IMPROVEMENTS ON RATE-DISTORTION PERFORMANCE OF H.264 RATE CONTROL IN LOW BIT RATE VIDEO CODING

ShuMin Zhou[1,2] *JinTao Li*[1] *YongDong Zhang*[1]

[1]Digital Technology Lab, Institute of Computing Technology, Chinese Academy of Sciences
P. O. Box 2704, Beijing, 100080, P. R. China, email: smzhou@ict.ac.cn
[2]Graduate School of the Chinese Academy of Sciences

ABSTRACT

This paper points out some defects in the techniques used in H.264 rate control and presents two new techniques to improve them. The improved scheme 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 paramter (QP) calculation is based on a simple encoding complexity prediction scheme, which is more robust and simple than the quadratic model used by H.264 in low bit rate video coding. Simulation results show that compared to rate control scheme in H.264, the improved scheme has significantly improved R-D performance (up to 1.29dB).

1. INTRODUCTION

Rate control in video coding has been the focus of research in recent years [1]-[6]. Most of the developed rate control schemes are composed of three major parts: target bits estimation (bits allocation), buffer level control and QP calculation. Accordingly, the rate control of H.264 is composed of three parts: a frame bits allocation scheme based on [1], a fluid flow buffer level control scheme and a QP calculation scheme based on [6]. Although the techniques used in the rate control of H.264 are state-of-the-art, there are serious defects in them, especially for low bit rate video coding. This paper will present two new techniques to deal with these defects so as to improve the rate-distortion (R-D) performance of rate control. In the following, we will have a brief look at the main techniques used in H.264 rate control and point out the defects in them.

The target bits estimation (also known as frame layer bits allocation) mainly focused on how to effectively allocate bits for each frame. In H.264, 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:

$$X_i \approx X_{i-1}, X_i = B_i * QP_i \tag{1}$$

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

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 lead to poor R-D performance when the local motion changes abruptly between the neighbor frames.

The buffer level control aims at preventing encoding buffer from overflowing or underflowing. In H.264, 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]:

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

where $Tbl(i,j)$ indicates the target buffer level for frame j in GOP i, B_s is the size of encoding buffer, N_p is the number of coded P frames in current GOP.

The QP caculation, which is the key part of rate control, is responsible for calculating QP according to the target bits and buffer level requirement. As to rate control in H.264, a quadratic rate-quantization paramter model (R-Q model), which is known as MPEG-4 Q2 [6], is used to calculate QP in basic unit layer (frame layer control can be seen as a special case of basic unit) [3]. However, the Q2 model is inaccurate in low bit rate video coding since the local motion information, which is distributed randomly over the whole GOP, is the major part in total bits in low bit rate coding. Thus, such an analytical model would deteriorate rate-distortion performance also.

In consideration of all the defects mentioned above, two new techniques are proposed to improve the rate-distortion performance of rate control for H.264. Firstly, 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. Secondly, 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

1-4244-0367-7/06/$20.00 ©2006 IEEE

according to the distribution of its neighbor *HOD* and encoding complexity.

The rest of the paper is organized as follows. In section 2, as to the defects in H.264 rate control mentioned above, two new techniques are presented in detail. Section 3 gives the improved rate control scheme. The experimental results are given in section 4. Finally, our summary and conclusion are presented in section 5.

2. PROPOSED TECHNIQUES

In the section, first, a *HOD*-based frame layer bits allocation scheme is given. Secondly, in order to deal with the inaccuracy of analytical R-Q model in low bit rate, we propose a simple QP calculation scheme based on encoding complexity prediction, which is performed on basic unit layer.

2.1. HOD-based Frame Layer Bits Allocation

In this section, a frame layer bits allocation scheme based on *HOD* is proposed. The *HOD* between frame *m* and *n* provides as follows [7]:

$$HOD(f_n, f_m) = \frac{\sum_{i \notin [-\alpha, \alpha]} hod(i)}{N_{pix}} \tag{3}$$

here *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 is threshold for level *i*.

As a matter of fact, Hwangjun Song *et al* [8] has proposed such a scheme, in which the target bits is allocated over the whole GOP [8]:

$$T_k = \left(1 + \frac{HOD_k - \mu_{hod}}{\mu_{hod}}\right) \cdot \frac{B_{gop}}{N_{gop}} \tag{4}$$

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, N_{gop} is the length of GOP.

Obviously, N_{gop} frames should be buffered before calculate target bits allocation for frame *k* in (4), which leads to long delay in encoding process. In low delay communication, such a scheme isn't feasible.

To eliminate the delay induced by (4), an alternative method is presented, in which the overall average *HOD* is replaced by progressive average.

$$T_k = \left(1 + \frac{HOD_k - \frac{1}{k}\sum_{i=1}^{k} HOD_i}{\frac{1}{k}\sum_{i=1}^{k} HOD_i}\right) \frac{B_r}{N_r} \tag{5}$$

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

Although (5) is sub-optimal compared to (4), it eliminates the encoding delay completely. Besides, compared to (4), the degradation of R-D performance in (5) is negligible, which will be shown in section 4.

When bits are used out in encoding process, the bits assignment will be negative which would lead to the quality deterioration of subsequent frames. Besides, 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 (5) for each frame:

$$T_k = \max(LEAST\_BITS, \min(T_k, MAX\_BITS)) \tag{6}$$

where the *LEAST_BITS* and *MAX_BITS* are set to be 96 and *2*bit_rate/frame_rate*, here the *bit_rate* indicates the current target bit rate and *frame_rate* indicates the frame frequency in video coding.

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

$$T_k^{'} = \frac{bit\_rate}{frame\_rate} - \gamma(BL_k - Tbl_k) \tag{7}$$

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:

$$T_k = \Delta T_k + (1 - \Delta) T_k^{'} \tag{8}$$

here Δ=0.5 is weighted coefficient.

2.2. New QP Estimation Scheme

In the section, 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.

2.2.1 Simple Encoding Complexity Estimation

Since motion information takes up a large part in total bits of inter frames in low bit rate, we have the following observations:

(1) Encoding complexity of current unit is approximately proportional to local motion degree, which can be indicated by *HOD*

(2) There are great correlation between the encoding complexity for current unit and its spatial and temporal neighbors

Thus, on one hand, the following relation is reasonable according to the first assumption:

$$\frac{HOD_{i-1,k-1}}{HOD_{i-1,k}} \approx \frac{HOD_{i,k-1}}{HOD_{i,k}} \Leftrightarrow \frac{X_{i-1,k-1}}{X_{i-1,k}} \approx \frac{X_{i,k-1}}{X_{i,k}} \tag{9}$$

here $X_{i,k}, HOD_{i,k}$ stand for the encoding complexity and *HOD* of *k*-th basic unit in *i*-th frame respectively. In the following, the subscipt *i,k* indicates *k*-th coding unit in *i*-th frame.

Therefore, the current encoding complexity can be provided:

$$\left|\frac{HOD_{i-1,k-1}}{HOD_{i-1,k}} - \frac{HOD_{i,k-1}}{HOD_{i,k}}\right| < \beta \Rightarrow X_{i,k} = X_{i,k-1}\frac{X_{i-1,k}}{X_{i-1,k-1}} \tag{10}$$

where β is threshold, which is an experimental threshold and set to be 0.2 in the paper

On the other hand, if the prerequisite of (10) 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 $\left|HOD_{i,k-1} - HOD_{i,k}\right| - \left|HOD_{i-1,k} - HOD_{i,k}\right| > 0.1HOD_{avr}$, then
    $X_{i,k} = (2 * X_{i-1,k} + X_{i,k-1})/3$
else if
    $\left|HOD_{i,k-1} - HOD_{i,k}\right| - \left|HOD_{i-1,k} - HOD_{i,k}\right| < -0.1HOD_{avr}$, then
    $X_{i,k} = (X_{i-1,k} + 2 * X_{i,k-1})/3$
else
    $X_{i,k} = (X_{i-1,k} + X_{i,k-1})/2$
```

here HOD_{avr} is the average *HOD* of current frame.

2.2.2 HOD-based Bits Estimation

As to the bits for current coding units, we can estimate it based on HOD, which is presented as follows:

$$Bits_{i,k} = \left(1 + \frac{HOD_{i,k} - HOD_{avr}}{HOD_{avr}}\right) * T_i \tag{11}$$

If the bits obtained from (11) is negative, QP for current unit is set to be *QPAverage+2* directly, here *QPAverage* is the average QP value. Otherwise, the QP estimation for current coding unit is given as:

$$QP_{i,k} = \max(2, \min(\frac{X_{i,k}}{Bits_{i,k}}, 51)) \tag{12}$$

3. 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.

Step 1. 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 proceeding one. With these *HOD* value, *HOD* between current frame and its proceeding one can be obtained. Then bits allocation T_k can be made out from (5).

Step 2. Buffer Level Control

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 (8).

Step 3. Basic unit layer QP modification

(1) 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.

(2) Otherwise, allocate bits for current basic unit by (11).

(3) If the bits allocated is negative, we set the current QP to *QPAverage+2* directly. Otherwise, predict the encoding complexity of current basic unit by using (10) and the pseudo code.

(4) Calculate QP for current basic unit by (12), which is indicated by *Qc*

Truncate *Qc* in order to keep smoothness of video quality:

$$Qc = MIN(Qc, QPAverage + 3) \tag{13}$$

4. SIMULATION RESULTS

The test sequences used are all in QCIF format, with I frame appearing every 50 frames. Note that no frame skip is concerned here. The total frames are 150 (3 GOP) for each sequence. The reference software H.264 JM76 [9] is used for simulation.

Table 1 shows that the comparison of control precision between the proposed scheme and H.264 rate control. All sequences are 30Hz. It can be seen that the precision of rate control is approximately the same.

Table 1. Comparison of control precision

Test Name	Video Sequence	Target Bit rate(kbps)	Actual bit rate(kbps)	
			H.264	Proposed
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

Fig. 1 compares the PSNR of sequence ‘News’ among the proposed scheme, H.264 (bit rate is 60kbps) and Hwangjun et al’s method [8]. It can be seen that, 1)compared to H.264 rate control, the overall PSNR of the proposed scheme is better; 2) as is mentioned in section II, in order to compare the performance between the bits allocation scheme in (4) and (5), we also plot the PSNR for sequence “foreman” in Fig.1, from which we can see that compared to the bit allocation scheme proposed by Hwangjun *et al* in [8], the suboptimal method proposed in the paper only leads to a little degradation (less than 0.05dB on average).

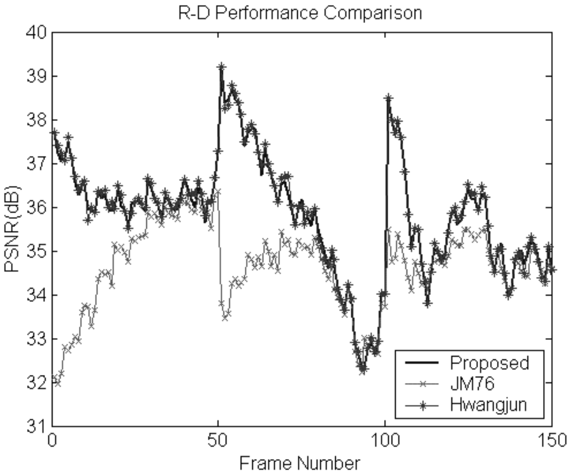

Fig.1 PSNR comparison among H.264, Hwangjun and the proposed rate control scheme (News)

Table 2 compares the R-D performance between the proposed scheme and H.264 rate control, from which we can see that the proposed scheme achieves obvious improvement on R-D performance in low bit rate (up to 1.29dB). Note that the best performance for Hall and News are obtained at 60kbps and 70kbps respectively, which says that the proposed scheme may have better performance not only in low bit rate but also median case.

5. CONCLUSION

As to the defects in rate control scheme of H.264, the paper presented a few new techniques to improve its R-D performance.

Experimental results show that: 1) compared to H.264 rate control, the proposed algorithm significantly improved the R-D performance in low bit rate while keeping the control precision; 2) compared to Hwangjun et al's bit allocation method, the proposed suboptimal methods can eliminate the coding delay while keeping the R-D performance to be nearly unchanged..

Although this paper only targets at low bit rate video coding in CBR applications, it is applicable to VBR as well.

Acknowledgement

This work is supported by National Nature Science Foundation of China under grant number 60302028 and the key program of International Science and Technology Cooperation under grant number 2005DFA11060.

Table 2. R-D simulation results for other sequences

Test Name	H.264 PSNR	Proposed PSNR	Gain
F60	32.51	32.71	+0.20
F75	33.37	33.64	+0.27
F100	34.77	34.80	+0.03
F150	36.43	36.44	+0.01
S45	34.71	34.84	+0.13
S65	36.16	36.30	+0.14
S95	37.86	37.89	+0.03
S135	39.28	39.32	+0.04
N25	30.16	30.55	+0.39
N60	34.63	35.76	+1.13
N90	38.00	38.02	+0.02
N130	40.06	40.27	+0.21
H25	32.25	32.69	+0.44
H45	35.35	36.26	+0.91
H70	37.27	38.56	+1.29
H100	40.01	39.99	-0.02

REFERENCES

[1] MPEG-2, Test Model 5, Doc.ISO/IEC JTC1/SC29/WG11/93-400, April, 1993.

[2] S.W.Ma, "Proposed draft description of rate control on JVT standard", ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6-G018, Awaji, 5-13, December, 2002.

[3] Z.G Li, F. Pan, K. P. Lim, G. Feng, X. Lin and S. Rahardja, "Adaptive Basic Unit Layer Rate Control for JVT", ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6-G012, Pattaya II ,Thailand, 7-14 March, 2003

[4] S.W. Ma et al, "Rate Control JVT Video Coding Scheme With HRD Considerations". IEEE International Conference on Image Processing (ICIP), vol.3, pp. 793-796, Spain , September 2003.

[5] Christos Grecos and Jianmin Jiang. "On-line improvement of the rate-distortion performance in MPEG-2 rate control", IEEE Trans. On CSVT, Vol . 13, No. 6, June 2003, pp.519~528

[6] Tihao Chiang and Ya-Qin Zhang, "A New Rate Control Scheme Using Quadratic Rate Distortion Model", IEEE Trans. On CSVT, Vol.7, No.1, Feb 1997, pp.246-250.

[7] Jungwoo Lee and Bradley W. Dickinson, "Temporally Adaptive Motion Interpolation Exploiting Temporal Masking in Visual Perception". IEEE Trans. On Image Processing, Vol 3. No.5. September 1994, p513-526

[8] Hwangjun Song and C.C Jay Kuo. "A Region-Based H.263+ Codec and Its Rate Control for Low VBR Video". IEEE Trans. On Multimedia, Vol. 6. No. 3. June 2004, p489~500

[9] JVT ,http://bs.hhi.de/~suehring/tml/download/jm76,2004