part of a sequence. Table 2 denotes the whole result. "Original" and "Proposed" represent the results of the rate control in H.264/AVC and proposed scheme, respectively. "Average PSNR" represents an average PSNR of 60 frames and "PSNR variation" is used to verify the PSNR difference between frames. For average PSNR, the proposed scheme performs better than the original scheme except Football. This Football overflows because the output bitrate exceeds of the first frame the target bitrate by more two times in the original scheme; however, the proposed scheme succeeded in achieving the expected target bitrate (543.92 kbits/s). Although the PSNR variation is not able to completely show PSNR change between frames, we present it to express our findings numerically. The proposed scheme has a smaller value in the part of the PSNR variation than the original scheme except Tempete, yet there is a negligible gap.

Figure 4 shows the PSNR performance for the sequence from frame 1 to 60 for Foreman and Paris, respectively. A result of Foreman that has QCIF resolution shows good performance for the proposed scheme. However, the algorithm of the H.264/AVC shows a wild fluctuation and it can take control of rate after encoding more than 50 frames. In the case of Paris with CIF resolution, PSNR continually increases to 30 frames for the original scheme because it uses a bigger QP_0. Figure 5 compares subjective quality. One more encoding time is added to the proposed algorithm compared to a fixed QP scheme, but it can be negligible as encoding time for the entire sequence.

Table 2. Results of the experiment for test sequences

Test sequence	Initial QP		Average PSNR (dB)		PSNR variation		Bitrate (kbits/s)	
	Original	Proposed	Original	Proposed	Original	Proposed	Original	Proposed
Football (QCIF)	10	27	39.27	35.32	47.09	1.08	1107.59	543.92
Carphone (QCIF)	10	17	44.46	45.16	11.87	0.80	545.55	545.61
Foreman (QCIF)	10	20	41.24	42.14	24.43	0.88	544.26	539.94
Paris (CIF)	35	32	33.19	33.77	0.56	0.50	546.70	548.03
Highway (CIF)	35	26	39.54	39.98	0.90	0.18	547.53	548.49
Tempete (CIF)	35	32	30.47	30.55	0.44	0.70	547.22	548.05

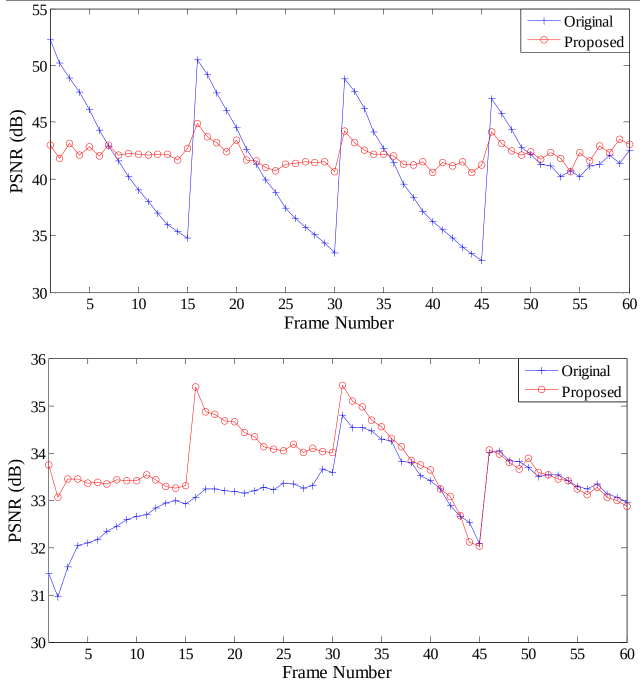

Figure 4. PSNR of each frame for sequences of Carphone and Paris, respectively

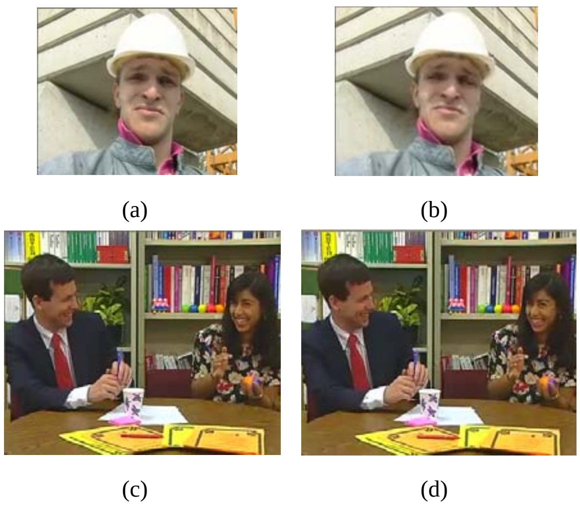

Figure 5. The real images for the sample video sequences. Foreman (a), (b); Paris (c), (d); The proposed scheme images are in the left column, and the original scheme images are in the right column.

5. CONCLUSIONS

In this paper, we proposed a linear QP_0 prediction model to decide optimal QP for the first frame in the H.264/AVC rate controller. The existing algorithm in the H.264/AVC standard is inefficient because it decides a constant value by using input parameters regardless of input video sequence features. There is an inverse relationship between QP value in the quantization process and bits produced after encoding. According to this relationship, output bits must be in proportion to optimal QP_0. So we predict QP_0' for the input video sequence by using bits produced after the first frame encoding. We demonstrated good performances on average PSNR and PSNR variation though experiments under TDMB standards that use H.264/AVC in Korea. Specifically, we proposed a scheme that can also take a rate control for the video sequence that was impossible to do in the state of the art scheme. After encoding the first frame, we could optional encode again for coding efficiency. Our next research project will involve an attempt reduce the computational complexity for additional encoding time and adaptive model variable decision algorithm under time varying channel bandwidth.

6. REFERENCES

[1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC), 2003

[2] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra. Overview of the H.264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 560–576, 2003.

[3] Iain E.G. Richardson. H.264 and MPEG-4 Video Compression. *WILEY* , 2003

[4] J. Ostermann et al., Video Coding with H.264/AVC: Tools, Performance, and Complexity. *IEEE Circuits Syst. Magazine*, vol.4, no.1, pp.7-28, Apr. 2004.

[5] Z. G. Li, W. Gao, F. Pan, et al. Adaptive Rate Control for H.264. *Journal of Visual Communication and Image Representation*, pp.377-4062005.

[6] Ahenggo Li, Wen Gao, Feng Pan. Adaptive Rate Control with HRD Consideration. JVT-H014, 8th Meeting: Geneva, May 20-26, 2003

[7] http://www.pixeltools.com/rate_control_paper.html

[8] Alexis Michael Tourapis. Revised H.264/MPEG-4 AVC Reference Software Manual; Section 4.8,JVT-Q042, 17th Meeting: Nice, FR,14-21 October, 2005

[9] http://iphome.hhi.de/suehring/tml/download