

## Research Article

### Optimal Rate Control in H.264 Video Coding Based on Video Quality Metric

<sup>1</sup>R. Karthikeyan, <sup>2</sup>G. Sainarayanan and <sup>3</sup>S.N. Deepa

<sup>1</sup>Engineering Manager (R&D), Teclever Solutions Pvt. Ltd., Bangalore, India

<sup>2</sup>Senior Technical Specialist, HCL Technologies Ltd., Chennai, India

<sup>3</sup>Department of EEE, Anna University, Regional Centre, Coimbatore, India

**Abstract:** The aim of this research is to find a method for providing better visual quality across the complete video sequence in H.264 video coding standard. H.264 video coding standard with its significantly improved coding efficiency finds important applications in various digital video streaming, storage and broadcast. To achieve comparable quality across the complete video sequence with the constraints on bandwidth availability and buffer fullness, it is important to allocate more bits to frames with high complexity or a scene change and fewer bits to other less complex frames. A frame layer bit allocation scheme is proposed based on the perceptual quality metric as indicator of the frame complexity. The proposed model computes the Quality Index ratio ( $QI_r$ ) of the predicted quality index of the current frame to the average quality index of all the previous frames in the group of pictures which is used for bit allocation to the current frame along with bits computed based on buffer availability. The standard deviation of the perceptual quality indicator MOS computed for the proposed model is significantly less which means the quality of the video sequence is identical throughout the full video sequence. Thus the experiment results shows that the proposed model effectively handles the scene changes and scenes with high motion for better visual quality.

**Keywords:** AVC, frame layer bit allocation, H.264, perceptual video quality, rate control

## INTRODUCTION

The success of digital video application commercially resets with its ability to deliver constant quality video which is better for the given bandwidth and performance constraints. The high computational processor availability reduces the impact of the performance constraints, delivering better quality constantly for a given bandwidth is of most important. All current standards including H.264 uses transform, motion estimation/prediction, quantization and variable block size coding as building blocks. A rate control model which decides the quantization step size and monitors the buffer overflow and underflow conditions is another important module in the video encoding. So rate control model is a two-step process, in the first step arrive at a frame layer bit allocation and in the second step calculate the quantization step size which meets the allocated buffer constraints. Even though this encoding module is not explained in the standard and it is left open for application specific implementation, normally it is associated with a buffer model specified in the video coding standard. A leaky bucket model is normally employed in encoder to characterize the Hypothetical Reference Decoder (HRD) and its input buffer called Coded Picture Buffer (CPB) to avoid the buffer overflow and underflow in the decoder.

Therefore all the video coding standards usually recommend their own informative rate control model during the standardization some of them are the MPEG-2 Test Model Version 5 (TM5) algorithm in ISO/IEC (1993), the H.263 Test Model Near-term version 8 (TMN8) algorithm in Ribas-Corbera and Lei (1999) the MPEG-4 Verification Model version 18 (VM18) algorithm in ISO/IEC (2001) and the H.264 in ITU-T (2005) fluid flow traffic model in Joint Model (JM) referred in Li *et al.* (2003) and Leontaris and Tourapis (2007) algorithms. Since the 50% more compression for the same picture quality of H.264 is achieved partially due to the arithmetic entropy coding, the other features like sub-pixel motion prediction, multiple reference frames, variable block size motion estimation and spatial intra prediction also contributes to improve the compression efficiency, the complexity of the H.264 rate control also increased considerably. So the rate control scheme is ineffective in providing same visual quality during scene change and high motion sequences. The main reason is that the frame layer bit allocation which determines the quantization parameters does not contain the frame complexity in the reference model.

The reference model considers only the buffer status parameter for the target bit allocation for a particular frame. Even though lot of research work has

been done for the bit allocation including the frame complexity based ones, these are based on Mean Absolute Difference (MAD) and Picture Signal to Noise Ratio (PSNR) which are less correlated to perceptual quality of video. The proposed idea in this study is to arrive at a new perceptual quality metric QIr based method to compute frame complexity and use the same for calculating the frame layer bit allocation.

## LITERATURE REVIEW

Traditional rate control algorithm briefed within the JM reference software as describe in Li *et al.* (2003) primary goal is to achieve the target bit rate without taken the video quality into consideration. In this, initially a bit allocation scheme calculates a bit target for the current picture and further adapted to achieve the target buffer level. An estimate of the header bits is computed and the same is subtracted from the target bits to arrive at a texture bits target. This texture bit target is translated to Quantization Parameter (QP) value with a quadratic model. The calculated QP value is used to encode all the slices of the current picture. Apart from this a total bit rate target is used all the frames and the target bitrate bit rate is computed by modulating QP. An bit allocation algorithm in Leontaris and Tourapis (2007) was proposed to achieve an accurate rate control when B or periodic I frames are introduced in real time encoding. In this individual frame level target bits are set for pictures of all coding types. If these target frame bit rates are achieved then the entire sequence target bit rate as well is achieved. Even though the referred work in Leontaris and Tourapis (2007) taken into consideration of the P coded or B coded picture to calculate the target bit for the frame, the frame complexity is not considered within the P picture.

The frame complexity based frame layer bit allocation is explained in Jiang *et al.* (2005), Jiang *et al.* (2004) and Roodaki *et al.* (2006) which are derived from PSNR drop ratio, MAD ratio and mode decision, respectively. In Jiang *et al.* (2005) model, the PSNR drop is defined as:

$$PSNR_{Drop-j} = PSNR_{j-1} + PSNR_{skip-j}$$

And the estimated PSNR drop ratio is the ratio between  $PSNR_{Drop}$  of the current frame to the average  $PSNR_{Drop}$  of all previously coded P frames in the video sequence. The higher  $PSNR_{Drop}$  ratio means higher frame complexity and the higher bit allocation. In Jiang *et al.* (2004) defines a frame complexity measure MAD ratio which is the ratio between predicted MAD of the current frame to that of the previous P coded frames average MAD:

$$MAD_{RATIO}(j) = \frac{MADP_j}{\frac{1}{j-1} \sum_{i=1}^{j-1} MAD_{A_i}}$$

Both these models are using the objective quality indicator PSNR and the error indicator MAD ratio, these are generally not depicts the true quality of the user experience. So proposed a method for frame layer bit allocation using the QIr based frame complexity.

## METHODOLOGY

### Proposed perceptual quality indicator based model:

**Measure of quality indicator:** In any application user video quality is an important factor for the users Quality of Experience (QoE). Based on our work, arrived a NR metric based perceptual quality assessment as part of the encoder without much complication for the in service assessment of quality of delivery. The NR metrics for video blockiness, blur and jerkiness are calculated in accordance with ITU-P910 and perceptual Quality Indicator (QI) is calculated based on these impairments. The QI for an optimal frame rate is defined in ITU-T G.1070 as follows:

$$QI = \left( v3 - \frac{v3}{1 + \left(\frac{Brv}{v4}\right)v5} \right)$$

The constant v4 is calculated as linear combination of the impairments together. So v4 is expressed as follows:

$$v4 = a.BM + b.BL + c.JR$$

**Measure of frame complexity:** Since the reference model is based on fluid flow linear tracking model, the target bits  $T_{Buffer}$  allocated for a particular frame with constraints on target buffer level, frame rate and bitrate. In the proposed work one group of picture is considered as one video sequence and the group of picture structure is first an I frame followed by P frames. The following definitions are used:

- $T_{Buffer}$  : Target bits
- $Br$  : Bitrate
- $Fr$  : Frame rate
- $B_t$  : Target buffer level
- $B_f$  : Current buffer fullness
- $\gamma$  : Constant
- $T_r$  : Remaining bits
- $R_r$  : No. of remaining bits for encoding
- $N_r$  : No. of remaining P frames for encoding
- $T$  : Final target bits
- $\beta$  : Constant

The optimal target bits  $T_{Buffer}$  is defined as follows:

$$T_{Buffer} = \frac{Br}{Fr} + \gamma (B_t - B_f) \quad (1)$$

In Eq. (1) the  $\gamma$  is a constant and its value is 0.75. The remaining bits are calculated as:

$$T_r = \frac{Rr}{Nr} \quad (2)$$

$$T_r = \frac{Rr}{Nr} * \frac{1}{QI_r} \quad (6)$$

The final target bits T is computed as follows:

$$T = \beta * T_r + \gamma (1 - \beta) * T_{Buffer} \quad (3)$$

where,  $\beta$  as constant and its typical value is 0.5 to give equal weightage to both buffer availability and the bandwidth requirement. In Eq. (2) the  $T_r$  calculation is without considering the frame complexity. So the frame complexity is added to get better bit allocation.

The frame complexity measure  $QI_r$  is defined as ratio between the predicted QI for the current frame  $PQI_j$  to average QI for the previous P frames in the video sequence. This can be easily calculated with the following equation:

$$QI_r(j) = \frac{PQI_j}{\frac{1}{j-1}(\sum_{i=1}^{j-1} AQI_i)} \quad (4)$$

where, the predicted QI for the current frame is computed as a linear extrapolation of the previous QI:

$$PQI_j = a * AQI_{j-i} + b \quad (5)$$

In Eq. (5) a and b are QI prediction coefficients for the current frame. The average QI of the all the previous P frames as indication of the video complexity of the sequence.

**Measure of frame target bits:** The Eq. (2) which computes the remaining bits become adaptive with the  $QI_r$  is used as scaling function to include the frame complexity:

In this equation, if the predictive QI is high means that the current frame require fewer bits to code and the inverse of the  $QI_r$  will adjust the bit accordingly. If the predictive QI is low means that the current frame is more complex and require more number of bits to code the frame.

### RESULT ANALYSIS AND DISCUSSION

For the experimentation, reference JM coder is used for the H.264 video encoding. The metric calculation is implemented as part of JM reference software. The proposed quality metric calculation is implemented in C language. The video resolution is of standard definition size and encoding is set to bitrates of 512 kbps. Three different standard definitions test videos are used for the experiment. The video sequences are “mobile and calendar”, “parkrun” and “shields” all are taken from media.xiph.org website. These test vectors have various spatial and temporal complexities in nature, the tests are carried out with the standard rate control and the proposed rate control. Based on this the results are analyzed and plotted for 100 frames in each of these three video sequences.

The performance of the proposed rate control is defined as frame rate control error which is defined as:

$$E_{frc} = \frac{A_c - T}{T} \times 100 \%$$

where,  $A_c$  and  $T$  are the actual and target bits for each frame. The experimented values for both the JVT and the proposed bit allocation model are plotted and shown in Fig. 1 to 3. The values are indicating that the

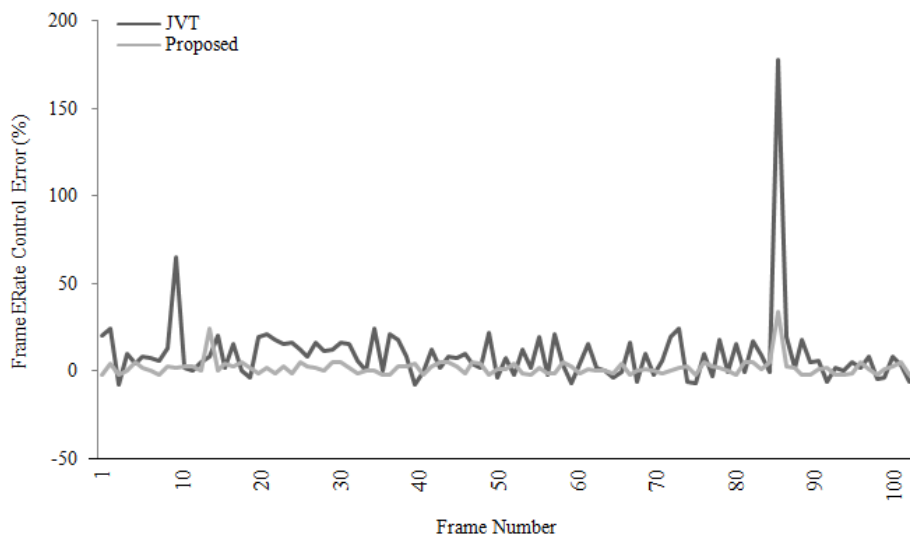


Fig. 1: Frame rate control error for the sequence “mobile and calendar”

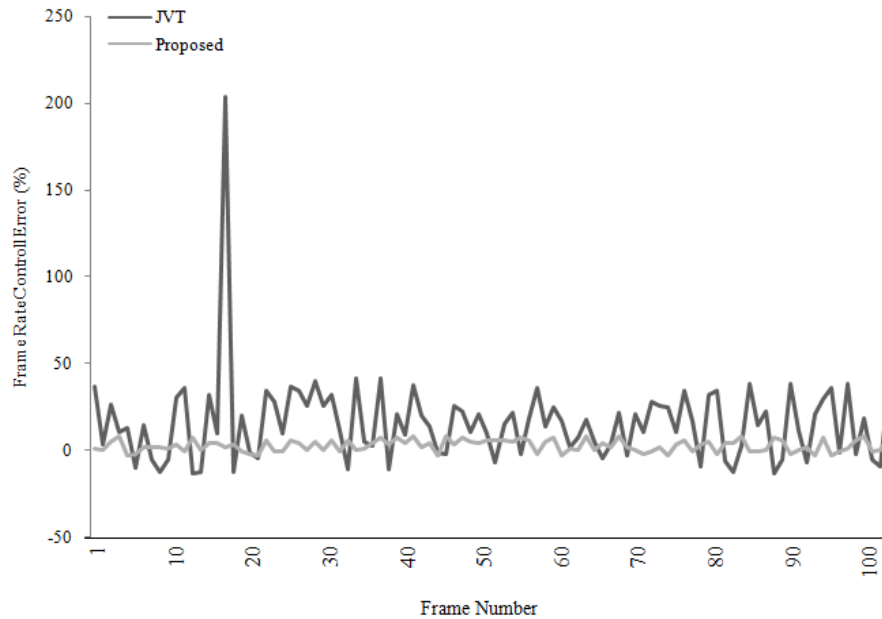


Fig. 2: Frame rate control error for the sequence “parkrun”

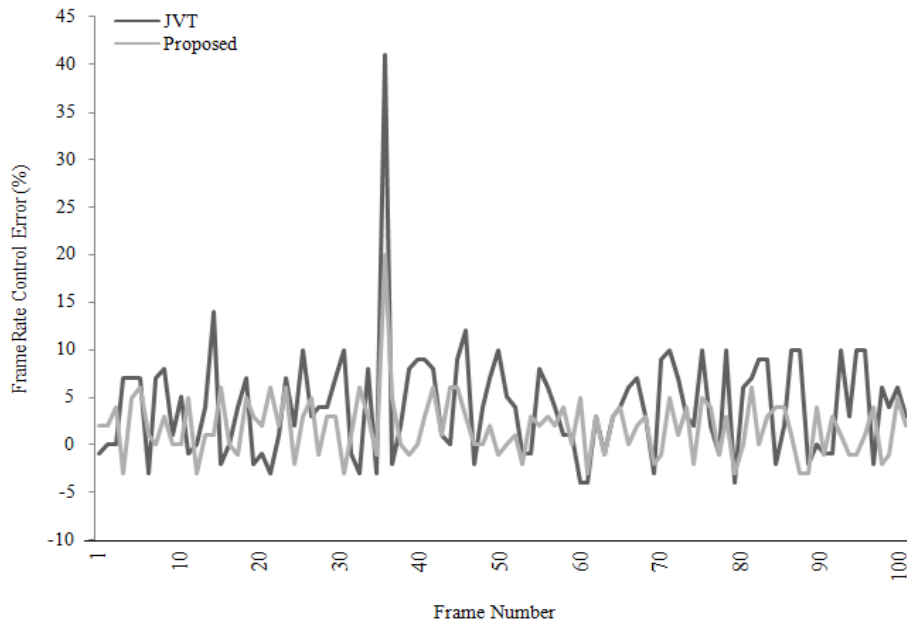


Fig. 3: Frame rate control error for the sequence “shields”

proposed model has relatively less frame rate control error than JVT G012 model.

The analysis indicates the perceptual quality variation is evident in Table 1 that the bit rate control as in JVT referred in Li *et al.* (2003) which does not include the frame complexity for the bit allocation and the perceptual quality variance is minimal for the proposed method.

The experimented values shown in Table 1 indicates that the average rate control error over each frame is minimal for the proposed method compared to

the one for JVT model with this the calculation of the target bits based on frame complexity is matching closely with the actual bits requirement for the coding. And the quality scores indicate a better quality index for the proposed model. Since the bits are allocated where it is required more which resulted in the quality improvement. Also the target bitrate is achieved more accurately compare to the JVT proposed model.

Since this perceptual quality index calculation does not include the calculation on the chroma and effect of the color component will not reflect on the quality

Table 1: Experimental results for the proposed rate control algorithm

Sequence	Average rate control error (%)		Average MOS (Qi)		Bitrate (kbps)	
	JVT	Our	JVT	Our	JVT	Our
Mobile and calendar	23.96	4.32	4.4	4.6	513.3	511.9
Parkrun	40.81	7.12	3.8	4.1	512.6	512.2
Shields	9.39	5.45	4.2	4.3	514.2	512.4

indicator which will in turn affect the frame complexity estimation to an extent. This study may be extended to include the chroma calculations for the quality indicator.

### CONCLUSION

In this research study, proposed a method of arriving frame complexity using perceptual quality indicator. The calculation of the frame layer bit allocation is based on frame complexity in addition to the buffer status based target bits calculation method. The implementation is done on the JM reference H.264 encoder and frame rate control error is calculated on the test sequences for the experimentation purpose. Experimental results show that our proposed method matched with the requested bitrate more closely and generates better quality than the reference model. Since the computation of the quality indicator is a no reference model, the computation complexity of the proposed model is less compare to the reference model.

### REFERENCES

ISO/IEC, 1993. MPEG-2 Video Test Model 5. ISO/IEC/JTC1/SC29/WG11MPEG93/457. Draft Version 1.  
 ISO/IEC, 2001. MPEG-4 Video Verification Model V18.0, Coding of Moving Pictures and Audio N3908. ISO/IEC/JTC1/SC29/WG11.

ITU-T, 2005. H.264 Advanced Video Coding for Generic Audiovisual Service. ITU-T. Retrieved from: [www.itu.int/rec/T-REC-H.264](http://www.itu.int/rec/T-REC-H.264).  
 Jiang, M., X. Yi and N. Ling, 2004. Improved frame-layer rate control for H.264 using MAD ratio. Proceeding of the IEEE International Symposium on Circuits and Systems, 3: 813-816.  
 Jiang, M., X. Yi and N. Ling, 2005. On enhancing H.264 rate control by PSNR-based frame complexity estimation. Proceeding of the IEEE International Conference on Consumer Electronics (ICCE, 2005), pp: 231-232.  
 Leontaris, A. and A.M. Tourapis, 2007. Rate control reorganization in the JM reference software. Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-W042, San Jose, California.  
 Li, Z.G., F. Pan, K.P. Lim, G.N. Feng, X. Lin and S. Rahardja, 2003. Adaptive basic unit layer rate control for JVT. Proceeding of the 7th Meeting of JVT-G012-r1, Pattaya II, Thailand.  
 Ribas-Corbera, J. and S. Lei, 1999. Rate control in DCT video coding for low-delay communications. IEEE T. Circ. Syst. Vid., 9: 172-185.  
 Roodaki, H., M.R. Hashemi and O. Fatemi, 2006. A frame layer bit allocation for H.264 based on frame complexity. Proceeding of the Canadian Conference on Electrical and Computer Engineering (CCECE '06). Ottawa, Ont, pp: 2025-2028.