Improvement on Rate-Distortion Performance of H.264 Rate Control in Low Bit Rate

ShuMin Zhou, JinTao Li, JinHao Fei, and YongDong Zhang

***Abstract*—This paper points out some defects in the techniques used in H.264 rate control and presents several new algorithms to improve them. The improved algorithm has the following main features: 1) the bits allocated to each P-frame is proportional to the local motion in it, i.e, more bits are allocated to a frame if the local motion in it is stronger; 2) the quantization parameter for I-frame is choosed based on a new bits allocation scheme for I-frame; 3) the quantization parameter calculation is based on a simple encoding complexity prediction scheme, which is more robust and of less complexity than the quadratic model used by H.264 in low bit rate video coding. Experimental results and analysis show that the improved rate-control scheme has significantly increased the average peak signal-to-noise ratio up to 1.53 dB, reduced the variation of buffer level, improved the perpetual quality of the reconstructed video and reduced the computation complexity.**

***Index Terms*—Bit allocation, low complexity, rate control, video coding.**

I. Introduction

VIDEO encoding rate control has been the focus of research in recent years [1]–[7]. The existing rate control schemes can be classified into three major categories according to the way of quantization step size calculation. 1) Analytical methods based on rate-quantization parameter (R-Q) model, e.g., the rate control scheme in MPEG-4 [6] and H.263 TMN8 [7]. 2) Empirical methods based on regressive search, e.g., the rate control scheme presented in [8]. 3) Simple parameter estimation methods based on buffer level or encoding complexity, e.g., the rate control scheme in MPEG-2 Test Model 5 (TM5) [1]. The rate control scheme in H.264 JM7.6 [10], which is proposed by Li [3], is the combination of the first and third category. Although the techniques used in Li's scheme are state of art, there are some defects in them, especially for low bit rate video coding.

Recently, Yuan *et al.* [7] proposed a rate-distortion (R-D) optimal solution to Li's scheme, which improved the R-D performance of rate control by Lagrange optimization. However, the defects in Li's scheme has not been solved, especially in low bit rate video coding. This paper will present several new techniques to deal with these defects so as to improve the R-D performance of rate control. In the following, firstly, we will have a brief look at the rate control scheme used in JM7.6. Then, the defects of this method will be pointed out. Based on the analysis, three new techniques are proposed.

Manuscript received November 10, 2004, revised August 26, 2006. This work is supported in part by the National Nature Science Foundation of China under Grant 60302028. This paper was recommended by Associate Editor H. Watanabe.

S. Zhou and J. Fei are with the Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518067, China (e-mail: jh.fei@siat.ac.cn; sm.zhou@siat.ac.cn).

J. Li and Y. Zhang are with the Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China (e-mail: jtli@ict.ac.cn; zhyd@ict.ac.cn.).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2007.903123

A. Rate Control Scheme in JM7.6

Fig. 1 illustrates the procedure of rate control scheme in JM7.6, from which we can see that there are three main step in this rate control scheme: a frame bits allocation scheme similar to the one in Test Model 5(TM5) of MPEG-2, a fluid flow buffer level control scheme and a quadratic formulation of the R-D function. It should be noted that in Fig. 1 we do not consider B-frame due to the fact that B-frame takes up only a little part in total encoded bit stream, which can be ignored in low bit rate. As a matter of fact, even in the rate control scheme of Li's, only a simple linear model is used to model B-frame. In addition, low delay is one of the main features emphasized in low bit rate video coding while B-frame would lead to extra delay. Therefore, we only consider I- and P-frame in the paper.

The target bits estimation (also known as frame layer bits allocation) mainly focused on how to effectively allocate bits for each frame, which is a dilemma since we cannot obtain the accurate bits count until the current unit has been encoded. In JM7.6, a popular method, which is known as MPEG-2 TM5 [1], is adopted to solve the problem by assuming that neighbor frames of same type have the similar encoding complexity [2], i.e,

$$X_k \approx X_{k-1}, X_i = B_i * \mathrm{QP}_i \tag{1}$$

here X_i is the encoding complexity for frame i, B_i stands for bits used in frame i, QP_i is quantization parameter.

The buffer level control aims at preventing encoding buffer from overflow or underflow. In JM7.6, a fluid flow control scheme [3] and hypothesis reference decoder (HRD) [4] are used to deal with the problem. In this scheme, target buffer level is modified by using a simple linear model as follows when there are no B-frames being considered [3]:

$$\mathrm{Tbl}(i,j+1) = \mathrm{Tbl}(i,j) - \frac{Tbl(i,2) - \frac{B_s}{8}}{N_p - 1} \tag{2}$$

where $\mathrm{Tbl}(i,j)$ stands for target buffer level when encoding the jth frame in the ith group of pictures (GOP), B_s is the size of encoding buffer, N_p is the number of coded P-frames in current GOP and $(i,2)$ indicates the second frame in GOP i.

1051-8215/$25.00 © 2007 IEEE

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.

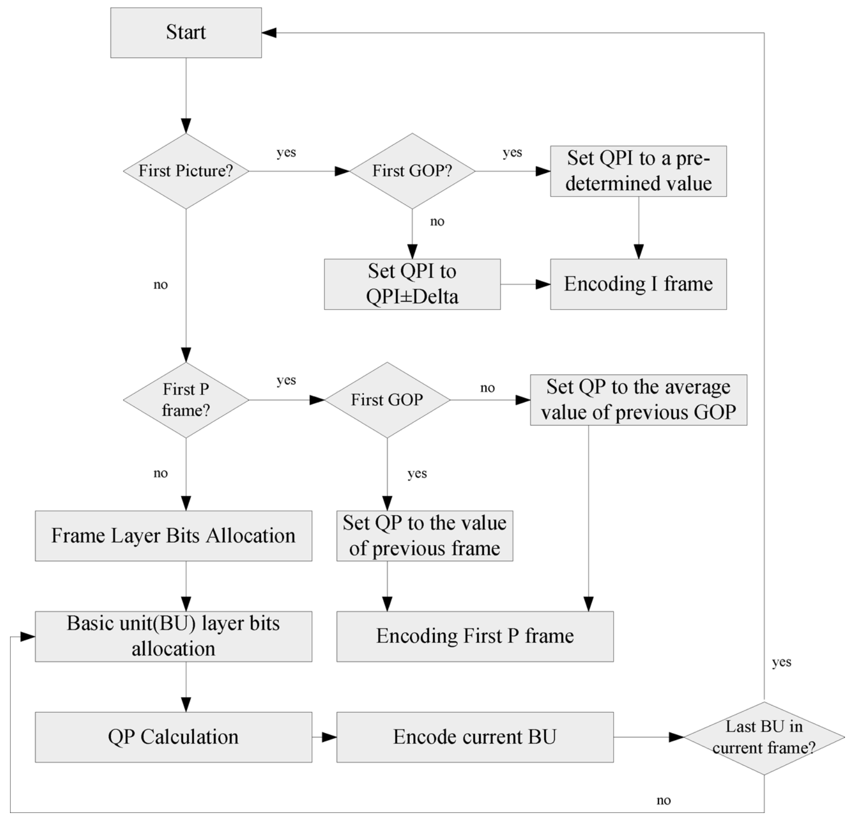

Fig. 1. Rate control scheme proposed by Li.

The QP calculation, which is the key part of rate control, is up to calculate the quantization step according to the target bits and buffer level requirement. As to Li's scheme, a quadratic rate-quantization step model, which is known as MPEG-4 Q2 [6], is used to calculate the quantization step in basic unit layer (frame layer control can be seen as a special case of basic unit) [3].

B. Defects Analysis in JM7.6 Rate Control

1) Defect in Bits Allocation: As is mentioned above, bits allocation in Li's scheme is based on the assumption that there is similar encoding complexity between two neighbor frames. However, the encoding complexity of neighbor frames in a video sequence could be quite different when there is large local motion change. Therefore, such a bits allocation scheme would have poor R-D performance when the local motion changes abruptly between the neighbor frames.

2) Defect in QP Calculation: Although progressive parameter prediction and Lagrange optimization are used in the process of QP calculation of Li's rate control scheme, the analytical model (MPEG-4 Q2) cannot obtain accurate enough bits estimation in low bit rate video coding since the local motion information, which is distributed randomly over the whole GOP and cannot be precisely described by using a single analytical model, takes up quite a large scale in total bits of inter frames. Thus, such an analytical model would deteriorate R-D performance greatly when local motion is not evenly distributed in a GOP.

3) Defect in Initial QP Estimation: In bit-rate constrained encoding, the initial quantization step selection also has obvious impact on R-D performance of rate control, which is illustrated in Fig. 2. It can be seen clearly from Fig. 2 that the optimal initial quantization step size is not only concerned with target bit-rate and GOP length, but also the sequence itself. However, the initial quantization step size is adjusted only according to target bit-rate and length of GOP in Li's scheme [3], which is surely not comprehensive.

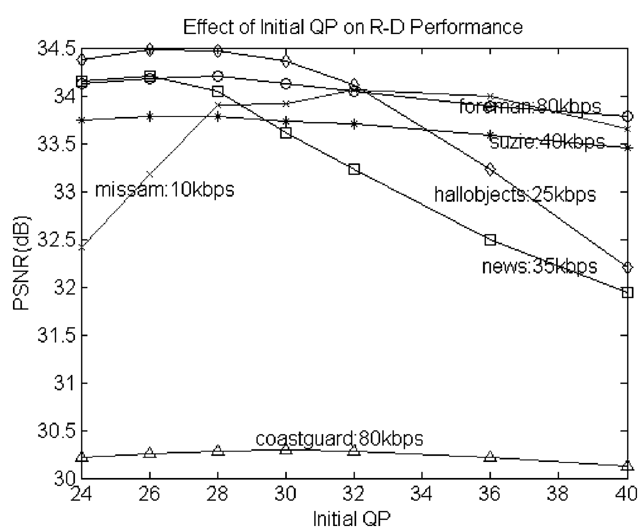

Fig. 2. Effect of initial QP on R-D performance, 100 frames (two GOP), all sequences here are QCIF, 30 Hz; the rate control scheme used here is Li's method.

In consideration of all the defects mentioned above, several new techniques are proposed to improve the R-D performance

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.

of rate control for H.264 in this paper. Firstly, we propose a new initial quantization step size selection scheme based on a balanced bits allocation strategy between I- and subsequent P-frames. Then, a new frame layer bit allocation scheme for P-frame is presented based on histogram of difference frame (HOD), in which target bit allocation is sensitive to local motion change rather than dull. Finally, as to the inefficiency of analytical R-Q model in low bit rate coding, a simple encoding complexity prediction scheme is proposed to calculate QP, in which encoding complexity of current coding unit is predicted according to the distribution of its neighbor HOD and encoding complexity.

Partial content of the paper has been presented in literature [11] and [12]. This paper will give a comprehensive of the three new techniques.

The rest of the paper is organized as follows. In Section II, as to the defects in AVC rate control mentioned above, three new techniques are presented in detail. Section III gives the improved rate control scheme with the proposed new techniques and analyze its complexity. The experimental results and discussions are given in Section IV. Finally, our summary and conclusion are presented in Section V.

II. Several New Techniques

In the section, first, a sequence adaptive initial quantization step size for I-frame is proposed. Then the HOD-based frame layer bits allocation scheme is given. Finally, in order to deal with the inaccuracy of analytical R-Q model in low bit rate, we propose the new quantization step size calculation scheme based on encoding complexity prediction, which is performed on basic unit layer.

A. R-D Optimal Initial QP Selection

As is shown in Fig. 2, the optimal initial QP for a GOP is not only concerned with target bit rate and GOP length but also sequence itself (e.g., motion strength and variance of I-frame in a sequence). As to the problem, Pan has proposed a method in which the initial QP is concerned with target bit rate, GOP length and the mean absolute value (MAV) of the discrete cosine transform (DCT) coefficient of the first frame [13]. Although they take feature of I-frame into consideration in the paper, the dependency between I- and P-frame has not been considered. In this subsection, we develop a new scheme in which the initial QP is obtained based on a balanced bits allocation scheme between I- and P-frame.

In the following, first we will present the balanced bit allocation scheme for I-frame based on some reasonable assumptions; then by using an accurate enough R-Q model, the QP can be derived. Both of these are developed in spatial rather than DCT domain, which is more favorable for H.264 than Pan's model.

1) Balanced Bits Allocation Between I- and P- Frame: Since target bit rate always can be predetermined, we present the balanced bits allocation for I- and P-frame under the constraint of certain target bit rate at first. Then, the adaptive bits allocation for I-frame among different target bit rate can be made out based on the constrained result by experimental analysis.

Prior to encoding a GOP under certain target bit rate, we assume that all P-frames use the same bits so as to get an analytical model for balanced bits allocation between I- and P-frames. In addition, to simplify the deduction, we also assume that I- and P-frame has the same analytical R-Q format in this section. Although such an assumption is not accurate, it is feasible theoretically at least.

To obtain optimal bits balance between I- and P-frame over a GOP, we introduce a new variable L which indicates the ratio of bits assignment between I- and P-frame

$$L = \frac{R_0}{R_p} \tag{3}$$

where R_0 stands for the bits allocation for I-frame, R_p indicates the average bits allocation for P-frames in a GOP.

Based on Lagrange relaxation R-D scheme, the R-D cost function of a GOP (format is IPPP) can be expressed as under the constraint of certain target bit rate

$$J = \sum_{i=0}^{M-1} D_i + \lambda \left(\sum_{i=0}^{M-1} R_i - B \right) \tag{4}$$

where D_i, R_i stand for distortion and bits for frame i, M is the length of GOP, and B is target bits.

D-Q model can be provided as follows [5]:

$$D_i(Q_i) = \frac{Q_i^2}{12}. \tag{5}$$

Here Q_i is quantizer step size of frame i. Note that frame 0 is I-frame in a GOP.

The R-Q model can be approximately formulated as follows in low bit-rate [5]:

$$R_i = \frac{e}{\ln 2} \frac{\delta_i^2}{Q_i^2}, 0 \leq i \leq M-1. \tag{6}$$

Here $\delta_i^2 = 1/N_{\rm pix} \sum_{k=1}^{N_{\rm pix}} (y_k - \bar{y})^2$ is the variance of source data, where y_k is the luminance value of kth pixel in frame i.

It should be noted that when a GOP is considered as a whole, the model in (6) is more feasible since we mainly focus on low bit rate video coding in the paper. Thus, we can deduce the following equation based on the above assumptions and (3)–(6) (See Appendix A):

$$L = \frac{\delta_0}{\delta_\mu}. \tag{7}$$

Here $\delta_\mu = \sum_{k=1}^{M-1} \delta_k / M - 1$ is the average standard deviation of P-frames in a GOP, δ_0 is the standard deviation of the first frame (I-frame).

The relationship in (7) is an ideal model of L under certain target bit rate, but it reflects the linear correlation between optimal L and the ratio of standard deviation (RSD) $= \delta_0/\delta_\mu$ at least. To confirm this, exhaustive experiments have been done, in which foreman, akiyo, mother&daughter, coastguard, container, hallobjects, miss_am, suzie, and news are tested under 25, 50,70, 100, 150, and 300 kbps, respectively. All of the sequences used are QCIF(30 Hz). In the simulation, rate control in JM7.6 is used to control the encoding bit rate; motion estimation (ME) accuracy is 1/4 pel (full block mode); the reference frame number is 2 and frame mode is IPP.

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.

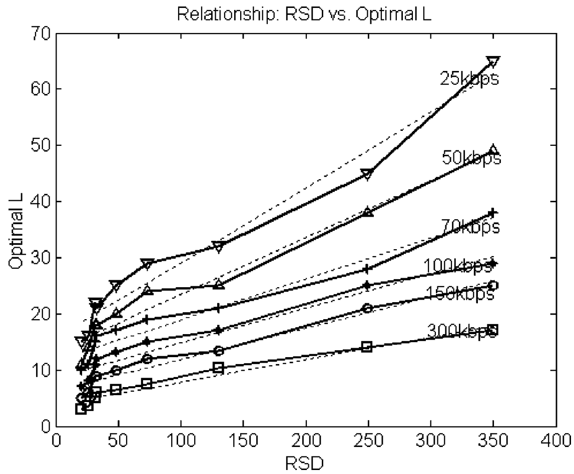

Fig. 3. Linear model of optimal L (first GOP for each sequence). Note that here the coding conditions are the same as in Fig. 2 except only one GOP is encoded (50 frames).

The optimal L is obtained by exhaustive search in the range of [1,100] (the search step is 2) for each sequence under certain target bit rate. The search scheme is detailed as follows.

Step 1) Select a value of L under certain target bit rate, e.g., $L = 50, 60$ kbps; then the bits allocation for I-frame can be made out.

Step 2) According to the bits allocation for I-frame, QP of I-frame can be calculated by use of R-Q model for I-frame, which will be shown in later section.

Step 3) Encode the sequence and take down average PSNR performance; then change L to another value, i.e., $L = 52$, go to step 1.

Step 4) Iterate the process in Steps 1–3 until all values of L are used out; choose the one which leads to highest PSNR performance as the optimal value of L in current target bit rate.

As to the average standard deviation of P-frames, since it is mainly determined by local motion strength in a sequence when target bit rate is fixed, we replace it with a statistic feature which can indicate the local motion efficiently and be proportional to δ_μ. The statistic feature will be detailed later.

Fig. 3 shows the relationship between optimal L and RSD under certain target bit rate, from which we can see the following.

1) The optimal L is approximately linear with RSD when target bit rate is fixed, which verifies the validness of (7), whose experimental model is given in (8)

$$L = A * \mathrm{RSD} + B, \begin{vmatrix} A \\ B \end{vmatrix} = \begin{vmatrix} a1 & a2 \\ b1 & b2 \end{vmatrix} \begin{vmatrix} \mathrm{TBR} \\ 1 \end{vmatrix} \tag{8}$$

where TBR is the abbreviation of target bit rate, $a1$, $a2$, $b1$, $b2$ are statistic constant.

2) Among different target bit rates, the gradient and offset of each line, i.e., parameter A and B in (8), declines as the target bit rate increases.

To make out the relationship between the parameters in (8) and target bit rate, further experiments have been done, whose results are shown in Fig. 4. It can be seen clearly in Fig. 4 that

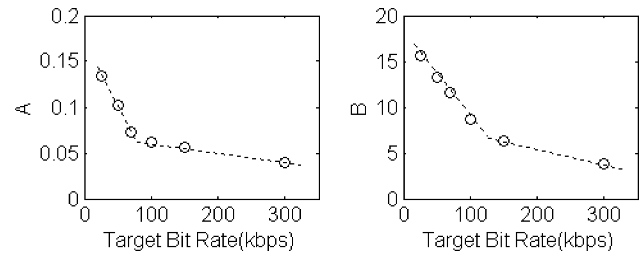

Fig. 4. Relationship between A, B, and target bit rate ($L = A * \mathrm{RSD} + B$).

both A and B are approximately linear with target bit rate except that there is an abrupt change in gradient and offset when the target bit rate is up to certain value. Thus, A and B can be expressed as the function of target bit rate, which is shown in (8) also. The parameters $a1$, $a2$, $b1$, $b2$ can be obtained as follows by linear regression in the paper:

$$(a1, a2) = \begin{cases} (-0.0014, 0.1688), \mathrm{TBR} < T_{\mathrm{tbr},a} \\ (-0.0001, 0.0724), \mathrm{TBR} \geq T_{\mathrm{tbr},a} \end{cases}$$

$$(b1, b2) = \begin{cases} (-0.0922, 17.9151), \mathrm{TBR} \leq T_{\mathrm{tbr},b} \\ (-0.0165, 8.7518), \mathrm{TBR} > T_{\mathrm{tbr},b} \end{cases}$$

where $\mathrm{T}_{\mathrm{tbr},a}$ and $\mathrm{T}_{\mathrm{tbr},b}$ are the target bit rate threshold of abrupt change in gradient and offset, respectively. In the paper, both of them are set to be 100 kbps through experiments.

With the L in (8), balanced bit allocation for I-frame can be calculated as follows:

$$R_0 = M \frac{\mathrm{bit\_rate}}{\mathrm{frame\_rate}} \frac{L}{(L + M - 1)} \tag{9}$$

where *bit_rate* is target bit rate, *frame_rate* is frame frequency.

2) Derived R-Q Model for I-Frame: According to R-Q model for I-frame in literature [9], relationship between R and Q can be expressed as follows in high bit rate:

$$R_0(Q_0) = \log_2 \left(2e^2 \frac{\delta_0^2}{Q_0^2} \right).$$

Note that here only I-frame is considered, which is the typical case of high bit rate coding. Therefore, the R-Q model for high bit rate is used in the situation.

To confirm the accuracy of the relationship above, foreman, akiyo and news are used to generate statistical data. All of them are encoded in I-frame only. The total test frame number is 100 for each sequence and QPs are set to be [24,26,28,30,32,35,38,40]. Fig. 5 shows the results that R is quadratic function of $\ln(Q)$ rather than linear. Thus, a more accurate R-Q function for I-frame is presented as follows:

$$R_0(Q_0) = a * \ln^2 \left(\frac{\delta_0}{Q_0} \right) + b * \ln \left(\frac{\delta_0}{Q_0} \right) + c \tag{10}$$

where a, b, and c are statistic parameters.

The value of a, b and c can be obtained by linear regression: $a = 0.2346$, $b = 0.5657$, $c = 0.6206$.

The relationship between QP and Q in H.264 can be approximately expressed as $Q = 2^{\mathrm{QP}/6}$. Therefore, QP_0 can be easily obtained once $\ln(\delta_0/Q_0)$ is made out from (10) at certain target bits R_0.

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.

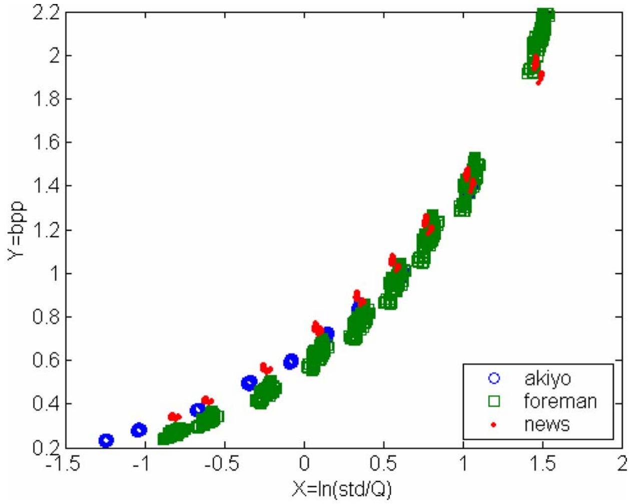

Fig. 5. Relationship between bits per pixel (bpp) and $\ln(\mathrm{std}/\mathrm{Q})$ for I-frame.

B. HOD-Based Frame Layer Bits Allocation

The target bit estimation scheme in H.264 JM7.6 rate control is proposed by Ma and Li [14], [3], in which target bits for current frame is allocated according to the encoding complexity of last frame and current buffer fullness. Once there is an abrupt motion change, however, such a scheme cannot adapt to the change effectively, which would deteriorate R-D performance greatly. To deal with the problem, an alternative method based on HOD image is proposed in this section. The HOD between frame m and n provides as follows [15]:

$$\mathrm{HOD}(f_n, f_m) = \frac{\sum_{i \notin [-\alpha,\alpha]} \mathrm{hod}(i)}{N_{\mathrm{pix}}} \quad (11)$$

here $\mathrm{hod}(i)$ is the histogram of level i in the difference frame between f_n and f_m, N_{pix} is image size in pixel. α stands for the difference level threshold in hod computation.

As a matter of fact, Song has proposed such a scheme, in which the target bits is allocated over the whole GOP [16]

$$T_k = \left(1 + \frac{\mathrm{HOD}_k - \mu_{\mathrm{hod}}}{\mu_{\mathrm{hod}}}\right) \cdot \frac{B_{\mathrm{gop}}}{M}. \quad (12)$$

Here T_k is the bits assignment for frame k, μ_{hod} is the average HOD in current GOP, B_{gop} is the total bits for a GOP, $\mathrm{HOD}_k = \mathrm{HOD}(k-1, k)$ is the HOD value of frame k.

Obviously, N_{gop} frames should be buffered before calculate target bits allocation for frame k in (12), which induces long delay in encoding process. In low delay communication, such a scheme is not feasible.

To eliminate the delay induced by (12), an alternative method is presented, in which the overall average HOD is replaced by progressive average

$$T_k = \left(1 + \frac{\mathrm{HOD}_k - \frac{1}{k}\sum_{i=1}^{k} \mathrm{HOD}_i}{\frac{1}{k}\sum_{i=1}^{k} \mathrm{HOD}_i}\right) \frac{B_r}{N_r} \quad (13)$$

where B_r and N_r are the remained bits and frames in current GOP before encoding frame k, respectively.

Although (13) is suboptimal compared to (12), it eliminates the encoding delay completely. In addition, as far as the adaptation to local motion change, both are effective.

When bits is used out in encoding process, the bits assignment will be negative which would lead to the quality deterioration of subsequent frames. In addition, if the bits assignment for current frame is too high, the remaining bits for the subsequent frames will decrease greatly, which would deteriorate the quality of subsequent frames also. Therefore, we truncate the value of bits assignment in (13) for each frame:

$$T_k = \max(\mathrm{LEAST\_BITS}, \min(T_k, \mathrm{MAX\_BITS})). \quad (14)$$

The LEAST_BITS and MAX_BITS are set to be 96 and $2 * \mathrm{bit\_}rate/\mathrm{frame\_}rate$.

In addition, in order to avoid buffer underflow or overflow, buffer occupancy should be considered in frame layer bits assignment, which can be given as

$$T'_k = \frac{\mathrm{bit\_rate}}{\mathrm{frame\_rate}} - \gamma(\mathrm{BL}_k - \mathrm{Tbl}_k)$$

where BL_k is the current buffer occupancy and Tbl_k is obtained from (2).

Then the final frame layer bits allocation is provided as

$$T_k = \Delta T_k + (1-\Delta) T'_k. \quad (15)$$

Here $\Delta = 0.5$ is weighted coefficient.

C. Low Complexity QP Estimation Scheme

The QP calculation scheme in H.264 rate control was proposed by Li, in which QP is estimated by use of quadratic R-Q model. As is mentioned above, such a method is inaccuracy for estimate the bits of P-frame in low bit rate. Therefore, a simple but effective QP estimation scheme is presented in this section in which encoding complexity is adopted to estimate QP. In the following, first we look at how to estimate the encoding complexity of current coding unit; then a HOD-based bits assignment scheme is presented, by which QP can be obtained.

1) Simple Encoding Complexity Estimation: Since motion information takes up a large part in total bits of interframe coding in low bit rate, we have the following assumptions.

- Encoding complexity of current unit is approximately proportional to local motion degree, which can be indicated by HOD
- There are great correlation between the encoding complexity for current unit and its spatial and temporal neighbors

The spatial and temporal neighbors of current coding unit is illustrated in Fig. 6. Thus, on one hand, the following relation is reasonable according to the first assumption:

$$\frac{\mathrm{HOD}_{i-1,k-1}}{\mathrm{HOD}_{i-1,k}} \approx \frac{\mathrm{HOD}_{i,k-1}}{\mathrm{HOD}_{i,k}} \Leftrightarrow \frac{X_{i-1,k-1}}{X_{i-1,k}} \approx \frac{X_{i,k-1}}{X_{i,k}} \quad (16)$$

where $X_{i,k}, \mathrm{HOD}_{i,k}$ stand for the encoding complexity and HOD of kth basic unit in the ith frame, respectively.

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.

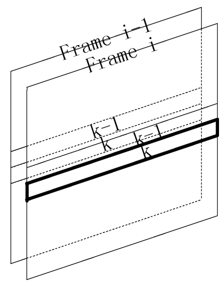

Fig. 6. Spatial and temporal neighbors of current basic unit k. Note that here basic unit is set to a line of MBs.

Therefore, the current encoding complexity can be provided

$$\left|\frac{\mathrm{HOD}_{i-1,k-1}}{\mathrm{HOD}_{i-1,k}}-\frac{\mathrm{HOD}_{i,k-1}}{\mathrm{HOD}_{i,k}}\right|<\alpha\Rightarrow X_{i,k}=X_{i,k-1}\frac{X_{i-1,k}}{X_{i-1,k-1}} \tag{17}$$

where α is threshold, which is set to be 0.2 in the paper.

On the other hand, if the prerequisite of (17) is invalid, the current encoding complexity can be predicted based on its neighbors according to the second assumption, which is detailed in pseudo code.

If

$$|\mathrm{HOD}_{i,k-1}–\mathrm{HOD}_{i,k}|-|\mathrm{HOD}_{i-1,k}–\mathrm{HOD}_{i,k}|>0.1\mathrm{HOD}_{avr}.$$

Then $X_{i,k}=(2*X_{i-1,k}+X_{i,k-1})/3$

Else if

$$|\mathrm{HOD}_{i,k-1}–\mathrm{HOD}_{i,k}|-|\mathrm{HOD}_{i-1,k}–\mathrm{HOD}_{i,k}|<-0.1\mathrm{HOD}_{\mathrm{avr}}$$

Then $X_{i,k}=(X_{i-1,k}+2*X_{i,k-1})/3$

Else $X_{i,k}=(X_{i-1,k}+X_{i,k-1})/2$

End If

here $\mathrm{HOD}_{\mathrm{avr}}$ is the average HOD of current frame.

2) HOD-Based Bits Estimation: As to the bits for current coding units, we can estimate it based on HOD, which is presented as follows:

$$\mathrm{Bits}_{i,k}=\left(1+\frac{\mathrm{HOD}_{i,k}-\mathrm{HOD}_{\mathrm{avr}}}{\mathrm{HOD}_{\mathrm{avr}}}\right)*T_i. \tag{18}$$

If the bits obtained from (18) is negative, QP for current unit is set to be QP $\;$Average+2 directly. Otherwise, the QP estimation for current coding unit is given as

$$\mathrm{QP}_{i,k}=\max\left(2,\min\left(\frac{X_{i,k}}{\mathrm{Bits}_{i,k}},51\right)\right). \tag{19}$$

III. Improved Rate Control Scheme for H.264 in Low Bit Rate

With the new techniques proposed in Section II, an improved rate control scheme is presented in this section. In addition, the analysis and comparison of complexity in computation and memory are also carried out between the proposed scheme and JM version 7.6 in this section.

A. Improved Rate Control Scheme

With the new techniques proposed above, the improved rate control process is developed as follows.

1) Step 1. Initial QP Estimation: In the paper, as is mentioned above, block variance difference [15] is used to replace the average standard deviation of P-frames, which can be formulated as follows:

$$\delta_\mu=\frac{1}{N_f-1}\sum_{k=1}^{N_f}\mathrm{BV}(f_{k-1},f_k)$$
$$\mathrm{BV}(f_n,f_m)=\frac{1}{N_{\mathrm{MB}}}\sum_{i=1}^{N_{\mathrm{MB}}}|\mathrm{var}_n(i)-\mathrm{var}_m(i)|$$

where BV is the abbreviation of block variance difference, N_{MB} is number of macroblocks (MBs) in a picture. $Var_n(i)$ indicates the block variance of MB i in frame n. N_f is the frames used to replace δ_μ.

If the current GOP is the first GOP, in order to reduce the delay, N_f is set to 2, i.e., two frames should be buffered at the beginning. With the two frames, we can make out δ_0 and estimate δ_μ.

If the current GOP is not the first GOP, only the first frame in current GOP needs to be buffered so as to make out δ_0. δ_μ is estimated by using the average block variance in last GOP.

With δ_0 and δ_μ, RSD can be obtained. Then, according to the current target bit rate, the gradient A and Offset B in (8) can be made out. With A and B, the optimal L can be estimated according to (8) and accordingly, bits allocation for I-frame can be derived by (9).

Finally, the initial QP of current GOP can be made out as follows:

$$\theta=\ln\left(\frac{\delta_0}{Q_0}\right),\mathrm{QP}_0=\frac{6}{\ln 2}Q_0\Rightarrow$$
$$\theta=\begin{cases}\frac{\sqrt{b^2-4a(c-R_0)}-b}{2a}, & b^2-4ac\geq 0\\ -\frac{b}{2a}, & \text{others}\end{cases}$$

where a, b, c are the parameters in (10), R_0 is the bits allocated.

2) Step 2. Frame Layer Bits Allocation: If the current frame is the first P-frame in a GOP, we just skip this step. Otherwise, calculate HOD for each basic unit between the current frame and its preceding one. With these HODs, HOD between current frame and its proceeding one can be obtained. Then bits allocation T_k can be made out from (13) and (14).

In addition, similar to H.264 frame layer bits allocation, we need to take buffer fullness into consideration. Calculate Tbl_k with (2) and get the final bits allocation for current frame by (15).

3) Step 3. Basic Unit Layer QP Modification: For the first P-frame in a GOP, QP is set to be the initial QP obtained from Step 1. Otherwise, if current basic unit is the first basic unit of current frame, current QP is set to be the average QP of the previous P-frames. In other cases, QP can be estimated as follows.

Step 1) Allocate bits for current basic unit by (18).

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.

Step 2) If the bits allocated is negative, we set the current QP to QP $Average + 2$ directly. Otherwise, predict the encoding complexity of current basic unit by using (17) and the pseudo code.

Step 3) Calculate QP for current basic unit by (19), which is indicated by Qc

Step 4) Truncate Qc in order to keep smoothness of video quality: $Qc = MIN(Qc, Average\text{QP} + 3)$

4) Step 4. Information Update: After encoding the current basic unit, calculate its encoding complexity and store it. If current frame is the last frame of a GOP, go to step 1; else if a frame is finished, go to step 2; otherwise, go to step 3.

B. Analysis of Complexity

In this subsection, we compare the complexity both in computation and memory between the proposed rate control scheme and the one in H.264.

1) In H.264 rate control, for each basic unit, 2 linear regression are carried out (MAD prediction and Q2 model update) [3]. Assume the computation of either is C_r, which is proportional to window size. In order to make out the QP, 1 square root operation is required for each basic unit. In addition, as to each pixel in a frame, MAD operation is carried on. Thus, the computation for a frame can be formulated as:

$$C_{264\text{RC}} = 2 * C_r * N_{\text{bu}} + C_{\text{mad}} + N_{\text{bu}} * C_{\text{sqrt}} + C_{\text{other}} * N_{\text{bu}}$$

where N_{bu} is the number of basic unit in a frame; C_{sqrt} is the computation for each square root operation; C_{other} stands for all the other operations in a basic unit; C_{mad} is the total computation of MAD operations in a frame, which is proportional to pixels number in a frame.

As for the improved rate control scheme, the computation of HOD is proportional to pixels number in a frame, which is approximately the same as C_{mad}. In addition, the computation in calculating QP for each basic unit is approximately the same as C_{other}. Thus, the total computation for a frame is about to be:

$$C_{\text{proposedRC}} \approx C_{\text{mad}} + N_{\text{bu}} * C_{\text{other}}.$$

Therefore, we can obtain the approximate computation ratio between H.264 rate control and the proposed scheme:

$$\rho = \frac{C_{264}}{C_{\text{proposedRC}}} \approx 1 + \frac{2 * C_r + C_{\text{sqrt}}}{\frac{1}{N_{\text{bu}}} C_{\text{mad}} + C_{\text{other}}}. \quad (20)$$

Since only C_{mad} is parallel to image size in (20), the larger image size is, the smaller ρ is. Therefore, the proposed scheme is favorable in low bit rate application in which image resolution is relatively small.

2) For H.264 rate control, two arrays are required to store MAD for each basic unit and another two arrays are required to store QP and bits for each basic unit in estimation window. For the proposed scheme, four arrays are required to store HOD and encoding complexity for each basic unit. Therefore, when the size of window is near to N_{bu}, both of them have the similar memory complexity.

TABLE I
COMPARISON OF CONTROL PRECISION

Test Name	Video Sequence	Target Bit rate(kbps)	H.264 RC	Proposed RC
			Bit rate	Bit rate
F60	Foreman	60	60.13	60.40
F75	Foreman	75	75.03	75.35
F100	Foreman	100	100.00	100.34
F150	Foreman	150	150.01	150.30
S45	Suzie	45	44.97	45.01
S65	Suzie	65	64.96	65.09
S95	Suzie	95	94.96	95.02
S135	Suzie	135	134.97	135.11
N25	News	25	25.30	25.21
N60	News	60	60.39	60.35
N90	News	90	90.37	90.26
N130	News	130	130.27	130.59
H25	Hall	25	25.06	25.08
H45	Hall	45	45.16	45.13
H70	Hall	70	70.08	70.16
H100	Hall	100	100.19	100.17
A20	Akiyo	20	20.05	20.11
A30	Akiyo	30	30.08	30.10
A50	Akiyo	50	49.95	50.08
A80	Akiyo	80	80.00	79.99
Mi25	Miss_am	25	25.09	25.13
Mi50	Miss_am	50	49.95	50.04
Mi80	Miss_am	80	79.95	80.10
Mi100	Miss_am	100	99.95	100.08
Mo30	M&D	30	30.12	30.14
Mo50	M&D	50	50.04	50.08
Mo70	M&D	70	69.88	70.11
Mo100	M&D	100	99.96	100.00
C100	Container	100	99.95	99.92
C80	Container	80	80.01	79.92
C50	Container	50	50.02	50.10
C30	Container	30	30.08	30.09

IV. EXPERIMENTS AND DISCUSSIONS

Numerous experiments have been conducted to evaluate the performance of the improved rate control algorithm. The test sequences used are in QCIF format, with I-frame appearing every 50 frames. Note that no frame skip is concerned here and ME accuracy is 1/4 pel for full block mode (all block types are considered). Search range in ME is 16 and the reference frame number is set to be 2. The entropy coding option is CAVLC if not specified. The total frame for each sequence is 150 (3 GOPs) if not specified particularly in these simulations. The target buffer level can be obtain by (2) and buffer size is equal to one half of target bit rate, i.e, there is 500-ms encoding delay. In the simulation, JM7.6 was used if not specifed.

Table I shows that the comparison of control precision between the proposed scheme and H.264 rate control. All sequences are 30 Hz (M&D is the abbreviation of mother and daughter). It can be seen that the precision of rate control is approximately the same.

Fig. 7 shows the PSNR of the reconstructed sequence "foreman" by using Li scheme and the improved rate control scheme. In order to observe the performance of the proposed method in the case that there is scene change, the total frame in this simulation is 350 (7 GOPs) because there is a scene change in foreman sequence from frame 287 to 320. It can be shown that, for the improved scheme: 1) the PSNR fluctuation of

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.

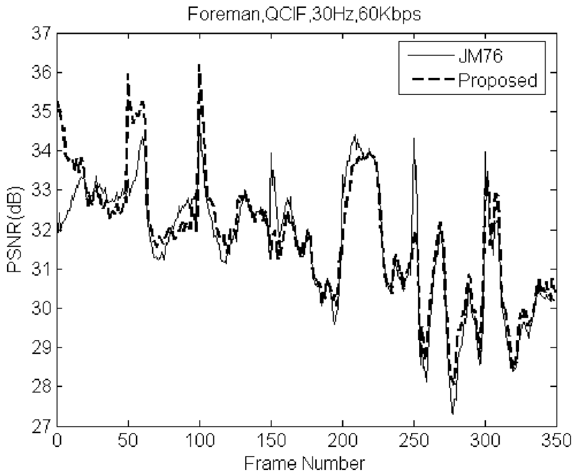

Fig. 7. PSNR results of sequence "foreman."

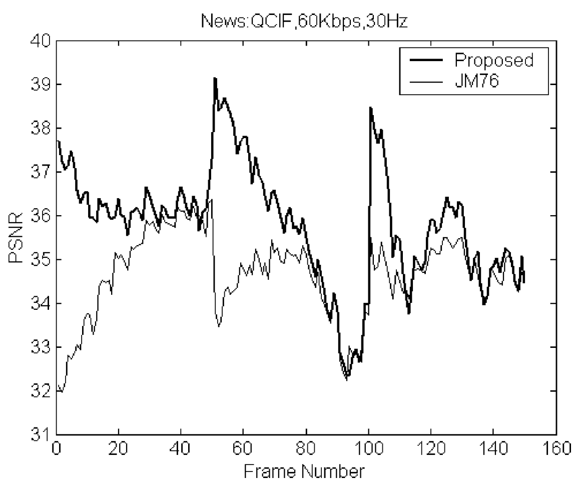

Fig. 8. PSNR results of sequence "news."

P-frames has been reduced; 2) it provides better adaptation capability to scene change than JM76; 3) the PSNR improvement in earlier GOPs (i.e., GOP 1 to 3) is higher than in later GOPs; and 4) the average PSNR has been improved about 0.14 dB.

Fig. 8 shows the PSNR of the reconstructed sequence "news." In the simulation, there is no scene change occurring. It is noted that for the proposed scheme, the PSNR of nearly all frames in a GOP are higher than Li's scheme. Since the local motion variation is strong in the sequence "news," which can be illustrated in Fig. 10, the improved scheme works much better than Li's scheme: the average PSNR gain is about 1.1 dB!

Fig. 9 shows that the buffer level variations of sequence "foreman" and "news" by using Li's scheme. It can be seen clearly that: 1) the buffer level surge occurs every time an I-frame is encoded; 2) the strength of surge varies from one to another; and 3) the actual buffer level deviates from target buffer level. Fig. 10 shows the similar diagram by using the improved scheme proposed in the paper. Compared to Fig. 9, the buffer occupancy is higher at the beginning of a GOP now, but the strength of variations reduced greatly. Also, the actual buffer level follows the target buffer level quite closely. The increase in the initial target buffer level is due to the fact that the initial a few P-frames are allocated with more bits than the average, which can better the R-D performance. In general, the proposed scheme have greatly reduced the buffer level variation and improved the accuracy of adaptation to target buffer level. Similar results could be obtained for all the other sequences that we have tested.

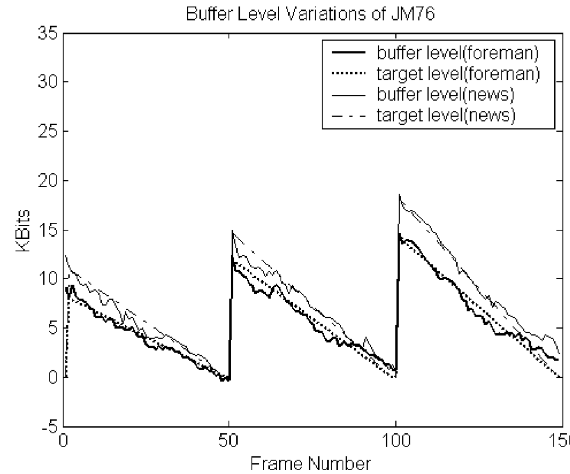

Fig. 9. Buffer level variations of JM76.

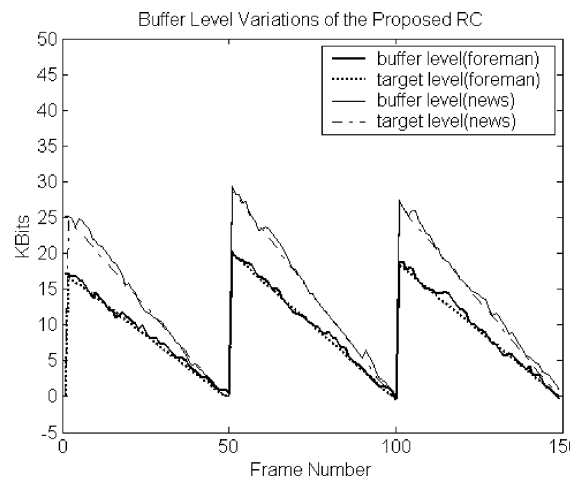

Fig. 10. Buffer level variations of the improved scheme.

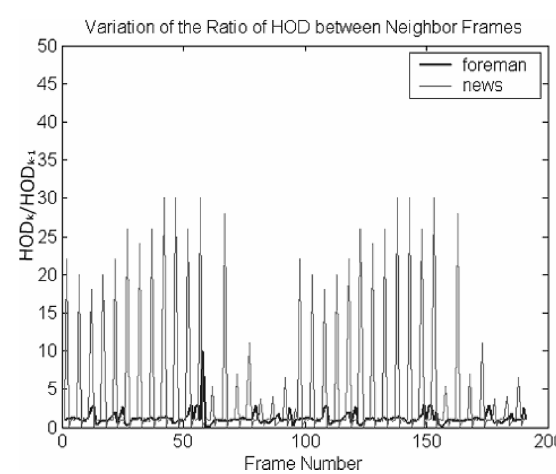

Fig. 11. Comparison of local motion variation in neighbor frames.

Fig. 11 shows the variations of local motion in neighbor frames for sequence "foreman" and "news." Here the local

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.

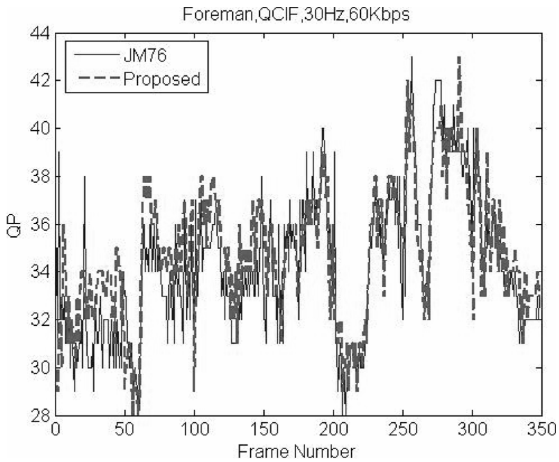

Fig. 12. Comparison of QP distribution variance.

motion variation (LMV) is indicated by the ratio of neighbor HODs:

$$\text{LMV} = \frac{\text{HOD}_{i-1}}{\max(MIN\text{HOD}, \quad \text{HOD}_i)}, N_p \geq i \geq 1$$

where MINHOD is set to be 0.0005 in the paper so as to prevent the divisor being zero. For the first P-frame in a GOP, LMV is set to 1.

It can be seen clearly that compared to sequence "news," the local motion variation of "foreman" is much smaller. Accordingly, the PSNR improvement of "news" by using the proposed scheme exceeds the "foreman" greatly. Generally, for the proposed scheme in the paper, the stronger local motion variation of a sequence is, the bigger R-D improvement can be achieved.

Fig. 12 shows the comparison of QP variations between the proposed scheme and JM7.6 over 3 GOPs for sequence "foreman" (target bit rate 40 kbps, 30 fps), from which it can be seen that strength of QP variations of the proposed scheme is a little bit higher than JM76. This is due to the fact that the proposed scheme is sensitive to local motion change.

Further experimental results are reported in Table II, in which two category experiments are carried out: with or without the new initial QP estimation scheme. The following is noted for the new techniques proposed in the paper. 1) Improvement of PSNR is higher in low bit rate than in high bit rate, which is due to the fact that in high bit rate, the encoding complexity estimation scheme is less effective since texture bits make up a large scale. 2) The initial QP estimation has quite a great impact on PSNR, which varies from sequence and target bit rate. 3) Generally speaking, the PSNR is improved (especially in low bit rate, it is up to 1.53 dB); 4) Subjective tests also show very significant improvement, with higher video quality and smoothness throughout the whole sequence thanks to the adaptation to local motion more effectively.

In order to illustrate the performance of the proposed scheme further, we compare the PSNR results of the proposed scheme with Yuan's method [8]. Since in Yuan *et al.* scheme JM10 is used as benchmark software, we also use JM10 in this simulation. The total frames encoded are 150 frames and the length of GOP is 50. The simulation results are given in Fig. 13, from which we can see that in low bit rate coding, the PSNR performance of the proposed scheme is a little better than Yuan's method; while in higher bit rate, the performance of the proposed scheme is inferior to Yuan's methods. Such a fact shows that the proposed scheme is more effective in low bit rate video coding.

TABLE II
EXPERIMENTS OF TEST SEQUENCES, WHERE "NEW QP" INDICATES THAT THE PROPOSED INITIAL QP ESTIMATION SCHEME IS USED AND "OLD QP" INDICATES THAT THE DEFAULT INITIAL QP SCHEME IN H.264 IS USED

Test Name	H.264 PSNR	Proposed PSNR		Gain	
		New QP	Old QP	New QP	Old QP
F60	32.51	32.71	32.59	+0.20	+0.08
F75	33.37	33.64	33.43	+0.27	+0.06
F100	34.77	34.80	34.79	+0.03	+0.02
F150	36.43	36.44	36.49	+0.01	+0.06
S45	34.71	34.84	34.80	+0.13	+0.09
S65	36.16	36.30	36.20	+0.14	+0.04
S95	37.86	37.89	37.93	+0.03	+0.07
S135	39.28	39.32	39.32	+0.04	+0.04
N25	30.16	30.55	30.36	+0.39	+0.20
N60	34.63	35.76	34.98	+1.13	+0.35
N90	38.00	38.02	38.05	+0.02	+0.05
N130	40.06	40.27	40.24	+0.21	+0.18
H25	32.25	32.69	32.46	+0.44	+0.21
H45	35.35	36.26	35.98	+0.91	+0.63
H70	37.27	38.56	38.01	+1.29	+0.74
H100	40.01	39.99	40.02	-0.02	+0.01
A20	34.96	35.11	35.08	+0.15	+0.12
A30	36.89	37.83	37.42	+0.97	+0.53
A50	39.10	40.63	39.65	+1.53	+0.55
A80	42.84	43.03	42.94	+0.19	+0.10
Mi25	38.51	38.66	38.49	+0.15	-0.02
Mi50	41.00	41.58	41.10	+0.58	+0.10
Mi80	43.07	43.16	43.08	+0.09	+0.01
Mi100	43.72	43.84	43.79	+0.12	+0.07
Mo30	32.70	32.79	32.72	+0.09	+0.02
Mo50	34.57	34.93	34.68	+0.36	+0.11
Mo70	35.85	36.42	35.96	+0.57	+0.11
Mo100	37.99	37.97	37.98	-0.02	-0.01
C100	39.28	39.24	39.23	-0.04	-0.05
C80	38.45	38.46	38.42	+0.01	-0.03
C50	35.96	36.58	36.00	+0.62	+0.02
C30	34.19	34.50	34.22	+0.31	+0.03

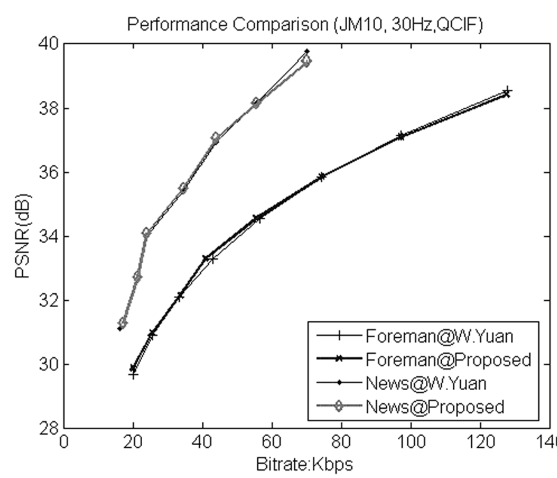

Fig. 13. Comparison of PSNR performance between the proposed scheme and Yuan's scheme. Note that JM10 is used in the simulation.

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.

V. Conclusion

As to the defects in rate control scheme of H.264, the paper presented a few new techniques to improve the rate control scheme proposed by Li *et al.* These improvements are summarized as follows.

1) Adaptive initial QP selection scheme is introduced to choose QP for I-frame according to target bit rate, the length of GOP and the statistic feature of sequence itself. Thus, a more reasonable initial QP can be estimated so as to obtain a higher R-D performance.
2) HOD-based frame layer target bits allocation is introduced so as to adapt bits allocation to local motion change more effectively. With this method, the frame layer bits allocation can be more closely to the actual bits used so that the buffer level control could be more accurate. In addition, R-D performance is also improved with such a scheme.
3) A simple QP modification scheme is proposed based on encoding complexity prediction, which overcome the defect of analytical R-Q model in low bit rate and improve the R-D performance.
4) Since the complexity of the proposed QP modification scheme is much lower than the one in JM7.6 (when the sequence is QCIF, the ratio of complexity reduction is about 3–4 times), the overall complexity of rate control is reduced greatly.

Experimental results show that the improved algorithm significantly improved the R-D performance in low bit rate, reduced the buffer level variation and improved the perceptual quality. The proposed scheme can achieve better performance than the most recent scheme proposed by Yuan *et al.* [8] in low bit rate. Although this paper only targeted at low bit rate video coding in constant bit rate (CBR) applications, it is applicable to VBR situations. In addition, the idea presented in the paper is also effective to the other video coding recommendations [17].

Appendix
Deduction of the Optimal L

In CBR applications, rate control aims at keeping bit-rate to be constant. Therefore, it is reasonable to assume that all P-frames in a GOP use the same bits approximately, which can be expressed as according to (6) in Section II-A

$$\frac{\delta_i^2}{Q_i^2} = \frac{\delta_{i+1}^2}{Q_{i+1}^2}, \qquad 1 \le i \le M-1.$$

Then, according to the assumptions in Section II-A, the following equations can be deduced from the assumptions (3), (5), and (6):

$$Q_i^2 = \frac{L\delta_i^2 Q_0^2}{\delta_0^2}, \qquad 1 \le i \le M-1$$
$$\sum_{i=0}^{M-1} \frac{\delta_i^2}{Q_i^2} = \left(\frac{M-1}{L}+1\right)\frac{\delta_0^2}{Q_0^2}. \tag{A.1}$$

Plug (5), (6), and (A.1) into (4), provides the result:

$$J(Q_0^2, L) = \frac{Q_0^2}{12} + \frac{LQ_0^2}{12\delta_0^2}\sum_{i=1}^{M-1}\delta_i^2 + \lambda\frac{e}{\ln 2}\left(\frac{M-1}{L}+1\right)\frac{\delta_0^2}{Q_0^2} - \lambda B. \tag{A.2}$$

Then, the minimum of J is obtained by setting its partial derivatives to 0, i.e.,

$$\begin{cases} \frac{\partial J}{\partial Q_0^2} = \frac{1}{12} + \frac{L(M-1)\delta_\mu^2}{12\delta_0^2} - \lambda\frac{e}{\ln 2}\left(\frac{(M-1)}{L}+1\right)\frac{\delta_0^2}{(Q_0^2)^2} = 0 \\ \frac{\partial J}{\partial L} = \frac{(M-1)\delta_\mu^2 Q_0^2}{12\delta_0^2} - \lambda\frac{e}{\ln 2}\frac{(M-1)\delta_0^2}{L^2 Q_0^2} = 0 \end{cases} \tag{A.3}$$

where $\delta_\mu^2 = 1/M-1\sum_{i=1}^{M-1}\delta_i^2$ is the mean variance of P-frames in a GOP.

From (A.3), (7) can be deduced easily.

Acknowledgment

The authors would like to thank Associate Editor H. Watanabe for his laborious work.

References

[1] *MPEG-2, Test Model 5*, Doc.ISO/IEC JTC1/SC29/WG11/93-400, 1993.
[2] S. W. Ma *et al.*, "Rate control for advance video coding (AVC) standard," in *Proc. ISCAS*, May 2003, vol. 2, pp. 892 –895.
[3] Z. G. Li, F. Pan, K. P. Lim, G. Feng, X. Lin, and S. Rahardja, *Adaptive Basic Unit Layer Rate Control for JVT* Pattaya II, Thailand, 2003, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6-G012.
[4] S. W. Ma *et al.*, "Rate control JVT video coding scheme with HRD considerations," in *Proc. IEEE ICIP*, Spain, Sep. 2003, vol. 3, pp. 793–796.
[5] C. Grecos and J. Jiang, "On-line improvement of the rate-distortion performance in MPEG-2 rate control," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 6, pp. 519–528, Jun. 2003.
[6] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 246–250, Feb. 1997.
[7] *ITU-T/SG15. Video Codec Test Model, Near-Term, TMN8* ITU Study Group 16, Video Coding Experts Groups, Portland, USA, 1997, Document Q15-A-59.
[8] W. Yuan *et al.*, "Optimum bit allocation and rate control for H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 6, pp. 705–715, Jun. 2006.
[9] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders," *IEEE Trans. Image Process.*, vol. 3, no. 5, pp. 533–545, Sep. 1994.
[10] JVT (2004) [Online]. Available: http://bs.hhi.de/~suehring/tml/download/jm76
[11] S. Zhou, J. Li, and Y. Zhang, "Efficient quantization step selection scheme for I-frame in rate-constrained video coding," in *Proc. IEEE ICIP*, Genova, Italy, Sep. 2005, pp. 321–324.
[12] S. Zhou, J. Li, and Y. Zhang, "Improvements on rate-distortion performance of H.264 rate control in low bit rate video coding," in *Proc. IEEE ICME*, Toronto, ON, Canada, Jul. 2006, pp. 681–684.
[13] F. Pan *et al.*, "An intra-frame quantization optimization algorithm for very low bit rate video coding" Inst. Infocomm Res., Singapore, 2002.
[14] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 172–185, Feb. 1999.
[15] S. W. Ma, *Proposed Draft Description of Rate Control on JVT Standard* Awaji, 2002, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6-G018.
[16] J. Lee and B. W. Dickinson, "Temporally adaptive motion interpolation exploiting temporal masking in visual perception," *IEEE Trans. Image Process.*, vol. 3, no. 5, pp. 513–526, Sep. 1994.
[17] H. Song and C. C. J. Kuo, "A region-based H.263+ codec and its rate control for low VBR video," *IEEE Trans. Multimedia*, vol. 6, no. 3, pp. 489–500, Jun. 2004.
[18] S. Zhou *et al.*, *Proposed Draft Description of Rate Control Scheme for Low Bit Rate Video Coding Applications* Digital Audio/Video Standard Workgroup, Su Zhou, China, 2004, AVS_M, Doc-1416.

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.

ShuMin Zhou received the Ph.D degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in June 2006.

Currently, he is an Assistant Professor in Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences. His major research fieldd include graphics and image processing, video codec, and wireless video delivery.

JinTao Li received the Ph.D degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China.

Currently, he is a Professor in the Institute of Computing Technology, Chinese Academy of Sciences. His major research fields include multimedia processing and VR technology.

JinHao Fei received the M.S. degree from Dalian University of Technology, Dalian, China. He is currently working toward the Ph.D. degree in the nstitute of Computing Technology, Chinese Academy of Sciences, Beijing, China.

His major research fields include multimedia processing and computer graphics.

YongDong Zhang received the Ph.D degree from Tianjin University, Tianjin, China, in 2002.

Currently, he is an Associate Professor in the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. His major research fields include image processing and video processing.

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 13, 2009 at 00:57 from IEEE Xplore. Restrictions apply.