



PSNR-based Initial QP Determination for Low Bit Rate Video Coding

Sanghyun Park*, *Member, KIICE*

Department of Multimedia Engineering, Sunchon National University, Sunchon 540-742, Korea

Abstract

In H.264/AVC, the first frame of a group of pictures (GOP) is encoded in intra mode which generates a large number of bits. The number of bits for the I-frame affects the qualities of the following frames of a GOP since they are encoded using the bits remaining among 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 quantization parameter (QP) affects the following frames as well as the first frame. In this paper, an adaptive peak signal to noise ratio (PSNR)-based initial QP determination algorithm is presented. In the proposed algorithm, a novel linear model is established based on the observation of the relation between the initial QPs and PSNRs of frames. Using the linear model and PSNR results of the encoded GOPs, the proposed algorithm accurately estimates the optimal initial QP which maximizes the PSNR of the current 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.

Index Terms: H.264/AVC, Initial quantization parameter, Rate control, Video compression

I. INTRODUCTION

Recently, the H.264/AVC standard, which was jointly developed by International Telecommunication Union (ITU) and Moving Picture Experts Group (MPEG), has been widely used in many applications for video coding. H.264/AVC outperforms previous coding standards and has many outstanding features, such as various intra/inter prediction modes, multiple reference frames, rate-distortion optimization, and variable block sizes [1]. However, the H.264/AVC standard does not take into consideration the issue of maintaining a constant bit rate (CBR) through the network channel. Hence, it is necessary to implement a rate control algorithm in the video encoder in order to transmit the coded video sequence without any abrupt variations of the bitrate over time under conditions of limited channel

bandwidth [2].

Usually, rate control aims to achieve good perceptual quality given the transmission bit rate constraint. That is, rate control regulates the amount of the coded bits by adjusting the quantization parameter (QP) while maximizing the video presentation quality. To achieve this, the rate-quantization (R-Q) model is often employed for representing the coded bits by means of QP and other parameters such as the mean absolute difference (MAD) of a residual MB and the percentage of zero quantized coefficients [3]. Unfortunately, using parameters such as MAD for R-Q modeling causes the chicken-and-egg dilemma because the Lagrangian method employed in H.264 needs to be available before mode decision but until the end of mode decision, rate control (RC) cannot access the statistics such as MAD for determining QP [4]. Li et al.

Received 15 June 2012, Revised 20 July 2012, Accepted 27 July 2012

*Corresponding Author E-mail: shark@scnu.ac.kr

Open Access <http://dx.doi.org/10.6109/jicce.2012.10.3.315>

print ISSN:2234-8255 online ISSN:2234-8883

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © The Korea Institute of Information and Communication Engineering

[3] in JVT-G012 have proposed an adaptive rate control framework for H.264/AVC, where a single-pass rate control method based on the quadratic R-Q model is used and a linear model for MAD prediction is employed to solve the above dilemma.

Recently, many rate control algorithms have been proposed for H.264/AVC to improve JVT-G012, but most of them only focus on P-frame coding. However, how to encode the I-frame of a group of pictures (GOP) is also a very important factor influencing the RC performance. Usually, the I-frame and the first P-frame of a GOP are encoded using the predetermined QP, which is called the initial QP. In many RC algorithms, the initial QP of the first GOP is determined only depending on the bits per pixel (BPP) as JVT-G012 does. From the second GOP, the initial QP for an I-frame depends on the average QP of the P-frames in the previous GOP. The potential problem of this scheme is that given a bit budget when encoding the current I-frame, it is difficult to accurately estimate the QP since the characteristics of the current GOP are not considered [4-6]. 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 the P- and B-frames in the same GOP due to frame skip and buffer overflow. On the other hand, a low-quality I-frame certainly degrades the video quality because the I-frame is used for encoding P- and B-frames. Usually, given the same BPP, a large initial QP is desired for video sequences with complex spatial details or high motion types, whereas, for video sequences with simple spatial contents or low motions, a small initial QP will be advantageous. Thus, the initial QP should be determined by considering BPP as well as the contents of the video sequence [4].

In this paper, an adaptive peak signal to noise ratio (PSNR)-based initial QP determination algorithm is proposed. By considering the characteristics of the contents, the proposed algorithm is capable of accurately estimating the initial QP for a GOP compared with the conventional methods. 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 II 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 III. Section IV demonstrates the experimental results for performance comparison. Finally, a conclusion is drawn in Section V.

II. EXISTING RATE CONTROL SCHEME

A rate control framework for H.264/AVC has been proposed in JVT-G012 [3] and recently modified in JVT-W057 [7]. The algorithm is used to create the stream satisfying the available bandwidth provided by a channel and is also compliant with a 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.

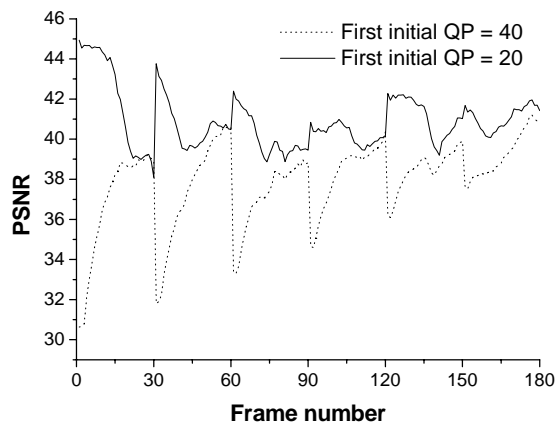


Fig. 1. Peak signal to noise ratio (PSNR) comparison versus frame number for Akiyo sequence.

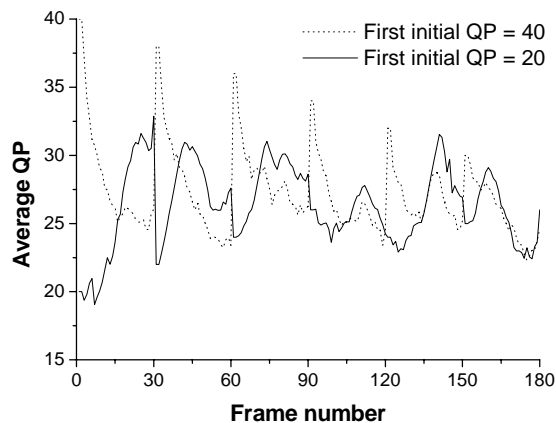


Fig. 2. Average quantization parameter (QP) comparison versus frame number for Akiyo sequence.

An initial QP $QP_i(1)$ is set for the IDR picture and the first stored picture of the i^{th} GOP. For the first GOP, $QP_1(1)$ is predefined based on the available channel bandwidth as follows:

$$QP_1(1) = \begin{cases} 40 & bpp \leq l1 \\ 30 & l1 < bpp \leq l2 \\ 20 & l2 < bpp \leq l3 \\ 10 & bpp > l3 \end{cases} \quad (1)$$

$$\text{where } bpp = \frac{R}{f \times N_{pixel}}$$

R , f , and N_{pixel} are the available bit rate, the frame rate, and the number of pixels in a frame, respectively. In this paper, it is assumed that three parameters have constant values. The three values of $l1$, $l2$, and $l3$ are recommended for quarter common intermediate format (QCIF)/CIF and a picture size larger than CIF in [7].

For the other GOPs, the initial QP's are calculated as follows:

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

$$QP_i(1) = \max\{QP_{i-1}(1) - 2, \min\{QP_{i-1}(1) + 2, QP'\}\}$$

where $N_p(i)$ is the total number of stored pictures in the i^{th} GOP, and $SumPQP(i)$ is the sum of the average QP's for all stored pictures in the i^{th} GOP. It is further adjusted as follows:

$$QP_i(1) = QP_i(1) - 1, \text{ if } QP_i(1) > QP_{i-1}(N_{i-1} - L) - 2 \quad (3)$$

where $QP_{i-1}(N_{i-1} - L)$ is the average QP of the last stored picture in the previous GOP, and L is the number of successive non-stored pictures between the two stored pictures.

Fig. 1 shows PSNR results of the QCIF Akiyo sequence when the GOP size is 30, the frame rate is 30 fps, and the bit rate is 60 kbps. The JVT algorithm determines the first initial QP according to Eq. (1), so the first initial QP is set to 40. For comparison, PSNR results are added when the first initial QP is 20. In the case of the Akiyo sequence, the first initial QP of 40 is too big, so the quality of the I-frame is not good. The bad quality of the 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 the I-frame is much higher than that of the previous case and the overall qualities of the GOP's are also better than those of the JVT algorithm. Fig. 2 shows the average QP of each frame of the sequence. From the second GOP, the initial QP is calculated by Eqs. (2) and (3), so the maximum difference between the two successive GOP's is 2. It is shown that the initial QP's vary gradually in the range of -2 to 2. Therefore, if the first initial QP is set to be too big or too small, the quality degradation is propagated to the following QOP's.

The selection of QP based on Eqs. (1), (2), and (3) 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.

III. PROPOSED RATE CONTROL SCHEME

This paper focuses on the determination of the initial QP of the GOP level rate control. In addition, rate control for real-time application is considered, so it is assumed that the frame structure is "IPPP..." without the B frame.

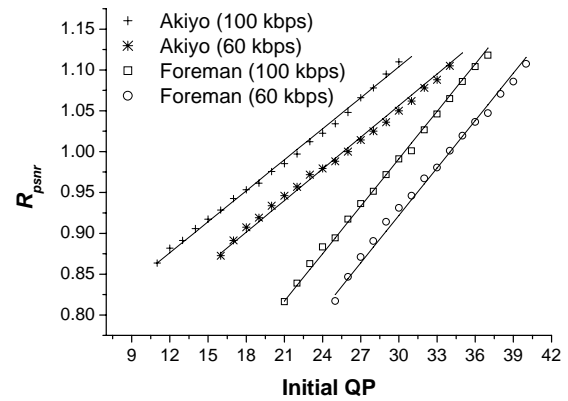


Fig. 3. Fitting accuracy of the linear model between the initial quantization parameter (QP) and peak signal to noise ratio for Akiyo and Foreman sequences with bit rates of 60 kbps and 100 kbps.

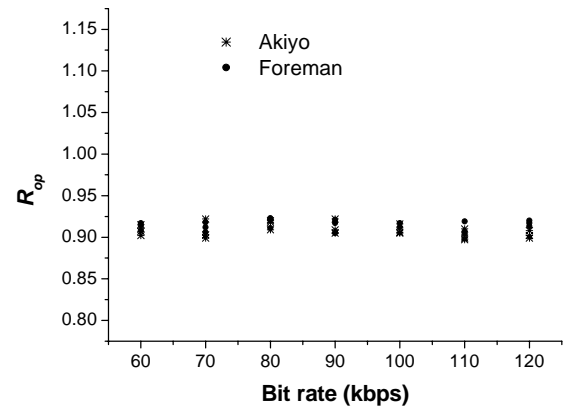


Fig. 4. Scatter plot of optimal quantization parameter (QP) ratio versus bit rate for Akiyo and Foreman sequences.

In the JVT rate control scheme, the QP for an I-frame depends on the average QP of the P-frames in the previous GOP as shown in Eq. (2). This initialization scheme is simple and adaptive to the available channel bandwidth,

but the initial QP converges to the optimal value very slowly. Also, it does not consider the characteristics of each video sequence. A more 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 sequence, such as the frame complexity and motion characteristics. However, the algorithm becomes more complex as the number of parameters is increased, and a 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 used in real-time applications.

Various test sequences have been encoded using different initial QP's in H.264/AVC, and PSNR characteristics of GOP's have been studied. As the initial QP decreases, the PSNR of the I-frame improves but that of the P-frame is degraded. This is because the I-frame consumes so many bits that there are not enough bits left for P-frames, which are encoded using the remaining bits. Let $PSNR_I(i)$ and $PSNR_P(i)$ denote the PSNR of the I-frame and the average PSNR of the P-frames of the i^{th} GOP. Based on the observations on a large number of benchmark video sequences, it is found that there is a linear relation between the initial QP and the ratio of $PSNR_I(i)$ and $PSNR_P(i)$. This linear relation can be formulated as

$$QP_i(1) = a \cdot R_{psnr}(i) + b \tag{4}$$

$$R_{psnr}(i) = PSNR_I(i) / PSNR_P(i)$$

where a and b are model parameters; $R_{psnr}(i)$ is the PSNR ratio of the i^{th} GOP. Fig. 3 shows the relation between the PSNR ratio and the initial QP for Akiyo and Foreman sequences with target bit rates of 60 kbps and 100 kbps. As can be seen from the figure, the PSNR ratio has a linear relation to the initial QP, but model parameters have different values according to the video sequence and the target bit rate.

The PSNR of a GOP varies with the change in the initial QP. When the initial QP has a small value, the entire PSNR of a GOP has a low value because of frame skip and buffer overflow. As the value increases, the entire PSNR also increases. Let QP_{op} denote the optimal QP, which maximizes the PSNR of a GOP. As the value increases beyond QP_{op} , the entire PSNR decreases. This is because the poor quality of the I-frame degrades the performance of the following intra coding.

Let R_{op} denote the PSNR ratio when the initial QP is QP_{op} . (QP_{op}, R_{op}) satisfies Eq. (4) so Eq. (4) is modified as follows:

$$QP_{op} = a \cdot R_{op} + b, \tag{5}$$

$$QP_i(1) = a \cdot \{R_{psnr}(i) - R_{op}\} + QP_{op}$$

Using the modified linear model, the proposed scheme determines the initial QP of the $i+1^{th}$ GOP as follows:

$$QP_{i+1}(1) = QP_i(1) - \{QP_i(1) - QP_{op}\} \\ = QP_i(1) - a \cdot \{R_{psnr}(i) - R_{op}\}, \quad i \geq 1 \tag{6}$$

In the proposed scheme, the first GOP of a sequence is encoded by the existing method, and from the second GOP, the initial QP's are determined by Eq. (6). However, there are two parameters a and R_{op} whose values are unknown. To estimate the value of a , the linear regression method is used. To apply the linear regression method, two or more data are needed. Thus, to estimate $QP_2(1)$ for the second GOP, the linear regression method cannot be applied. Thus, for the second GOP, the proposed scheme uses the value of 40 for slope a , which is determined from experimental observation. From the third GOP, a is determined by the linear regression method using $\{QP_i(1), R_{psnr}(i)\}$ pairs of the previous GOPs as follows:

$$a = \frac{\sum_{i=1}^N \{QP_i(1) - QP_{avg}\} \cdot \{R_{psnr}(i) - R_{avg}\}}{\sum_{i=1}^N \{QP_i(1) - QP_{avg}\}^2}, \tag{7}$$

$$QP_{avg} = \frac{\sum_{i=1}^N QP_i(1)}{N},$$

$$R_{avg} = \frac{\sum_{i=1}^N R_{psnr}(i)}{N},$$

where N is the number of encoded GOP's.

In Eq. (6), R_{op} can be set to a desired PSNR ratio. That is, if the PSNR of the I-frame is desired to have the same value of the average PSNR of the P-frame, R_{op} will be set to one. However, this setting will degrade the quality of the I-frame. The low quality of an I-frame will, in turn, degrade the entire quality of a GOP. Usually, R_{op} has a value of less than one. R_{op} can also be updated with the PSNR ratio of the GOP whose PSNR is greatest among encoded GOP's.

The characteristics of R_{op} have also been studied using different initial QP's. From extensive experiments, it has been found that R_{op} varies with the change of GOP size, but the impact of other parameters such as the target bit rate or video sequence on R_{op} is trivial. Fig. 4 shows the optimal PSNR ratios (R_{op}) of Akiyo and Foreman sequences where the GOP size is 30 and the target bit rate varies from 60 kbps to 120 kbps. From Fig. 4, it is shown that R_{op} is almost fixed for different sequences with different target bit rates. R_{op} values are around 0.92 regardless of the video sequence and the target bit rate. This means that the PSNR of a GOP is maximized when the average PSNR of P-frames is around

Table 1. Performance comparisons of the GOP level rate control algorithms in JVT-W057 and the proposed rate control algorithm in terms of average PSNR when the bit rates are 60, 80, and 100 kbps

Video	Bit rate (kbps)	JVT-W057								Proposed							
		2 nd GOP		3 rd GOP		4 th GOP		5 th GOP		2 nd GOP		3 rd GOP		4 th GOP		5 th GOP	
		PSNR	IQP	PSNR	IQP	PSNR	IQP	PSNR	IQP	PSNR	IQP	PSNR	IQP	PSNR	IQP	PSNR	IQP
Akiyo	60	37.40	38	37.14	36	38.22	34	38.44	32	39.44	28	39.85	24	40.93	22	41.64	21
	80	38.77	38	38.44	36	39.57	34	40.00	32	41.03	26	41.32	22	42.33	20	43.34	19
	100	40.10	38	39.55	36	40.77	34	40.84	32	42.13	24	42.49	20	43.32	18	44.48	17
Carphone	60	33.97	38	33.38	36	34.79	34	34.97	32	34.73	32	33.82	29	35.13	30	35.13	28
	80	35.23	38	34.75	36	35.93	34	36.06	32	36.12	30	35.13	27	36.47	28	36.42	26
	100	36.25	38	35.70	36	36.95	34	37.11	32	37.35	28	36.25	25	37.64	26	37.54	25
Foreman	60	31.78	38	31.00	36	31.80	34	32.19	34	32.65	32	31.05	30	31.80	32	32.32	31
	80	33.13	38	32.25	36	32.97	34	33.59	32	33.87	30	32.62	28	33.25	29	33.78	29
	100	34.03	38	33.17	36	33.90	34	34.56	32	34.79	29	33.76	27	34.24	28	34.74	27

GOP: group of picture, PSNR: peak signal to noise ratio, IQP: initial quantization parameter.

92% of the I-frame PSNR. For simplicity, R_{op} is set to a constant of 0.92 in the proposed algorithm when the GOP size is 30.

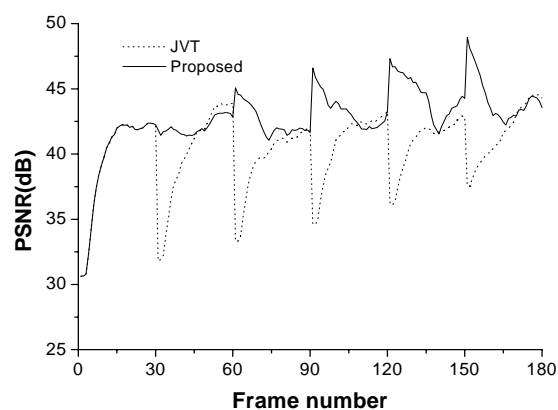
IV. EXPERIMENTAL RESULTS

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 a baseline profile. The results achieved here are compared with those achieved using the JVT-W057 rate control algorithm adopted by JM18.3.

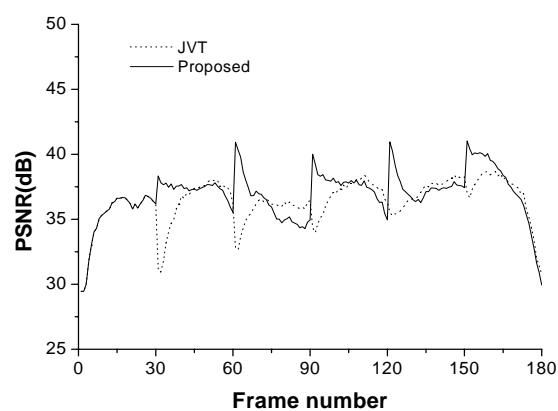
The same encoding parameters are used for both algorithms in order to ensure that the comparison is fair. For the experiments, the following test conditions are used: an “IPPPP...” GOP structure with a GOP size of 30 is used, the motion vector search range and the number of multiple reference frames 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 180 frames of three QCIF test sequences of Akiyo, Carphone, and Foreman. 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.

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 QOP is calculated and listed in Table 1 in order to provide an objective evaluation of the video quality. The proposed scheme uses the JVT algorithm for the first GOP, so the PSNR results of the first GOP's are not included in Table 1, where IQP denotes the initial QP. The proposed scheme shows better video quality than the rate

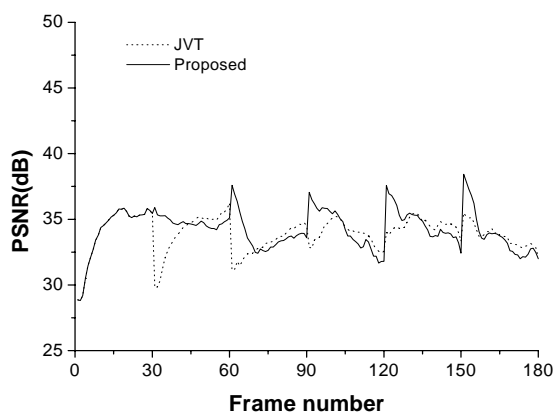
control of the JVT algorithm in terms of the average PSNR values.



(a)



(b)



(c)

Fig. 5. The peak signal to noise ratio (PSNR) results in using the group of pictures level rate control algorithm of JVT-W057 and the proposed algorithm for three video sequences with the bit rate of 100 kbps: (a) Akiyo, (b) Carphone, and (c) Foreman.

The frame-to-frame PSNR results of three sequences are shown in Fig. 5, where it is shown that better results are obtained by the proposed scheme than the JVT algorithm. Under the conditions of these simulations, the initial QP of the first GOP is set to 40 by Eq. (1), but this value is bigger than the optimal value, which maximizes the average PSNR value of a GOP. In the JVT algorithm, the initial QP value decreases by 2, so it takes several GOP's to reach the optimal value. On the other hand, it is shown that the proposed scheme can find the optimal value more quickly than the JVT algorithm.

The proposed scheme can also be applied to the scene change situation because the initial QP calculation by Eq. (1) is used for the first GOP after the scene change as well as the first GOP of a sequence. After scene changes, the proposed scheme can improve the visual qualities by finding the optimal initial QP more quickly.

V. CONCLUSIONS

In this paper, an adaptive PSNR-based initial QP determination algorithm for H.264/AVC is proposed. The proposed algorithm takes the characteristics of each video sequence into consideration by using the linear relation between the initial QP and the PSNR ratio, so 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 the Akiyo sequence, the proposed algorithm improves the average PSNR of GOPs up to about 2 dB.

REFERENCES

- [1] 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, no. 7, pp. 560-576, 2003.
- [2] S. C. Lim, H. R. Na, and Y. L. Lee, "Rate control based on linear regression for H.264/MPEG-4 AVC," *Image Communication*, vol. 22, no. 1, pp. 39-58, 2007.
- [3] Z. Li, F. Pan, and K. Pang, "Adaptive basic unit layer rate control for JVT," *Proceedings of Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-G012*, Pattaya, Thailand, 2003.
- [4] H. Wang and S. Kwong, "Rate-distortion optimization of rate control for H.264 with adaptive initial quantization parameter determination," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 1, pp. 140-144, 2008.
- [5] X. Jing and L. P. Chau, and W. C. Siu, "Frame complexity-based rate-quantization model for H.264/AVC intraframe rate control," *IEEE Signal Processing Letters*, vol. 15, pp. 373-376, 2008.
- [6] B. Yan and M. Wang, "Adaptive distortion-based intra-rate estimation for H.264/AVC rate control," *IEEE Signal Processing Letters*, vol. 16, no. 3, pp. 145-148, 2009.
- [7] K. P. Lim, G. Sullivan, and T. Wiegand, "Text description of joint model reference encoding methods and decoding concealment methods," *Proceeding of Joint Video Team (JVT) of ISO/IEC MPEG and TUT-T VCEG, JVT-W057*, San Jose, CA, 2007.



Sanghyun Park

received the 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 Suncheon National University, Suncheon, Korea, where he is currently an associate professor. His research interests include image processing, pattern recognition, and multimedia communication.