PAPER

Adaptive Rate Control Mechanism in H.264/AVC for Scene Changes

Jiunn-Tsair FANG[†a], **Zong-Yi CHEN**[††b], **Chen-Cheng CHAN**[††c], ***Nonmembers***, ***and*** **Pao-Chi CHANG**[††d], ***Member***

SUMMARY Rate control that is required to regulate the bitrate of video coding is critical to time-sensitive video applications used over networks. However, the H.264/AVC standard does not respond to scene changes, and this causes the transmission quality to deteriorate as a scene change occurs. In this work, a scene change is detected by comparing the ratio of the sum of absolute difference (SAD) between two consecutive frames. As the scene change is detected, the proposed method, which is modified from the reference software of H.264/AVC, re-assigns a quantization parameter (QP) value to regulate the bitrate. Because the inter-prediction works poorly for the scene-changed frame, the proposed method estimates its frame complexity based on the content, and further creates another Q-R model to assign QP. The adaptive rate control mechanism presented in this study can quickly respond to the heavy bitrate increment caused by a change of scene. Simulation results show that the proposed method improves the average peak signal noise ratio (PSNR) to approximately 1.1 dB, with a smaller buffer size compared with the performance of the reference software JM version 17.2.

key words: *H.264/AVC, rate control, scene change, target bits, target buffer level.*

1. Introduction

The H.264/AVC standard [1] has been applied to various consumer electronics, such as smart phones, video conferences, and IPTVs. Because of this diversity of application, the need for the rate control to meet several requirements has increased. The rate control of H.264/AVC provides adequate strategies to maintain transmission quality. Its bitrate regulation is mainly based on the available bandwidth and a predefined buffer size [2], [3].

The rate control of H.264/AVC allocates a bitrate budget to each of a group of pictures (GOP) [2], [3]. For each frame within a GOP to distribute the bitrate budge, the rate control is usually required to estimate its target bits and frame complexity, and then to assign a proper quantization parameter (QP) value for the current encoding frame. The QP value is calculated based on a quadratic rate-distortion (R-D) model [4], and the QP assignment directly affects the bitrate and frame quality. Thus, it is an essential procedure for the rate control to determine the QP value for the current encoding frame [5].

The H.264/AVC standard has no particular response to a scene change. As a scene change occurs, the estimation of its frame complexity is no longer accurate. The inaccurate estimation of the frame complexity causes the quadratic R-D model to fail to assign an appropriate QP value. Furthermore, as the scene change occurs, more bits are created than the target bits estimation causing a deterioration in the quality of succeeding frames, and the buffer to overflow. Several methods to detect scene changes have been explored. The histogram-based detection method [6], the gray value-based detection method [7], and the intra macroblock-based detection method [8] use various methods to measure the dissimilarity between two consecutive frames.

For real-time video applications, the detection method must be fast and low in complexity. Lee et al. calculated the peak signal to noise ratio (PSNR) difference between two consecutive frames, and then compared the result with the average PSNR value of previously encoded frames within the same GOP [11]. Their method is simple but may cause errors. Ding and Yang utilized the sum of absolute transformed difference (SATD) for scene change detection [10]. SATD can be applied to predict the bitrate more precisely than the SAD, but it increases the coding complexity. In this work, the SAD ratio (SADR) between two consecutive frames was used as the criterion of scene change detection. The SAD is an existing coding parameter in the inter-prediction procedure, and therefore calculating the SADR increases with little complexity.

Several methods have been proposed for modifying JM, the reference software of H.264/AVC, to improve the rate control performance as a scene change is detected. An effective method is to create a transition GOP for the scene-changed frame and the remaining frames in the original GOP [9], [12]. This method works because the content of the scene-changed frame is dissimilar from the contents of its previous frames, and thus most of the blocks in the scene-changed frame are encoded using intra-coding.

A procedure is needed to assign a QP value to the scene-changed frame, and the JM method can be applied to determine the assignments for the rest of P frames of the transition GOP. Lee et al. selected the QP value from the intra-frame of the first GOP or the previous GOP [11]. However, the content between the scene-changed frame and its previous frames is no longer similar. As for the estimation of the content complexity, Jing et al. applied the sum of the gradient values of pixels to represent the complex-

[†]The author is with Department of Electronic Engineering, Ming Chuan University, Taoyuan, Taiwan.
[††]The authors are with the Department of Communication Engineering, National Central University, Zhongli, Taiwan.
a) E-mail: fang@mail.mcu.edu.tw
b) E-mail: zychen@vaplab.ce.ncu.edu.tw
c) E-mail: airteanet@yahoo.com.tw
d) E-mail: pcchang@ce.ncu.edu.tw
DOI: 10.1587/transfun.E0.A.1

ity of an intra-frame [13]. Tsai and Chou modified their method to formulate a rate-quantization (R-Q) model [14]. However, their model is only applicable for all intra-frames sequences. As for the target bits estimation, Lee et al. modified the JM parameters using a ratio of intra-coded bits to inter-coded bits [9]. They also classified the complexity of sequences into three levels to respectively assign various QP values. Recently, Chen and Liu applied a scene change factor (SCF) to estimate the frame complexity, and the target bits estimation was classified into five levels based on the SCF value [12]

In this work, some experiments were designed to illustrate the influence of a scene change to the H.264/AVC performance. The proposed method was to reinforce this unaddressed part of H.264/AVC, and was performed by two steps. The first step was for the scene change detection, and the second step was for the rate control mechanism realized by QP adjustment. Two schemes were applied for the QP assignment according to the proposed rate control mechanism. One was for the empty buffer condition, and the other was for the non-empty buffer condition. Our previous work [22], represented the preliminary result for the non-empty buffer condition in this manuscript. Compared with the work [22], this work proposed an efficient scene change detection method and a complete rate control mechanism.

In Section II of this paper, the influence of the scene change on the rate control performance of the H.264/AVC is described. Section III illustrates the proposed methods for scene change detection and for the rate control mechanism. The simulation results are shown in Section IV. Finally, the conclusion is presented in Section V.

2. SCENE CHANGE INFLUENCE ON THE RATE CONTROL PERFORMANCE

Before discussing the influence of the scene change on the rate control performance, the rate control algorithms of H.264/AVC, focusing on the related formula of the QP assignment, are reviewed. Then, the test sequence Trevor is encoded using JM. Trevor contains 100 frames, and has a scene change at the 59th frame. The impact of the scene change on the rate control performance for the rest of the frames is shown and discussed. In particular, the influences on the buffer management, QP assignment, and PSNR performance are illustrated.

2.1 REVIEW OF RATE CONTROL ALGORITHMS OF H.264/AVC

Rate control refers to maintaining the stability of the transmission quality. It regulates bitrates to prevent the buffer from overflow and underflow [15], [16]. From the rate control algorithm of H.264/AVC, the QP assignments to the I frame and the P frame are different. For the I frame, the QP assignment mainly depends on the channel bandwidth and the number of pixels in a frame [2], [3]. Its assignment is independent of the content of the frame. However, for the P frame, the estimations of the target bits and the frame complexity are needed.

According to the frame layer rate control of H.264/AVC, the estimation of the target bits for an IPPP...P sequence is a weighted combination of two parameters [2], [3]. The first parameter is based on the available channel bandwidth, the frame rate, the target buffer level, and the actual buffer occupancy. In the frame layer of the rate control, the channel bandwidth and the frame rate have already been determined. This parameter is thus mainly determined by the target buffer level and the actual buffer occupancy. Formula (1) shows that the first parameter modifies the estimation difference of the buffer level from the previous frame:

$$\tilde{f}(n) = \frac{u(n)}{F_r} + \gamma(Tbl(n) - ABF(n)), \tag{1}$$

where $u(n)$ represents the available channel bandwidth for the nth frame, F_r is the frame rate, Tbl is the target buffer level (TBL), ABF is the actual buffer fullness (ABF), and γ is a constant number [2], [3]. TBL is the expected value of the buffer fullness, and is always set to reduce a fixed size of bits after encoding a frame, except the I and the first P frames [17]. After encoding each frame, the ABF is compared with the TBL. If the TBL is overestimated, which means that more bits are encoded than expected, the target bits should be increased for the succeeding frame. If it is underestimated, the target bits should be reduced for the succeeding frame. Formula (1) shows how to adjust the estimation of the target bits based on the TBL.

The second parameter of the target bits estimation is to modify the estimation of the coding bits from the previous frame. This parameter is determined by the remaining bits after encoding the previous frames within the same GOP. The parameter is derived in (2):

$$\hat{f}(n) = \frac{B(n)}{m - n}, \tag{2}$$

where m, $B(n)$, and $(m\text{-}n)$ are the number of frames, the number of remaining bits, and the un-coded frames in the current GOP, respectively. Formula (2) shows that the remaining bits in the current GOP are equally distributed to the un-coded frames. Finally, the target bits are a weighted combination of these two parameters, as described in (3):

$$f(n) = \beta \times \hat{f}(n) + (1 - \beta) \times \tilde{f}(n), \tag{3}$$

where β is set to 0.5 using JM.

In addition to the target bits estimation, the frame complexity must also be estimated. For the inter-frame prediction, H.264/AVC provides variable block sizes from a 16×16 sample region to a 4×4 sample region for the motion estimation. Motion estimation refers to finding a block in the reference frames that closely matches the current block. After searching for the predicted block, each pixel in the current block is subtracted from the corresponding pixel of the predicted block. The residual values of pixels form a residual block, which is to be encoded or transmitted. The

SAD for the current block is the sum of the absolute value of each pixel in its residual block. The SAD for the current block is represented in (4):

$$SAD_{cb} = \sum_{i=0}^{N_1-1} \sum_{i=0}^{N_2-1} |a_{i,j} - \hat{a}_{i,j}|, \tag{4}$$

where $a_{i,j}$ and $\hat{a}_{i,j}$ are the (i, j) pixels of the current block cb and the predicted block, respectively, and N_1 and N_2 are the number of pixels of the block in the x-axis and the y-axis, respectively.

The mean absolute difference (MAD) can be applied to estimate the frame complexity. The MAD in the current block is an average value of the SAD_{cb} value, and is represented in (5):

$$MAD_{cb} = \frac{SAD_{cb}}{N_1 N_2}, \tag{5}$$

However, without the QP value, the SAD cannot be calculated. H.264/AVC thus applies a linear prediction model to predict the MAD of the current block [18], as illustrated in (6):

$$MAD_{cb} = a_1 \times MAD_{pb} + a_2, \tag{6}$$

where a_1 and a_2 are coefficients of the prediction model, and MAD_{pb} is the actual MAD value of the co-located block in the previous frame. Furthermore, MAD is an average value of total blocks in the current frame. It is illustrated in (7):

$$MAD = \frac{1}{N} \sum_{N} MAD_{cb}, \tag{7}$$

where N is the number of blocks in the current frame. Finally, the QP assignment is based on a quadratic rate-distortion model [4]. Formula (8) is associated with target bits and frame complexity:

$$f = C_1 \times MAD/QP + C_2 \times MAD/QP^2, \tag{8}$$

where C_1 and C_2 are constant. Furthermore, to maintain the stability of the video, the maximal difference of the QP value between two consecutive frames is restricted to 2. The related work on the rate control algorithms is covered in [8]-[12].

2.2 SCENE CHANGE INFLUENCE ON THE RATE CONTROL PERFORMANCE

The rate control of H.264/AVC has no particular response to the scene change. During the scene change, the content of the scene-changed frame differs from the content of its previous frames. The content difference causes the inaccurate estimation of the current frame complexity, which fails to satisfy the formula of the quadratic R-D model. As a consequence, using JM to assign QP to the scene-changed frame is probably not correct. Furthermore, as the scene change occurs, more bits are created. The ABF is thus much higher than the expected TBL, which results in the reduction of the target bits allocation. A larger QP value thus should be assigned to the next frame. However, the QP clipper restricts the range of QP adjustment, and several succeeding frames have to sacrifice their target bits to absorb these extra bits. As a result, the scene change deteriorates not only the quality of the scene-changed frame, but also the quality of its succeeding frames. It also adds to the risk of buffer overflow.

To show the influence of the scene change on the rate control performance, the test sequence Trevor was used as an example. Simulations were executed with JM of version 17.2.The channel bandwidth was 128 K, and all frames were in quarter common intermediate format (QCIF). The test sequence contained 100 frames and a scene change occurred at the 59th frame. The TBL and ABF were plotted for each frame. Fig. 1(a) shows that the ABF was much higher than the TBL at the 59th frame, lasting for 7-8 frames.

The QP and PSNR values from each frame are plotted in Fig. 1(b). It shows that the QP value started to increase after the 59th frame was encoded, which resulted in the drop of the PSNR performance. The succeeding frames after the scene change were with lower motions. They should be assigned smaller QP values to obtain higher PSNR performance. However, their PSNR values dropped. Furthermore, the QP clipper restricts the QP adjustment. It takes more than 10 frames for the QP adjustment to maintain stable PSNR performance.

Fig. 1(b) also shows that the QP value assigned to the first frame, the initial I frame, was too low. This is because the QP assignment mainly depends on the channel bandwidth and the number of pixels in a frame. H.264/AVC does not consider the complexity of the frame content. The weakness of JM during a scene change was illustrated. In the next section, JM is modified and another R-Q model is developed to manage scene changes.

3. SCENE CHANGE DETECTION AND ADAPTIVE RATE CONTROL MECHANISM

The influence of a scene change on the rate control performance is mainly because of the heavy extra bits that are generated when scene-changed frames are encoded. The proposed method for scene change detection is to identify the content dissimilarity between two consecutive frames. As a scene change is detected, JM is modified to absorb the extra bits.

3.1 SCENE CHANGE DETECTION

The scene change refers to content that significantly changes between two consecutive frames. A scene change can be the result of the fast motion of a recorded object, camera movement, frame concatenation, or fade-ins or fade-outs. It is not easy to define the scene change appropriately [21], not to mention the detection. This work focused on rate

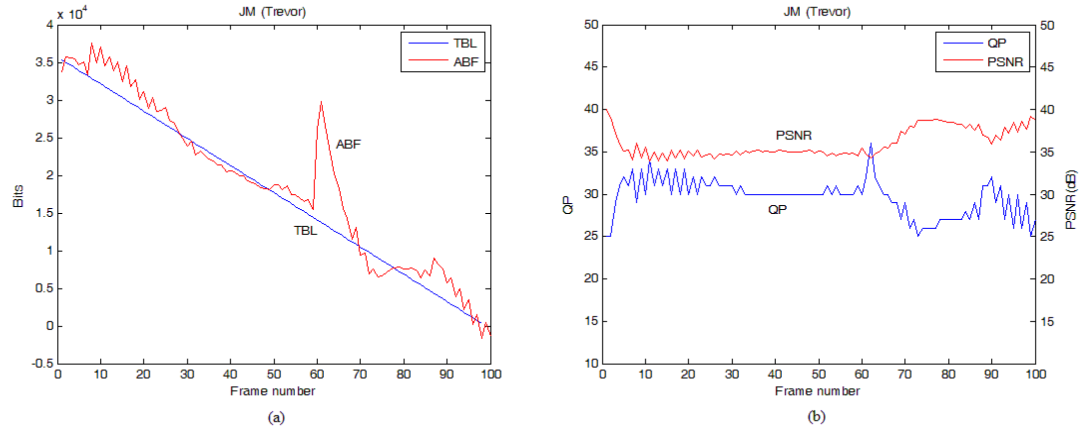

Fig. 1 Scene change influences on JM by the sequence Trevor: (a) ABF and TBL, (b) QP and PSNR.

control performance, which is significantly affected by the sudden heavy bit generation. It is necessary for H.264/AVC rate control to have an adaptive mechanism to regulate the bitrate. Instead of detecting the scene-changed frames, this work focused on detecting the frames producing heavy bits.

SADR between two consecutive frames was used to measure the dissimilarity of the temporal correlation. The SAD for the current block can be calculated with (4) and the total SAD value for the current frame can be calculated with (9):

$$SAD = \sum_{N} SAD_{cb}. \tag{9}$$

The SADR is the ratio of the total SAD value of the current frame divided by its previous frame. This relationship is represented in (10):

$$SADR = \frac{SAD_i}{SAD_{i-1}}, \tag{10}$$

where i is the current frame. To reduce the complexity, only a 16×16 luminance block for inter-prediction was applied.

Calculating the SADR is an additional procedure. However, it does not increase the encoding complexity if no scene change is detected because the SAD calculation is contained within the inter-frame prediction procedure. In the event that a scene change is detected, the SADR calculation does increase the complexity. As a result, the complexity increases slightly to the proposed method.

3.2 ADAPTIVE RATE CONTROL MECHANISM

As a scene change is detected, the scene-changed frame and the remaining frames of the original GOP are treated as a transition GOP. This is because the content of the scene-changed frame differs from the content of its previous frames, and the parameters in (6) and (8) are no longer applicable. A procedure is then needed to assign a QP value to the scene-changed frame, which is encoded as an I frame, and the JM method can be applied to determine the assignments for the rest of P frames of the transition GOP.

The transition GOP and the ordinary GOP differ in two respects. First, the length of the transition GOP is variable, whereas it is fixed for an ordinary GOP. The location of the scene change affects the target bits estimation and buffer management. Second, the buffer level in the ordinary GOP is always empty at the start, whereas this is not the case in the transition GOP. The non-empty buffer level also influences the target bits estimation. The rate control mechanism accounted for these two differences.

To assign the QP value to the scene-changed frame, two steps were followed. First, JM was modified for the QP assignment to the I frame of an ordinary GOP. Second, an R-Q model was developed for the QP assignment to the scene-changed frame of a transition GOP.

3.2.1 QP initialization for an ordinary GOP

According to JM, the QP initialization for a GOP is based on the channel bandwidth, frame rate, and the frame size [2], [3]. Various sequences are assigned to the same QP value. Tsai and Chou thus proposed a new R-Q model [14], as shown in (11):

$$\frac{B(0)}{m} = a \times G \times QS^{b}, \tag{11}$$

where a and b are constant, G is the average gradient per pixel, QS is the quantization step size, and $B(0)$ and m are the bit budget and the number of frames for a GOP, respectively. The average gradient per pixel, G, is defined as

$$G = \frac{1}{H \times V} \sum_{i=0}^{H-1} \sum_{j=0}^{V-1} |I_{i,j} - I_{i+1,j}| + |I_{i,j} - I_{i,j+1}|, \tag{12}$$

where H and V are the horizontal and vertical sizes of the frame, respectively, and $I_{i,j}$ is the luminance value of the pixel at location (i, j) [14]. Similar to the quadratic rate-distortion model [4], (12) also connects the target bits; QP, which the quantization step size expresses; and frame complexity, which the average gradient per pixel estimates. The

Table 1 The QP initialization for JM and our method

Sequence	Method	Gradient value (I frame)	QP value (I frame)
Soccer	JM		25
	Ours	15.2	26
Carphone	JM		25
	Ours	13.7	25
Forman	JM		25
	Ours	14.6	26
Akiyo	JM		25
	Ours	9.57	21
Football	JM		25
	Ours	30.3	34

target bits can be equally distributed to each frame, because (12) is only applicable to all I-frame sequences.

The average gradient per pixel was applied to estimate the frame complexity. However, the first I frame in the IPPP...P sequence requires more target bits than the P frames. The R-Q model [14] was thus modified by multiplying a constant number z as shown in (13):

$$z\frac{B(0)}{m} = a \times G \times QS^b. \tag{13}$$

After testing several sequences using various channel rates, a suitable value for the constant number z was found to be 6.5. A simulation was set up to test the sequences and to compare the QP initialization with JM. The simulation was performed using a channel rate of 128 K and the QCIF format. The sequence contained 100 frames. The results are listed in Table 1. In JM, QP initialization always occurs at 25, whereas the proposed model set value is based on the gradient value.

3.2.2 QP initialization for a transition GOP

An R-Q model was developed for the QP initialization of a transition GOP. The buffer status and the location of the scene change were considered. Similar to our previous work [22], two parameters were applied to estimate the target bits. The first parameter estimates the remaining bits, and is derived by (14):

$$T_1 = z \times \frac{B(n)}{m-n}. \tag{14}$$

Formula (14) is similar to (2) because the remaining bits are equally distributed to the un-coded frames. However, (14) contains the weighting constant z for the IPPP...P sequence. Formula (14) differs from (11) because the location of the scene-changed frame in the original GOP is considered. The second parameter T_2 is to consider the buffer status. It is derived as

$$T_2 = \frac{u(n)}{F_r} + (Tbl(n) - ABF(n)). \tag{15}$$

Formula (15) was modified from (1) by removing the constant γ, which equals to 0.5. Bits may be left in the buffer at the start of a transition GOP, and the increase in the weighting of the buffer status absorbs the extra bits quickly. Finally, the target bits are the combination of the two parameters, derived as

$$T = (1-\alpha) \times T_1 + \alpha \times T_2, \tag{16}$$

where $\alpha = \frac{n}{m}$ is to increase the weighting of the target bits from the buffer status as n increases, which causes the remaining bits to decrease. Note that the α assignment is different from the assignment of [22]. Furthermore, it becomes more difficult for the buffer to absorb the extra bits if the scene change occurs near the end of the GOP. Formula (16) emphasizes the location of the scene change and increases the weighting of the target bits from the buffer status. In the extreme case that n is equal to zero, the transition GOP is the same as an ordinary GOP, and (14) is the same as the target bits estimation in (13). Finally, the QP-clipper is removed from the scene-changed frame so that the QP value can be assigned adaptably.

To calculate the target bits estimation, (13) and (16) are combined to form (17):

$$T = a \times G \times QS^b, \tag{17}$$

Formula (17) was developed to assign the QP for the scene-changed frame of a transition GOP. After simulating several test sequences, the suitable values for the parameters were $a = 14,500$ and $b = -0.8$.

4. Simulation and Experience Results

Simulations were designed to detect the scene-changed frames, or more precisely, the frames producing heavy bits, and then to show the rate control performance. The simulation conditions were illustrated in table 2. Scene-changed sequences were created using either two concatenated sequences or four concatenated sequences. For the two concatenated sequences, each sequence contained 50 frames, and for the four concatenated sequences, each sequence contained 25 frames.

Several concatenated sequences were tested. The average and maximum SADRs from all frames, except the scene-changed frames, of each sequence are listed in Table 3. The corresponding frame numbers are in the parenthesis. It shows that the maximum SADR values are approximately 1.1 to 1.4. A threshold of 2 was thus set for the SADR to determine the scene change. As a consequence, if the SAD value is twice or more than twice of its previous frame, the proposed adaptive mechanism starts to operate because this frame may create heavy bits to influence the rate control performance. The QCIF format and concatenated sequences were used in simulations with this detection criterion. Finally, Table 3 also lists the SADR values

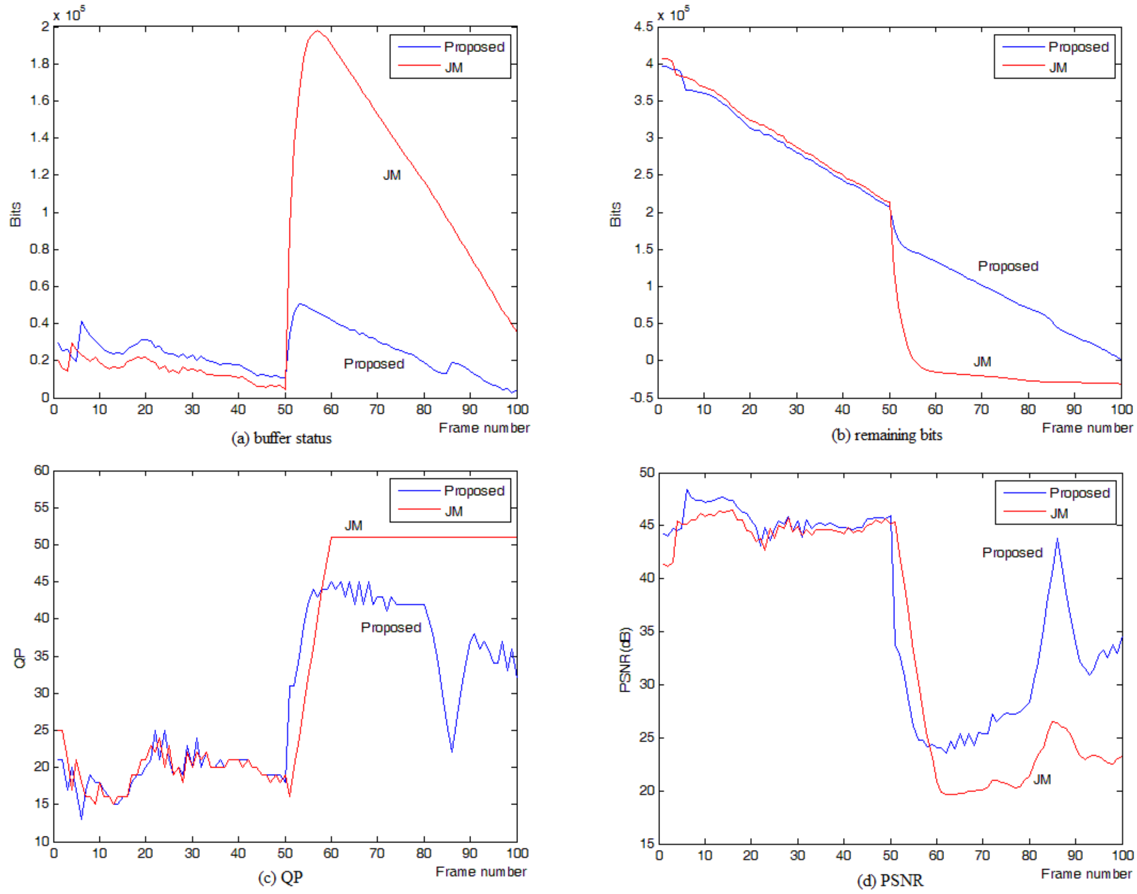

Fig. 2 Performance comparisons between JM and our method for a concatenated sequence Akiyo_Football at 128Kbps: (a) buffer status, (b) remaining bits, (c) QP, and (d) PSNR.

Table 2 Simulation conditions

JM Version	17.2
Profile	Baseline
Resolution	QCIF
Frame Rate	30 fps
Encoded Frames	100
GOP Size	1
Sequence Type	IPPP...P
Search Range	22
Reference Frames	1
RDO	On (High complexity mode)
Rate Control	On
Target Bitrate	64Kbps, 96Kbps, 128Kbps
Search Mode	Full Search

Table 3 Scene change detection for different concatenated sequences

Sequence	**Scene-changed frame(s) (SADR >2.0)**	**Other Frames**	
		Average SADR	**Maximum SADR**
Silent_Garden _Suzie_Container	16.7 (26) 2.6 (51) 12.6 (76)	0.97	1.3 (53)
Akiyo_Football	26.8 (51)	0.99	1.3 (87)
Football_Akiyo	2.3 (51)	0.99	1.1 (4)
Carphone_Foreman	15.9 (51)	0.99	1.4 (43)
Foreman_Carphone	9.0 (51)	0.99	1.3 (55)
Claire_Coastguard	31.8 (51)	1.00	1.3 (48)
Coastguard_Claire	6.81(51)	1.00	1.3 (75)

for the scene-changed frames. Most of the SADR values are extremely high.

The concatenated sequences, Akiyo and Football, illustrated a comparison and discussion of the rate control performance. The results for other sequences are listed in Table 4. Akiyo is a low-motion sequence, whereas Football is a high-motion sequence. Figs. 2(a), (b), (c), and (d) compare the rate control performance between JM and the proposed method for the buffer status, remaining bits, QP, and the PSNR, respectively.

After encoding the 51st frame, which was the scene-changed frame, heavy bits in the buffer increased significantly when using JM. As a result, the QP values of the 52nd frame and the next frames should be much larger than

Table 4 Performance comparisons with JM and [11]

Sequence	Rate (Kbps)	Y-PSNR (dB)			ΔY-PSNR gain (dB)		Bit rate (Kbps)			ARD(%)		
		JM	[11]	Proposed	JM vs Proposed	[11] vs Proposed	JM	[11]	Proposed	JM	[11]	Proposed
Akiyo_Football	128	34.48	36.47	37.69	3.21	1.22	138.68	130.28	129.37	8.34	1.78	1.07
	96	33.76	35.22	36.27	2.51	1.05	111.00	98.93	97.29	15.62	3.05	1.35
	64	31.82	32.27	34.36	2.55	2.10	74.39	90.42	64.80	16.24	41.28	1.25
Football_Akiyo	128	32.89	33.72	34.34	1.44	0.62	128.06	127.46	128.37	0.05	0.42	0.29
	96	30.53	32.07	32.97	2.44	0.91	95.77	96.63	96.30	0.24	0.66	0.31
	64	29.95	31.31	31.17	1.22	-0.15	64.75	64.39	65.00	1.17	0.61	1.57
Foreman_Carphone	128	36.49	36.53	36.55	0.06	0.02	128.30	128.19	128.40	0.23	0.15	0.31
	96	35.11	35.10	35.23	0.12	0.13	95.98	95.79	96.18	0.03	0.21	0.19
	64	33.33	33.25	33.47	0.13	0.21	64.28	64.80	64.56	0.44	1.25	0.87
Coastguard_Claire	128	38.42	38.71	38.99	0.57	0.28	128.17	128.45	128.93	0.13	0.35	0.72
	96	37.10	37.48	37.71	0.62	0.23	96.27	96.25	96.62	0.29	0.26	0.64
	64	35.82	35.76	36.01	0.18	0.25	64.26	64.04	64.88	0.41	0.07	1.38
Claire_Coastguard	128	37.90	38.10	38.46	0.55	0.35	128.03	128.68	129.25	0.02	0.53	0.98
	96	36.83	36.93	37.27	0.44	0.34	97.53	97.10	96.45	1.60	1.15	0.46
	64	34.51	35.06	35.46	0.95	0.40	66.44	64.50	64.44	3.82	0.78	0.68
Average		34.64	35.20	35.73	1.10	0.53				3.24	3.50	0.80

the value of the 51st frame. However, the QP clipper confined the increment to only 2 causing these heavy bits continuously to accumulate in the buffer. Furthermore, encoding Football created more bits compared with Akiyo, and larger QP values should be assigned to the Football frames. But, too small QP value was assigned to the 51st frame, the Football frame, and caused heavy bits creation in the buffer. The QP values for the 52nd frame and the next frames were still too small, restricted by the QP clipper, which resulted in these frames rapidly consuming the remaining bits. The heavy bits in the buffer and the shortage of remaining bits lasted to the end of the GOP, as shown in Figs. 2(a) and (b), respectively. For the rest of the frames, the PSNR performance deteriorated even though the QP values were adjusted to the maximum value of 51 because of the shortage of remaining bits. The results are shown in Figs. 2(c) and (d).

Using the proposed system, the 51st frame was detected as a scene change and a suitably high QP value was assigned without the QP restriction. With a high QP value, less extra bits were contained in the buffer, compared with JM. The following few frames absorbed the extra bits quickly because the QP value was large enough. At the end of the GOP, no extra bits were left in the buffer, as shown in Fig. 2(a). In other words, Fig. 2(a) shows that the proposed method requires a smaller buffer size than JM. Fig. 2(b) shows that, after the scene change, the remaining bits were still sufficient. In addition, a large QP value was assigned to the scene-changed frame, so suitable QPs could be assigned to the next frames even under the confinement of the QP clipper. The comparisons of QP and PSNR performance are shown in Figs. 2(c) and (d), respectively.

Finally, the performance comparisons with JM and [11] are listed under alternative sequences and channel rates. The PSNR and absolute rate difference (ARD) performances are shown in Table 4, where

$$ARD = \frac{|encoded\ bit\ rate - target\ bit\ rate|}{target\ bit\ rate}. \quad (18)$$

Compared with JM and [11], the proposed method outperforms JM and [11] with a PSNR of 1.1 dB and 0.5 dB, and with a reduction of an ARD of 2.4 % and to 2.7 %, on average, respectively. In other words, the proposed method has superior PSNR performance and maintains improved bitrate accuracy.

5. CONCLUSION

In this work, the weakness of JM was analyzed in terms of the influence of a scene change. A simple method was designed to efficiently detect the frames producing heavy bits. Based on the weakness of JM, its algorithm was modified. The modification included the QP assignment to the first I frame of an ordinary GOP and the first I frame of a transition GOP. Simulation results show that the proposed method improves rate control performance, including buffer level management, remaining bits, and PSNR performance. The proposed method can improve the PSNR to approximately 1.1 dB on average as well as the bitrate accuracy with a small extra buffer.

References

[1] T. Wiegand, G. J. Sullivan, G. Bjonte, and A. Luthra, ”Overview of

the H.264/AVC video coding standard," EEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560-576, July 2003.

[2] Z. Li, F. Pan, K. P. Lim, G. Feng, X. Lin, and S. Rahardja, "Adaptive basic unit layer rate control for JVT," JVT-G012-r1,7th Meeting, Pattaya II, Thailand, Mar. 2003.

[3] S. Ma, Z. Li, and F. Wu, "Proposed Draft of Adaptive Rate Control for JVT,"

[4] T. Chiang, and Y. Q. Zhang, "A new rate control scheme using quadratic rate-distortion modeling," IEEE Trans. Circuits and Systems for Video Technology, vol.7, no. 1, pp. 246- 50, Feb.1997.

[5] J.-Y. Lee, and H. W. Park, "A rate control algorithm for DCT-based video coding using simple rate estimation and linear source model," IEEE Trans. Circuits and Systems for Video Technology, vol.15, no. 9 pp. 1077- 86, Sep. 2005.

[6] W. A. C. Fernando, C. N. Canagarajah, and D. R. Bull, "Fade-in and fade-out detection in video sequence using histograms," IEEE International Symposium on Circuit and Systems, Geneva, vol. 4, pp. 709 - 12, 2000.

[7] L.-J. Luo, C.-R. Zhou, and Z.-Y. He, "A new algorithm on MPEG-2 target bit-number allocation at scene changes," IEEE Trans. Circuit. Syst. Video Technol. vol.7, no. 5, pp.136- 47, Oct.1997.

[8] H.-C. Hwang, and D.-G. Kim, "Shot detection from MPEG compressed video," IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences, vol. E87-A, no. 6, pp. 1509-13, 2004.

[9] C. Lee, S. Lee, Y. Oh, and J. Kim, "Real-time H.264 rate control for scene-changed video at low bit rate," International Conference on Consumer Electronics (ICCE2007), pp. 10- 14, Jan. 2007.

[10] J.-R. Ding, and J.-F. Yang, "Adaptive group-of-pictures and scene change detection methods based on existing H.264 advanced video coding information", IET Image Process., vol. 2, no. 2, pp. 85- 94, 2008.

[11] C. Lee, Y. Jung, S. Lee, Y. Oh, and J. Kim, "Real-time frame-layer H.264 rate control for scene-transition video at low bit rate," IEEE Trans. Consumer Electronics, vol. 53, no. 3, pp. 1084- 91, Aug. 2007.

[12] X. Chen, and F. Lu, "Adaptive rate control algorithm for H.264/AVC considering scene change," Mathematical Problems in Engineering, vol. 2013, article ID 373689, 2013.

[13] X. Jing, 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- 76, 2008.

[14] W.-J. Tsai, and T.-L. Chou, "Scene change aware Intra-frame rate control for H.264," IEEE Trans. Circuits and Systems for Video Technology, vol. 20, no. 12, pp. 1882 - 86, Dec., 2010

[15] H.-J. Lee, Tihao Chiang, and Ya-Qin Zhang, "Scalable Rate Control for MPEG-4 Video," IEEE Trans. Circuit Syst. Video Technology, vol.10, no. 6, pp. 878- 94, Sep. 2000.

[16] Z. Li, C. Zhu, N. Ling, X. Yang, G. Feng, S. Wu, and F. Pan, "A unified architecture for real-time video-coding systems," IEEE Trans. Circuit Syst. Video Technol., vol. 13, pp. 472- 87, Jun. 2003.

[17] Z. G. Li, L. Xiao, C. Zhu and P. Feng, "A novel rate control scheme for video over the Internet," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Florida, USA, pp. 2065- 68, May 2002.

[18] Z. G. Li , W. Gao , F. Pan , S. W. Ma, K. P. Lim , G. N. Feng, X. Lin, S. Rahardja, H. Q. Lu and Y. Lu, "Adaptive rate control for H.264," J. Vis. Commun. Image Represent., vol. 17, pp.376- 406, 2006.

[19] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," IEEE Trans. Circuit Syst. Video Technology, vol. 9, no. 1, pp. 172- 85, Feb. 1999.

[20] A. Vetro, H. Sun, Y., and Wang, "MPEG-4 rate control for multiple video objects," IEEE Trans. Circuit Syst. Video Technol. vol. 9, no. 1, pp. 186-99 Feb. 1999.

[21] U. Gargi, R. Kasturi, and S. H. Strayer,"Performance characterization of video shot change detection methods," IEEE Trans. Circuits Syst. Video Technol., vol. 10, no. 1, pp. 1- 13, Feb. 2000.

[22] J.-T. Fang, C.-C. Chan, S.-R. Hsu and P.-C. Chang, "Robust rate control mechanism against scene change for H.264/AVC, " The 17th IEEE International Symposium on Consumer Electronics (ISCE 2013), Hsinchu, Taiwan, R.O.C., pp. 241-242, Jun. 2013.

Jiunn-Tsair Fang received the B.S. degree in Physics from National Taiwan University in 1987, and Ph.D degree in electrical engineering from National Chung-Cheng University, Taiwan in 2004. Currently, he is an assistant professor in the Department of Electronic Engineering at Ming Chuan University, Taiwan. His researching interest includes video/image coding, channel coding, and joint source and channel coding.

Zong-Yi Chen received the B.S. degree in electrical engineering and the M.S. degree in communication engineering from the National Central University (NCU), Taiwan, in 2005 and 2007, respectively. He is currently pursuing the Ph. D. degree at the Video-Audio Processing Laboratory (VAPLab) in communication engineering, NCU, Taiwan. His research interests include video/image processing and video compression.

Chen-Cheng Chan received the B.S. degree in business administration and the M.S. degree in communication engineering from the National Sun Yat-sen University and NCU, Taiwan, in 1997 and 2012, respectively. His research interest is video coding.

Pao-Chi Chang received the B.S. and M.S. degrees from National Chiao Tung University, Taiwan, in 1977 and 1979, respectively, and the Ph. D. degree from Stanford University, California, 1986, all in electrical engineering. From 1986 to 1993, he was a research staff member of the department of communications at IBM T. J. Watson Research Center, Hawthorne, New York. At Watson, his work centered on high speed switching systems, efficient network design algorithms, and multimedia conferencing. In 1993, he joined the faculty of NCU, Taiwan, where he is presently a Professor in the Department of Communication Engineering and the Department of Electrical Engineering, as well as the Advisor of the VAPLab. He was a visiting professor at Stanford University in 2000 and 2004, respectively. His main research interests include speech/audio coding, video/image compression, scalable coding, error resilient coding, digital watermarking and data hiding.