An Adaptive Rate Control Initialization Method for H.264 at Low Bit Rate

Jinbao Cai
Faculty of Foreign Studies, Jiangxi University of Science and Technology
Ganzhou, Jiangxi, China
Email: 12484788@qq.com

Yalin Wu *
* Department of Smart information systems, Jeonju University
Jeonju, Jeonbuk, Korea
Email: wuyalin@hanmail.net

***Abstract*—This paper presents an adaptive rate control initialization algorithm for H.264 at low bit rate. First we search the best initialization quantization parameters that used to generate a reconstructed video sequence with highest quality and reconstructed video sequence quality without extremely change of *GOP* (Group of Pictures) in different characteristic video sequence and different low bit rate. And then we propose a model to express the relationship of bit rate and best initialization quantization parameter. Moreover we use the proposed model to easily determine the best initialization quantization parameter of target bit rate. Compared with two noted algorithms, the measurement result of the objective *PSNR* and the subjective image quality of the proposed method has the best initialization quantization parameter.**

***Index Terms*—initialization quantization parameter, rate control, H.264, MAD, DSCQS**

I. INTRODUCTION

The supply of various multimedia services, such as pictures and videos, has been developed according to the recent development in digital signal processing, storage media and device, and transmission technologies including the wide distribution of computers and Internet. In particular, the application of video services has been significantly increased due to the coding technology for huge video data. The video coding standard for supplying various compression bitrates according to various communication networks and its terminals were established as MPEG-1, 2, 4, H.261, H.262, H.263, and H.264 [3].

The video coding can be carried out through a coding process in which it is essential to use a quantization algorithm. The quantization algorithm is classified into scalar and vector quantizers [4-5]. In an MPEG-2 coding method [6] or establishing a class of digital TV quality uses a scalar quantizer that has a quantization step-size from 1 to 32. In the case of the coding process established in the present time [4, 5], such as H.264 and MPEG-4 Part 10, it uses a total of 52 quantization step-sizes.

In addition rate control involves modifying the encoding parameters in order to maintain a target output bit rate. The most obvious parameter to vary is the quantiser parameter or step size (*QP*) since increasing *QP* reduces coded bit rate (at the expense of lower decoded quality) and vice versa. Simultaneously, the rate control initialization is a very importance section in the rate control strategy. It implements to select initialization quantization parameter for the first instantaneous decoding refresh (*IDR*) picture in a video sequence. In JVT-G012 [1], the initialization quantization parameter is decided by the bits per pixel (*BPP*). And *BPP* is determined according to bit rate, frame rate and frame resolution. However this method does not take into account that tested video sequence's feature and complexity. Therefore, we obtain the reconstructed video sequence is in low quality or reconstructed video sequence quality with extreme changing. In order to resolve existed problems in JVT-G012 [1], the method in [2] not only used *BPP* but also added two *MAD* values that are used to indicate the spatial and temporal characteristics of a video sequence to determine initialization quantization parameter. However, predicted mode is restricted and do not consider reconstructed video sequence quality balance.

In this paper, an adaptive method that can automatically decide the initialization quantization parameter is proposed. First we find the best initialization quantization parameters that used to generate a reconstructed video sequence with highest quality and reconstructed video sequence quality without extremely change of *GOP* (Group of Pictures) in different characteristic video sequence and different bit rate. And then we proposed a model to express the relationship of bit rate and best initialization quantization parameters. Moreover we used proposed model to easily find best initial quantization parameters of target bit rate.

Copyright credit, project number: YY1119, corresponding author: Wu Ya Lin.

doi:10.4304/jsw.8.10.2587-2593

This study consists of five sections. Section 2 shows existing method of initialization quantization parameter determination. Section 3 explains proposed method. Section 4 represents the results of the analysis of experiments. Finally, Section 5 shows the conclusion of this study.

II. Existing Method of Initialization Quantization Parameter Determination

At the beginning of the rate control stage, we have to carry out initialization process. In this stage, we should give an initialization quantization parameter *(QP)* for encoding the first intra frame. Process of initialization quantization parameter is manually decided. The value of initialization *QP* can be determined the quality of reconstructed video sequence. As a relative smaller bit rate, a small initialization quantization parameter is desired for the video sequence with complex spatial details and high motions like Bus and Flower video sequences. We can see the circumstance which is the first reconstructed frame brings about relative higher amount of bits and has higher quality of image. Due to this situation, other images of *GOP* are unable to assign to appropriate amount of bits and images of coarse quality appear in a *GOP*. TABLE Ⅰ and Ⅱ show this situation.

TABLE I.
Relationship of initialization quantization parameter and *PSNR* and difference of maximum and minimum quantization parameter in one *GOP* as bit rate is 0.4Mbps for Bus video sequence

Qp	PSNR	*bits*	*def Qp*
50	26.72	402400	12
49	26.81	403344	10
48	26.8	403272	9
47	26.8	403560	7
46	26.79	403912	7
45	26.9	403936	6
44	26.91	403832	5
43	26.92	404704	3
42	27.04	403224	3
41	27.09	410376	5
40	27.17	408448	8
39	27.22	406816	8
38	27.32	405824	9
37	27.28	407232	12
36	27.28	409912	13
35	27.24	412232	14
34	27.15	412472	17
33	27.06	414384	18
32	26.82	422176	19
31	26.49	431040	20
30	26.18	438184	21

TABLE II.
Relationship of initialization quantization parameter and *PSNR* and difference of maximum and minimum quantization parameter in one *GOP* as bit rate is 0.4Mbps for waterfall video sequence

Qp	*PSNR*	*bits*	*def Qp*
50	30.86	402512	17
49	30.91	402568	16
48	30.88	403000	15
47	30.96	402608	14
46	30.98	402520	13
45	30.93	403016	12
44	31.02	402696	11
43	31.14	402704	10
42	31.13	403088	9
41	31.24	403528	8
40	31.37	403416	7
39	31.50	404016	6
38	31.65	403984	5
37	31.87	405104	4
36	32.04	406288	3
35	32.25	407256	2
34	32.44	409152	1
33	32.70	409280	6
32	32.77	410496	5
31	33.04	407536	8
30	33.19	410144	14

In TABLE Ⅰ, as the value of initialization quantization parameter is decided to 30, the video sequences of *GOP* of higher quality cannot be obtained. Moreover, in this *GOP*, we can see the difference of maximum and minimum quantization parameter in one *GOP* is 21. This situation expounds reconstructed video sequence quality is not balance. However, value of initialization quantization parameter is decided to 38 and we can obtain video sequence of relatively higher quality and balance. TABLE Ⅰ shows this situation.

In TABLE Ⅱ, as the same bite rate with Bus video sequence, value of initialization quantization parameter is decided to 34, the video sequences of *GOP* of higher quality can be obtained. Moreover, in this *GOP*, we can see the difference of maximum and minimum quantization parameter in one *GOP* is 1. This situation expounds reconstructed video sequence have relatively higher quality and balance. TABLE Ⅱ shows this situation. However, for the beginner, it is too difficult to manually decide the value of initialization quantization parameter.

As the initialization *QP* is given as 0, it can be automatically determined according to *BPP*. So we review two methods which are used to decide initialization *QP* in recent literatures: the method of JVT-G012 [1] and the method of Wu [2]. The method of JVT-G012 [1] that can automatically decide an initialization *QP* is widely used in reference software of H.264 [1, 7].The method of JVT-G012 [1] uses only *BPP* to determine an initialization *QP*, while the method of Wu [2] uses the characteristics of video sequences as well as *BPP*.

In the method of JVT-G012 [1], an initialization *QP* can be determined according to *BPP* as follow:

$$initialQ_P = \begin{cases} 35 & BPP \geq L_1 \\ 25 & L_1 < BPP \leq L_2 \\ 20 & L_2 < BPP \leq L_3 \\ 10 & BPP > L_3 \end{cases}. \quad (1)$$

$$BPP = \frac{R}{f \times V_s}. \quad (2)$$

where $initialQ_p$ is an initialization *QP* value to encode the first I-frame. Initially, we work out the value of BPP through R , f and V_s being the bit rate, frame rate and the size of image of tested video sequence in (2). Secondly, L_1, L_2 and L_3 are the parameters. Moreover, the values of parameters are defined to 0.1, 0.3 and 0.6 for QCIF video sequences and 0.2, 0.6 and 1.2 for CIF video sequences and 0.6, 1.4 and 2.4 for other type of video sequences respectively in the H.264 reference software JM10.2 [8]. Although this initialization scheme is simple, yet it is not accurate enough.

In the method of Wu [2], the value of $initialQ_p$ is computed on the basis of *BPP*, the *MAD* value of the first *IDR* frame and the average *MAD* value of the second, third and fourth frames which are inter frames in video sequence as follows:

$$initialQ_P = F_1 + F_2\left(MAD_{IDR}, MAD_P\right). \quad (3)$$

$$F_1\left(BPP\right) = \begin{cases} b1 \cdot BPP + b2 & BPP \geq 0.4 \\ b3 \cdot BPP + b4 & 0.4 > BPP \geq 0.2 \\ b5 \cdot BPP^2 + b6 \cdot BPP + b7 & 0.2 \geq BPP \end{cases}. \quad (4)$$

$$\begin{aligned} F_2\left(MAD_{IDR}, MAD_p\right) = {} & C1 \cdot [\alpha \cdot MAD_{IDR} + (1-\alpha) \\ & \cdot MAD_p]^2 + C2 \cdot [\alpha \cdot MAD_{IDR} \\ & + (1-\alpha) \cdot MAD_p] + C3 \end{aligned}. \quad (5)$$

where the valve of α is 0.05 and θ = {*b1*, *b2*, *b3*, *b4*, *b5*, *b6*, *b7*, *C1*, *C2*, *C3*} is the parameter which can be given as θ = {-15.21, 25, -30.41, 31, 462.47, -189.86, 44.60, -0.78, 8.48, -10.90} for CIF video sequences and θ = { -12.67, 24, -25.34, 29, 546.59, -221.55, 46.41, 0.36, 1.30, -2.78} for QCIF video sequences. We can compute the value of *BPP* with (2) and calculate the value of $initialQ_p$ using (3).

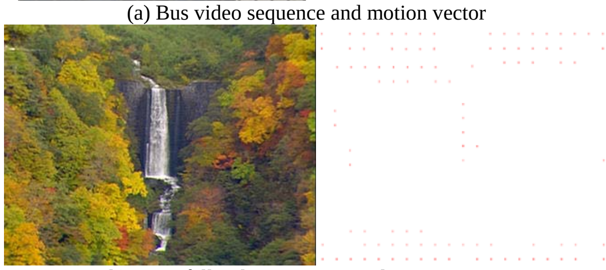
The parameters used in the method of Wu [2] are calculated by using News, Foreman and Mobile video sequences. Although they have different spatial and temporal complexities, it is difficult to say that they represent a large number of video sequences. Moreover, this method has not considered image quality balance of *GOP*.

III. Proposed Adaptive Method of Initialization Quantization Parameter Determination

A. Spatial and Temporal Characteristics of Video Sequence

Spatial and temporal characteristics of video sequence can be determined using motion vectors. Figure. 1 shows the motion vectors according to the characteristics of videos. The Bus video was made by moving a camera from right to left where motion vectors are presented at the central of the video. So we can see the complex spatial details and high motion in this video. And the waterfall video is gradually expanded at the contour of the video. So it shows simple spatial contents and low motions.

(a) Bus video sequence and motion vector



(b) Waterfall video sequence and motion vector

Figure 1. Motion vector for each video.

Obviously, two video sequences have different characteristic and complexity. In that way, given the same bit rate, the large initialization quantization parameter is decided for the video sequence with complex spatial and temporal characteristics. On the contrary, the small initialization quantization parameter should be used for the video sequence with simple spatial contents and low motion vectors. So we use method of JVT-G012 [1] to automatically decide initialization *QP* of Waterfall video is accurate.

On the contrary, we cannot gain an accurate value of initialization *QP* for Bus video by method of JVT-G012 [1]. Figure. 2 shows *QP* values of all pictures in one *GOP* as 0.8MBps bit rate for the Bus video and Waterfall video.

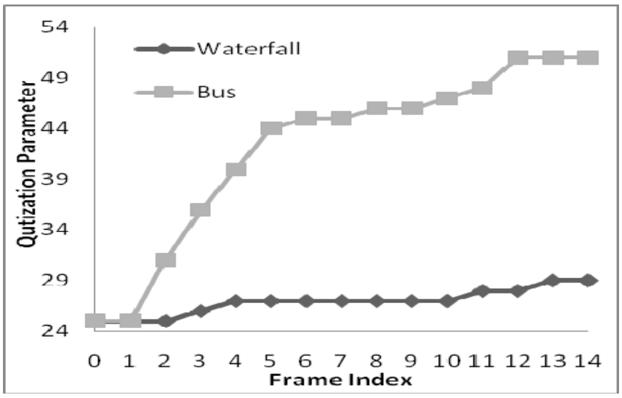


Figure 2. *QP* values of all pictures in one *GOP*.

As a result, we can see the quality of Waterfall reconstructed video sequence is high and balance. Due to difference between the maximum and the minimum values of *QP*s of all pictures in *GOP* is 3. However, the difference is 25 in bus video sequence. Quality of Bus reconstructed video sequence is not equilibrium and gruff.

In [2], although values of two *MAD* are added to indicate the spatial and temporal characteristics of video sequence, restricted predicted mode lead to influence quality of reconstructed video sequence. Figure. 3 shows difference of *MAD* as the restricted predicted mode is used and all modes are used for Bus video sequence.

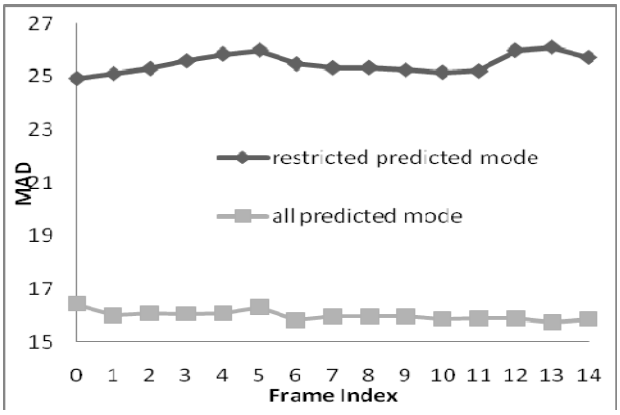


Figure 3. *MAD* values of all the I frames in one *GOP* for Bus video sequence.

The smaller value of *MAD* represents the higher quality of reconstructed video sequence.

B. Model of Best Initialization Quantization Parameter and Bit Rate

In order to effectively and accurately decide the initialization quantization parameter for video sequence. In the first place, we find the best initialization quantization parameter as 0.4, 0.5, 0.6, 0.7 and 0.8Mbps bite rate for Bus and Waterfall video sequence through experiments.

The target of sample video sequences is the optimal initialization *QP* at the given target bit rate. Therefore, the selection method for the optimal initialization *QP* is introduced.

On the basis of *RC* algorithm of H.264 [1], the initialization *QP* and target bit rate bear direct relevance for performance of encoding. The strategy of this proposed method is that the reconstructed video sequences have consistent and superior quality and the bit rate closer to target bit rate in various tested target bit rates by the optimal initialization *QP*. In order to realize this algorithm, primarily, we obtain average *PSNR* and bit rate of front 60 frames of testing sample video sequences and the differences of *QP*s in a *GOP* of 52 initialization *QP*s of given the target bit rate. The *PSNR*, bit rate and differences of *QP*s represent picture quality, amount of data and stationary quality of a *GOP*. Therefore, we can calculate the optimal initialization *QP*s of given the target bit rate of all sample videos according to pick up the specific initialization *QP* that can be used to generate the maximum *PSNR* and minimums of bit rate and differences of *QP*s. However, the *PSNR*, bit rate and difference of *QP* are not same magnitude. As a result, *PSNR*, bit rate and difference of *QP* should be respectively normalized. The process of normalization is expressed as follows:

$$NPSNR_{initial_{Qp}} = \frac{PSNR_{initial_{Qp}} - \underset{initial_{Qp}=0,\cdots,51}{MIN}(PSNR_{initial_{Qp}})}{\underset{initial_{Qp}=0,\cdots,51}{MAX}(PSNR_{initial_{Qp}}) - \underset{initial_{Qp}=0,\cdots,51}{MIN}(PSNR_{initial_{Qp}})},$$
$$Initial_{Qp} = 0,\cdots,51. \quad (6)$$

$$NBIT_{initial_{Qp}} = \frac{BIT_{initial_{Qp}} - \underset{initial_{Qp}=0,\cdots,51}{MIN}(BIT_{initial_{Qp}})}{\underset{initial_{Qp}=0,\cdots,51}{MAX}(BIT_{initial_{Qp}}) - \underset{initial_{Qp}=0,\cdots,51}{MIN}(BIT_{initial_{Qp}})},$$
$$Initial_{Qp} = 0,\cdots,51. \quad (7)$$

$$NDIFFQp_{initial_{Qp}} = \frac{DIFFQp_{initial_{Qp}} - \underset{initial_{Qp}=0,\cdots,51}{MIN}(DIFFQp_{initial_{Qp}})}{\underset{initial_{Qp}=0,\cdots,51}{MAX}(DIFFQp_{initial_{Qp}}) - \underset{initial_{Qp}=0,\cdots,51}{MIN}(DIFFQp_{initial_{Qp}})},$$
$$Initial_{Qp} = 0,\cdots,51. \quad (8)$$

where $NPSNR_{initial_{QP}}$, $NBIT_{initial_{QP}}$ and $NDIFFQp_{initial_{QP}}$ are normalized PSNR, bit rate and differences of *QP*. $PSNR_{initial_{QP}}$, $BIT_{initial_{QP}}$ and $DIFFQp_{initial_{QP}}$ are the values that are before normalized. Afterwards, the selection algorithm of the optimal initialization *QP* is designed as follow:

$$BestInitial_{Qp} = \underset{initial_{Qp}=0,\cdots,51}{\arg\min}\left(\frac{1}{NPSNR_{initial_{Qp}}} + NBIT_{initial_{Qp}} + NDIFF_{initial_{Qp}}\right). \quad (9)$$

where the $BestInitial_{QP}$ is the Optimal initialization *QP* of given target bit rate of sample videos.

Next, we find the relationship of bit rate and best initialization *QP* in Figure. 4.

And then we can find a model to express the relation in (10).

$$FinalInitial_{Qp} = \lceil I_a \times R^2 - I_b \times R + I_c \rceil \ . \tag{10}$$

where R is target bit rate, and the parameters of I_a, I_b and I_c can be given as 42.85, -75.43 and 57.570 for Bus and 21.43, -34.71 and 48.286 for Waterfall video. So we can use (10) to calculate the best $FinalInitial_{Qp}$ form 0.4Mbps to 0.8Mbps.

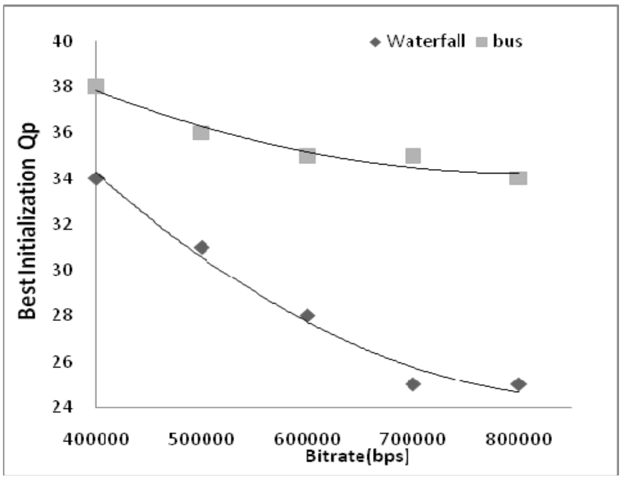


Figure 4. Relationship of bit rate and best initialization *QP*.

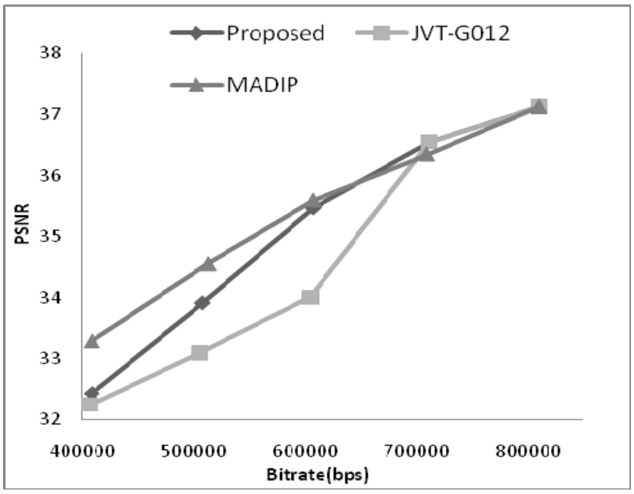
IV. Experimental Results

The JM10.2 [8] that is a standard coding software tool of the H.264 was used to the experiment for the method proposed in this study where Bus and Waterfall videos with horizontal and vertical resolutions of 352 and 228 pixels, respectively, were selected.

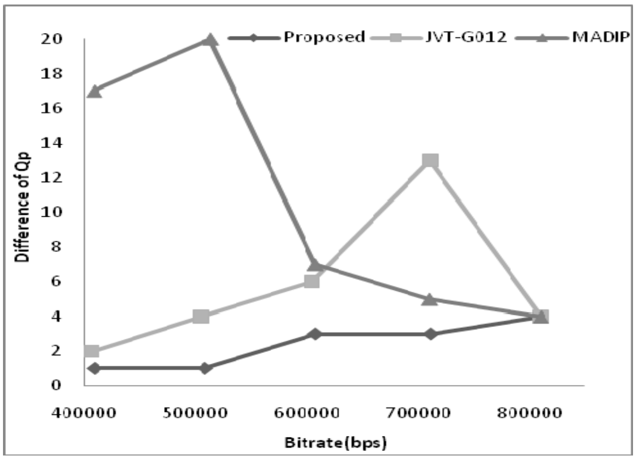
The B-picture is not included due to the use of the H.264 baseline porfile, and 15 pictures are configured as one *GOP* in which each video applies 30 pictures. The number of slices are the same as for each picture and that are determined by 18 along the vertical direction. The test target bit rates (units: mbps) are 0.4, 0.5, 0.6, 0.7, 0.8. Each of the 2 video sequences is intra coded for the first frame and followed with subsequent inter and intra coded with all prediction modes.

A. Measurement of the Objective PSNR

This study investigated the *PSNR* (peak signal to noise ratio) that is an objective picture quality criterion. In Figure. 5(a), we can see that the method [2] achieves better *PSNR* performance than proposed method and JVT-G012 [1] method for Waterfall video. However, in Figure. 5(b), we find the proposed method gain the better balance quality of reconstructed video sequence in *GOP*. Due to the smaller *QP* difference, achieve the better balance quality.



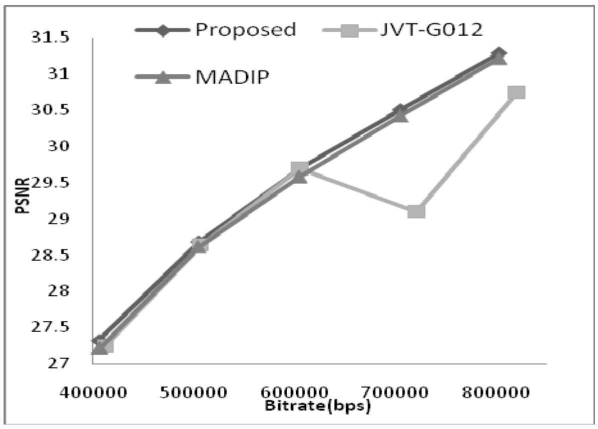
(a) Relationship of bit rate and *PSNR*



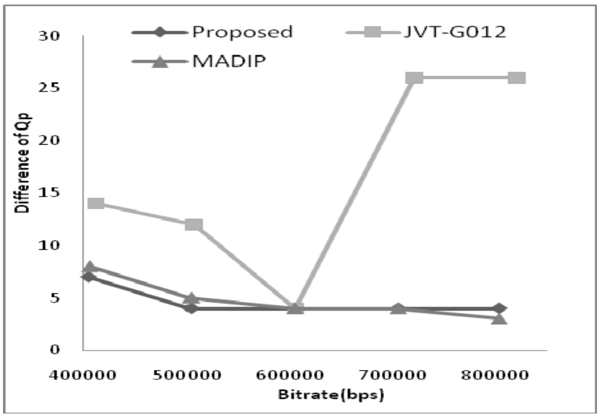
(b) Relationship of initialization *QP* and difference initialization *QP*

Figure 5. Waterfall video

In Figure. 6 (a), we can see that the proposed method [2] achieves better *PSNR* performance and better balance quality of reconstructed video sequences in *GOP* than JVT-G012 [1] and method [2].



(a) Relationship of bit rate and *PSNR*



(b) Relationship of initialization *QP* and difference initialization *QP*

Figure 6. Bus video

In brief, because of the method of JVT-G012 [1] does not consider the characteristic and complexity of tested video sequence; better quality of reconstructed video sequence is not achieved. Although method [2] considers the spatial and temporal characteristics of video sequence and we use it and can gain the better quality of video sequence, achieved quality reconstructed video sequence is not balance.

B. Measurement of Subjectively Concerned Picture Quality

In this paper, we used DSCQS (Double Stimulus Continuous Quality Scale) method to measurement of subjectively concerned picture quality [7]. According this method, the assessors can observe two videos that one is an original video and the other is a processed video by existing method or proposing method at the monitor. Assessors evaluated pictures quality of both videos using an ITU-R quality scales (Excellent=5, Good=4, Fair=3, Poor=2, Bad=1) [9].

TABLE III.
RESULTS OF THE SUBJECTIVELY CONCERNED PICTURE QUALITY

Video Name	Proposed method	Existing method	
		JVTG012 [1]	Method [2]
Water fall	3.8	3.4	1.2
Bus	3.6	3.0	3.6

The presentation order of original and process videos was random. The assessors are asked to evaluate the videos and sequence. And assessors are distanced from the video as much as three times of the diagonal length of the monitor used in the play of videos. Final score of subjectively picture quality of assessed video was calculated as follows：

$$U = \frac{1}{N}\sum_{i=1}^{N} u_i \ . \qquad (11)$$

where u_i is score which is determined by each assessor, N is quantity of assessor.

In the results, we know the proposed method gains the best grade in three methods. Due to this we use proposed method to achieve the reconstructed video sequence with highest quality and reconstructed video sequence quality without extremely change of *GOP*. Figure. 7 shows worst image quality of reconstructed video sequence in *GOP* with three different methods for Waterfall video as the bit rate is 0.4Mbps.

As the result we can see that the quality of image of Figure 7(c) is too crude to clearly see the contents of image. However, in Figure 5(a), it achieves the best *PSNR* of three methods.

This due to *PSNR* is an average value of 30 pictures in video sequence. In first 15 Images, from number 0f 0 to 10, the quantization parameter range is from 29 to 32. However from number of 11 to 14, the value of quantization parameter is 46.

As we measure the quality of video, we should use the objective *PSNR* and subjectively concerned picture quality.

(a) Quantization Parameter is 31 by proposed method

(b) Quantization Parameter is 34 by JVT-G012 [1] method

(c) Quantization Parameter is 34 by method [2]

Figure. 7. Worst image quality of reconstructed video sequence in *GOP*

V. CONCLUSIONS

In this paper, simple and effective rate control initialization method which decides initialization quantization parameter according to the model that expresses the relationship of bit rate and best initialization quantization parameters. Moreover we can apply this model to decide initialization quantization parameter for rate control as target bit rate range from 0.4Mbps to 0.8Mbps. In the experiment there are two different types of video sequences to be tested. Finally we use the measurement of the objective *PSNR* and the subjective image quality of result presented that the proposed method can achieve better quality of video sequences performance with the bit rate constraint than JVT-G012 [1] and existing method [2]

ACKNOWLEDGMENT

This paper is financially supported by The Education Department Of Jiangxi Province. The title is “A Study on the Construction of English Corpus Based on the Export-oriented Development of Metallurgical Industry”. The grant number is YY1119.

REFERENCES

[1] Z. Li, F. Pan, K. P. Lim, G. F, X. Lin, and et al. “Adaptive Basic Unit Layer Rate Control for JVT”, *Doc. JVT-G012, Thailand*, Mar. 2003.

[2] W. Wu, and H. K. Kim. “A Novel Rate Control Initialization Algorithm for H.264”, *Consumer Electronics, IEEE Transactions on*, vol. 55, no. 2, pp. 665-669, 2009.

[3] S. K. Kwon, A. Tamhankar, and K. R. Rao, “Overview of H.264/MPEG-4 Part 10”, *Journal of Visual Communications and Picture Representation*, vol. 17, no. 2, pp. 186-216, 2006.

[4] ITU-T Rec. H.264 / ISO/IEC 11496-10, “Advanced Video Coding”, *Final Committee Draft, Document JVT-F100*, Dec 2002.

[5] T. Widegand, G. Sullivan, G. Bjonttegaard, and A. Luthra, “Overview of the H.264/AVC Video Coding Standard”, *IEEE Trans. Circuit Syst. Video Technology*, vol. 3, no. 7, pp. 560-576, 2003.

[6] MPEG-2: ISO/IEC JTC1/SC29/WG11 and ITU-T, ISO/IEC 13818-2, Information Technology_Generic Coding of Moving Pictures and Associated Audio Information: Video, *ISO/IEC and ITU-T International Standard Draft*, 1994.

[7] ITU-R Rec. BT. 500, "Methodology for the subjective assessment of the quality of television pictures", *Geneva*, Jun 2002.

[8] JM Reference Software Version 10.2, http:// iphome.hhi.de/suehring/download.

[9] M. Bernas, “Image Quality Evalution”, *Video/Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom*, pp.133-136, 2002.

Cai Jinbao is an associate professor, graduated from University of Plymouth (UK), is working in Jiangxi University of Science and Technology (P.R.China). He devotes to linguistic teaching &research and the information processing of natural language. Now he is working with some researchers to do the study on the construction of metallurgical English-corpus based on the computer knowledge and the study on the pattern recognition, etc.

Wu Ya Lin completed B.S, M.S at the Department of Computer Software Engineering of Dongeui University In 2008 and 2010. Now he has been a doctor degree candidate at Jeonju University, Jeonju, Republic of Korea. His research interests are in the areas of image processing, video coding, and video transmission.