

# A Novel Rate Control Scheme for Constant Bit Rate Video Streaming

Venkata Phani Kumar M  
Department of E & ECE  
IIT Kharagpur, India  
phanikumar@iitkgp.ac.in

K C Ravi  
Department of E & ECE  
IIT Kharagpur, India  
kcravi912@gmail.com

Sudipta Mahapatra  
Department of E & ECE  
IIT Kharagpur, India  
sudipta.mahapatra@gmail.com

## ABSTRACT

In this paper, a novel rate control mechanism is proposed for constant bit rate video streaming. The initial quantization parameter used for encoding a video sequence is determined using the average spatio-temporal complexity of the sequence, its resolution and the target bit rate. Simple linear estimation models are then used to predict the number of bits that would be necessary to encode a frame for a given complexity and quantization parameter. The experimental results demonstrate that our proposed rate control mechanism significantly outperforms the existing rate control scheme in the Joint Model (JM) reference software in terms of Peak Signal to Noise Ratio (PSNR) and consistent perceptual visual quality while achieving the target bit rate. Furthermore, the proposed scheme is validated through implementation on a miniature test-bed.

## Categories and Subject Descriptors

H.4.3 [Information Systems Applications]: Communications Applications

## General Terms

Algorithms, Experimentation

## Keywords

Video Streaming, Spatio-Temporal Complexity, PSNR

## 1. INTRODUCTION

With the advancement of video coding and communication technologies, video streaming applications have become more and more popular in consumer electronics devices. These can be divided into real-time video streaming and non-real-time video streaming. Encoding speed is imperative in real-time video streaming because video needs to be transmitted on the fly. Over the past few decades, digital video compression techniques are extensively used in video stream-

ing, television broadcast and digital storage. Several international industry standards have been established such as MPEG-1/2/4, H.263, H.264 for efficient video coding. The H.264 video coding standard, a successor to MPEG-4 Part 2, that was jointly developed by the ISO/IEC Moving Picture Experts Group (MPEG) and ITU-T Video Coding Experts Group (VCEG), provides good video quality at substantially lower bit rates than previous standards [5].

For all of the video encoders, rate control plays a key role in video coding. In general, rate control algorithms regulate rate by adjusting the Quantization Parameter (QP). In H.264 and many other video compression standards, the QP can be changed at both the frame and the macroblock levels. Rate control should provide consistent perceptual visual quality under buffer constraints and a given target bit rate. For video streaming applications, it is desirable that video should be encoded not only with a good average quality, but also with fewer quality fluctuations among the adjacent frames. In this paper, a novel rate control scheme is proposed for video streaming applications constrained by channel bandwidth. Since the proposed algorithm results in high-speed encoding and has a low complexity, it is more suitable for real-time applications. The proposed rate control model provides consistent visual quality for a given target bit rate. After encoding each frame, the rate control parameters are updated with linear estimation models to predict the QP accurately for the next frame. The remainder of this paper is organized as follows. Section II presents the background and related work. Section III explains the proposed coding technique. The proposed rate control scheme is discussed in Section IV. Section V presents experimental results for the proposed algorithm. Finally, Section VI concludes the paper.

## 2. BACKGROUND AND RELATED WORK

Use of hierarchical pictures leads to an improved coding efficiency compared to coding with the classical IBBP or IPPP structures for a wide range of test sequences [4]. Furthermore, the usage of hierarchical pictures improves not only the objective quality for the majority of test sequences, but also the subjective quality, especially for sequences with a high spatial detail and slow motion. The coding structures with hierarchical pictures have been integrated with H.264/AVC to improve its coding performance. Currently, hierarchical pictures have been adopted in SVC as an extension of the H.264/AVC standard to improve the coding efficiency as well as to provide different temporal resolutions. Hierarchical pictures are also used as the coding struc-

ture both in temporal and inter-view predictions in Multi-view Video Coding (MVC) in order to exploit the temporal and inter-view redundancies respectively. In the past few years, some research efforts on the rate control for hierarchical frames coding in H.264/AVC have been reported [3, 1]. We consider the rate control algorithm within the JM reference software [2] as the contemporary rate control algorithm. This rate control scheme's primary goal is to achieve the target bit rate without considering the visual quality. In this scheme, the bit budget for the current frame is first computed as the ratio of the target bit rate to the frame rate of the sequence. An estimate of the header bits is subtracted from the bit budget to determine the texture bit budget. Header bits denote all of the bits in the bit stream apart from those used to encode the texture. These include the bits needed to represent mode and motion information. The texture bit target is translated into a QP value with a quadratic model that links the QP to the texture bits. A linear model estimates the current picture prediction error from that of the previous picture. Finally, the derived QP value is used to encode all of the slices in the current picture. Statistics such as the actual prediction error and the number of header and texture bits are gathered for the coded picture. These are then used to update both the linear and the quadratic models.

In this paper, a new rate control scheme for hierarchical pictures is proposed. To achieve the target bit rate with the best quality, we start rate control with a QP based on the complexity of the sequence, its resolution and the target bit rate. Simple linear estimation models are used to predict the number of bits that would be necessary to encode a frame for a given complexity and quantization parameter. The complexity of the proposed rate control mechanism is also much lower than that of the contemporary rate control scheme since the proposed rate control employs a simple linear prediction model rather than the complicated MAD prediction used in the latter. The proposed algorithm also achieves consistent visual quality while satisfying the long term constant bit rate constraint.

### 3. PROPOSED CODING SCHEME

In the proposed coding scheme each of the frames to be encoded, is assigned a unique rank. Further, we categorize the coding scheme into either Hierarchical B Rank (HBR) or Hierarchical P Rank (HPR). In HBR coding scheme all the I frames and P frames are assigned a rank zero while ranks assigned to B frames vary from rank one to a maximum rank value. In HPR coding scheme all of the I frames are assigned a rank zero, B frames are assigned the maximum rank, while ranks assigned for P frames vary from rank zero to a rank less than the maximum rank value as shown in Fig. 1 and Fig. 2 respectively.

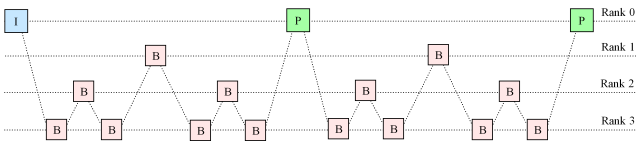


Fig. 1: HBR Coding Scheme

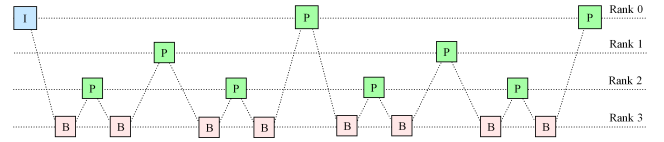


Fig. 2: HPR Coding Scheme

The QP for any frame of rank  $i$  is derived by adding a small value to the QP of the corresponding reference frames of rank  $i - 1$  irrespective of the frame type. We maintain the QP variation between zero rank reference frames as shown in Fig. 1 and Fig. 2. Further, the proposed rate control scheme modulates the zero rank frame QPs according to the target bit rate while the rest of QPs will be derived as discussed above. With this coding structure, a good rate control model can be designed which can keep the actual bit rate close to the target bit rate while mainly regulating the QPs for zero rank frames.

### 4. PROPOSED RATE CONTROL SCHEME

The simplest method of rate control is to start with a predefined initial QP for any sequence irrespective of the target bit rate. This method is not appropriate for the proposed hierarchical coding scheme because the QP for any frame of rank  $i$  is derived from the QP of the corresponding reference frames of rank  $i - 1$ . Therefore, an adaptive initial QP estimation method is developed for the hierarchical coding structures. In [3], for hierarchical coding structures, an analytical model is established to set an appropriate initial QP as follows:

$$\log \hat{Q} = a * \log K + b \quad (1)$$

where  $a$  and  $b$  are constants,  $\hat{Q}$  is the initially estimated QP and  $K$  is the ratio of the target bit rate per frame rate to the resolution of the sequence.

However to achieve the target bit rate with the best quality, it is necessary to start rate control with the QP based on the complexity of the sequence. Following the work reported in [3], we encoded several sequences of various resolutions with fixed QPs. Based on the experimental results we constructed an analytical model to set an appropriate initial QP as follows:

$$\log \hat{Q} = k_1 * \log C + k_2 \quad (2)$$

where  $k_1$  and  $k_2$  are constants,  $\hat{Q}$  is the initially estimated QP and  $C$  is the ratio of the average spatio-temporal complexity to the target bit rate. The average spatio-temporal complexity is an estimate of the spatio-temporal complexity of the future frames. The QP derived in Equation (2) is used as the initial QP for encoding. This QP is assigned as zero rank P frame QP, while the rest of QPs are derived as discussed above.

Let  $\bar{P}_{QP}$  be the average of zero rank P frame QPs encoded so far,  $\bar{cplx}$  be the average of spatio-temporal complexity for all the frames encoded so far,  $bit_{act}$  be the total bits consumed for all the frames encoded so far,  $cplx$  be the spatio-temporal complexity of the current frame and  $bit_{tgt}$  denote the total target bits for all of the frames including the current frame. Then, the QP required to encode the zero rank P frame such that  $bit_{act}$  are close to  $bit_{tgt}$  is given by

$$P_{QP} = \frac{bit_{act} * \bar{P}_{QP} * cplx}{bit_{tgt} * \bar{cplx}} \quad (3)$$

Simple linear estimation models are used in rate control to predict the number of bits that would be necessary to encode a frame for a given complexity and quantization parameter of specific frame type and rank. For a given frame of specific type and rank, the estimated number of bits is given by

$$bit_{est} = \frac{\alpha * cplx}{\beta * QP} \quad (4)$$

where  $\alpha$  and  $\beta$  are estimation model parameters, that corresponds to frame type and rank and are updated after encoding each frame. Let  $P_{QP_T}$  be the target  $P_{QP}$  that gives the ideal target bit rate. Before encoding a frame, a buffer verifier is called to estimate the status of the buffer if the current frame is encoded with the QP derived from  $P_{QP}$ . Depending on the estimated buffer status, the current  $P_{QP}$  is either left unchanged or is modified so as to avoid the buffer overflow or underflow. Buffer control tries to modulate the current  $P_{QP}$  to keep it close to  $P_{QP_T}$  whenever there is buffer fluctuation due to a variation in scene complexity. Whenever buffer control modifies the current  $P_{QP}$ , all the QPs corresponding to different frame types and ranks are updated. A Coding Group (CG) is considered to represent a set of frames from the frame succeeding the past zero rank frame to the future zero rank frames. For all the frames in the current CG, bits are estimated using Equation (4) depending on the current state of coding and the QP derived from the current  $P_{QP}$ . If the current buffer level is  $buf_{cur}$ , then continuing with the current  $P_{QP}$  will lead to an estimated buffer level,  $buf_{est}$  given by

$$buf_{est} = buf_{cur} + buf_{incr} \quad (5)$$

where  $buf_{incr}$  is increment in buffer level. If  $buf_{est}$  is close to the target buffer,  $buf_{tgt}$  for the CG, the current  $P_{QP}$  should be used for encoding the CG. Otherwise, the algorithm derives a small increment in  $P_{QP}$  ( $\zeta_{qp}$ ) that needs to be applied to the current  $P_{QP}$  to take the buffer close to the target buffer. This increment is derived as follows:

$$\zeta_{qp} = P_{QP} * \zeta_{qpf} \quad (6)$$

where the step factor,  $\zeta_{qpf} = ((1 + \zeta_{br}) * (1 + \zeta_{vq})) - 1$ . The two quantities  $\zeta_{br}$  and  $\zeta_{vq}$  respectively correspond to the bit rate constraint and the consistent visual quality constraint as discussed below.

- $\zeta_{br}$  depends on how close the estimated buffer level is to the target buffer level. It is computed from the saturation status of the buffer ( $buf_{sat}$ ) as given below.

$$buf_{sat} = \frac{(buf_{est} - buf_{tgt})}{buf_{tgt}} \quad (7)$$

where  $buf_{tgt}$  is the target buffer level that gives the ideal target bit rate.

- $\zeta_{vq}$  corresponds to scene complexity and is computed from  $buf_{incr}$  as follows:

$$buf_{incr} = buf_{est} - buf_{last} \quad (8)$$

where  $buf_{last}$  is the buffer status at the beginning of the current CG. The algorithm updates  $buf_{est}$  and  $buf_{last}$  at the instance of each zero rank reference frame compute the buffer increment from this point for each frame in the CG.

After estimating the increment in  $P_{QP}$ , it is modulated as follows:

$$P'_{QP} = P_{QP} + \zeta_{qp} \quad (9)$$

where  $P'_{QP}$  is the updated  $P_{QP}$ . After encoding of each of the frames, the coding state and the buffer status are updated depending on the actual number of bits produced.

$$buf_{cur} = buf_{cur} + (bit_{act} - bit_{tgt}) \quad (10)$$

## 5. EXPERIMENTAL RESULTS

The proposed rate control (PRC) scheme was used with JM 18.6 reference software. The Rate-Distortion (RD) performance of all the test algorithms is shown in Fig. 3. From Fig. 4, it can be observed that the proposed rate control scheme achieves a smoother PSNR variation than the traditional rate control (TRC) mechanism. The QP and PSNR variations are represented by their Standard Deviations (SDs) are tabulated in Table 1 for all rate control schemes. We observe that use of the proposed rate control schemes lead to a smoother variation in QP and hence in PSNR of the test sequences.

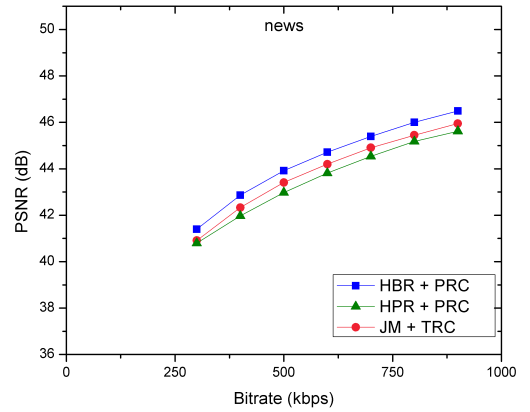


Fig. 3: RD performance for the rate control schemes

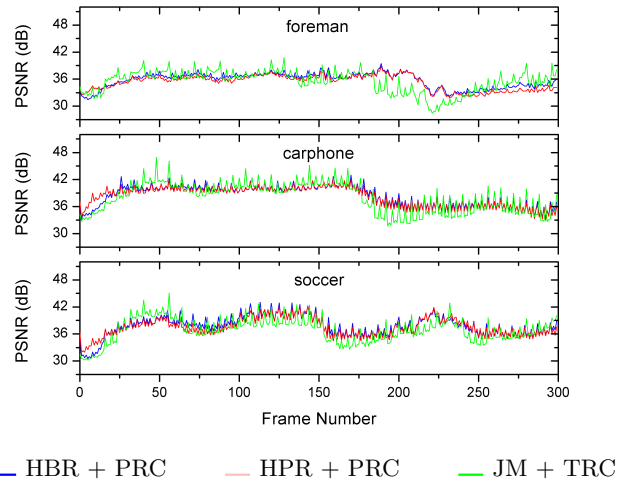


Fig. 4: PSNR curves for CIF Test Sequences

A miniature Next Generation Wireless Internet (NGWI) test-bed is also deployed for evaluating the performance of

Table 1: Standard Deviation of QP and PSNR variations for CIF Test Sequences

Sequence	Bit rate (kbps)	SD (QP)			SD (PSNR)		
		HBR + PRC	HPR + PRC	JM + TRC	HBR + PRC	HPR + PRC	JM + TRC
foreman	256	3.74	3.49	3.86	1.68	1.64	2.45
carphone	512	4.54	4.22	4.89	2.28	2.19	3.29
soccer	600	4.17	4.14	4.26	2.28	2.06	2.71

Table 2: Experimental Results on CIF Test Sequences

Sequence	Target bit rate (kbps)	Scheme	Coded bit rate (kbps)	PSNRY (dB)	PSNRY Improvement (dB)	SSIM
foreman	256	HBR + PRC	254.62	36.04	0.61	0.93
		HPR + PRC	253.51	35.60	0.17	0.92
		JM + TRC	253.83	35.43	—	0.92
carphone	512	HBR + PRC	516.93	38.39	0.45	0.95
		HPR + PRC	514.62	38.11	0.17	0.95
		JM + TRC	516.12	37.94	—	0.95
soccer	600	HBR + PRC	604.63	38.08	1.20	0.94
		HPR + PRC	601.42	37.84	0.96	0.93
		JM + TRC	602.09	36.88	—	0.93

the proposed algorithm. The snapshots of the video encoder and the mobile client while streaming a test video over the miniature testbed is shown in Fig. 5 and Fig. 6 respectively. For CBR cases, a better performance is observed for the proposed algorithm with respect to visual quality smoothness. The proposed algorithm provides consistent perceptual visual quality while satisfying the long term constant bit rate constraint. From the perceptual experience of viewers, the performance of the proposed algorithm is validated including consistent visual quality and the constant bit rate constraint.

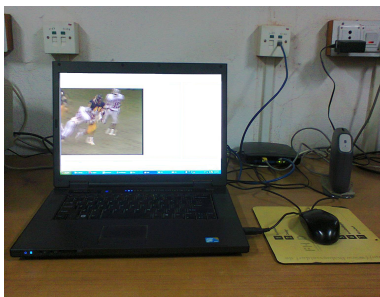


Fig. 5: Video Encoder



Fig. 6: Mobile Client

## 6. CONCLUSIONS

In this paper, a novel rate control scheme is proposed for video streaming applications constrained by a channel bandwidth. The proposed algorithm results in a consistent visual quality and satisfies the bit rate constraint. The experimental results demonstrate that our proposed rate control scheme outperforms the existing rate control scheme in JM 18.6 in terms of PSNR and consistent perceptual visual quality while maintaining the target bit rate. A video streaming system with proposed rate control is implemented over a miniature test bed. The experimental evaluation of the proposed rate control carried over the miniature NGWI test-bed validate our proposed rate control mechanism.

## 7. REFERENCES

- [1] C. Kan, Y. Bo, M. Ai-Dong, and Z. Wen-Hao. Rate Control for Hierarchical B-frames in H. 264/AVC. *The Journal of China Universities of Posts and Telecommunications*, 17(5):116–121, 2010.
- [2] A. Leontaris and A. M. Tourapis. Rate Control reorganization in the Joint Model (JM) reference software. In *JVT-W042, 23rd meeting, San Jose, California, USA*, pages 21–27, 2007.
- [3] M. Li, Y. Chang, F. Yang, S. Wan, S. Lin, and L. Xiong. Frame Layer Rate Control for H. 264/AVC with Hierarchical B-frames. *Signal Processing: Image Communication*, 24(3):177–199, 2009.
- [4] H. Schwarz, D. Marpe, and T. Wiegand. Hierarchical B Pictures. *Joint Video Team, Doc. JVT-P014, Poznan, Poland*, 2005.
- [5] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra. Overview of the H.264/AVC Video Coding Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7):560–576, July 2003.