Coding Complexity Prediction for H.264/AVC Rate Control

Yimin ZHOU(a) and Ling TIAN(b), Members

SUMMARY Coding complexity is a crucial parameter in rate control scheme. Traditional measures for coding complexity are based on statistic and estimation. This way may cause the imprecise coding complexity and finally bring inaccurate output bit rate more or less. To resolve this problem, we propose a hypothetical virtual coding complexity to imitate the real coding complexity. Based on the proposed coding complexity measure, a novel rate control algorithm is proposed either. Experimental results and analysis show that the proposed measure for coding complexity is effective, and our scheme outperforms the JVT-W042 solution by providing more accurate QP prediction, reducing frame skipping, and improving visual quality.

key words: rate control, H.264/AVC, coding complexity, bit allocation

1. Introduction

Rate control is employed in H.264/AVC to regulates the bit stream according to the limited network bandwidth and a predefined buffer size so that the video quality is kept as high as possible. As the latest video compression standard, H.264/AVC employs a number of innovative coding techniques to achieve significant improvements in rate distortion efficiency and coding performance [1]. However, these new features make rate control in H.264/AVC more complex. These are hot research issues in H.264/AVC.

The calculation of quantization parameter (QP) value is a key part of rate control, because the QP value may affect the coding efficiency and the visual quality. To obtain an appropriate QP value, the coding complexity of a frame, which indicates how hard to compress a video source, should be measured prior to encoding. Most of the time, the coding complexity is due to motions, shapes and textures.

Many coding complexity measurements are proposed in the literature. The mean absolute difference (MAD) of the residual signal has been widely used to estimate the coding complexity, such as in MPEG4 [2] and H.264/AVC [3]. Other approaches such as mean squared error (MSE) [4], mean absolute error (MAE) [5], sum of absolute difference (SAD) [6] for the residual block are also employed. However, the straightforward calculation of any kind of coding complexity is not economical. Since these traditional measurements are based on the statistic, the complexity of calculation is relatively high. Moreover, in H.264/AVC video coding, the mentioned coding complexities, like MAD or MSE, of the current frame can only be obtained after rate distortion Optimization (RDO), but an unique QP is required before the manipulation of RDO. Hence the current coding complexity can only be estimated in existing H.264/AVC rate control algorithms. Furthermore, the estimated coding complexity of the current frame is brought into rate-distortion (R-D) models to derive the QP value. The accumulated errors of the complexity estimation and the model adaptation may result in a suboptimal QP.

To reduce the imprecision of QP estimation caused by inaccurate coding complexity measures, we propose a hypothetical virtual coding complexity to simulate the real coding complexity in this paper. The novelty is that our scheme does not need a real statistic like mentioned above for coding complexity. The coding complexity is a symbol in the deductions. Moreover, to obtain the current frame’s coding complexity, we adopt an innovative and effective prediction approach.

The remainder of this paper is organized as follows. Section 2 describes the coding complexity prediction method. Section 3 introduces our rate control algorithm, including bit allocation strategy, QP calculation for intra- and inter-frames, and the model update approach. Section 4 presents the experimental results, and followed by conclusions in Sect. 5.

2. Coding Complexity Prediction

In our previous work [7], we proposed a simple but effective exponential R-D model, which do very well performance in H.264/AVC under JM8.6. In this section, we rewrite the model and extract the coding complexity from the R-D model. Finally we adopt a new approach to estimate the coding complexity.

2.1 R-D Model

Without the consideration of the coding complexity, we suppose the relationship between ln(bitrate) and quantization parameter is linear. And the relationship between lnR and Q can be expressed by:

\[ \ln R = -a \cdot Q + b \] (1)
where $a$ and $b$ are model parameters, the negative $a$ stands for the model slope and $b$ indicates the intercept of the linear model. The Eq. (1) can be reorganized by:

$$R = e^{-aQ+b} = e^b \cdot e^{-aQ}$$

(2)

To investigate the relationship between the parameters in Eq. (2), further experiments have been done. In Table 1 we list the experimental results on various test sequences. In order to compare the accuracy of the R-D model, the correlation coefficients Co. is adopted to display how close a model is to the actual data. The results reveal that the correlation coefficient of our model for both intra- and inter-frames are almost high up to 1. This indicates the R-D model effectively simulate the real dates. It also can be seen from Table 1 that the model slopes $a$ for each sequence are almost similar. Such results make sense the proposed model is robust.

### 2.2 Extraction of Coding Complexity

Different QP brings different bit rate for certain frame, meanwhile consecutive frames with different coding complexities output different bit rates under a fixed QP value. To simplify the deduction, we assume that the coding complexity is independent of QP. Although such an assumption may not accurate, it is feasible theoretically at least.

Let’s define $C$ be coding complexity, and $Q$ be quantization parameter. Since $C$ is independent of $Q$, the general relationship among the bit rate $R$, $C$ and $Q$ is represented as follows:

$$R = R(C, Q) = f(C) \cdot g(Q)$$

(3)

where $f(C)$ and $g(Q)$ express the complexity function and bitrate-QP function respectively.

According to Eq. (2) and (3), let $g(Q) = e^{-aQ}$, hence $f(C) = e^b$. It is to say $f(C) = e^b$ denotes the coding complexity. To simplify, we adopt a parameter $\varphi$ to replace $e^b$, so the R-D model (2) can be rewritten by:

$$R = \varphi \cdot e^{-aQ}$$

(4)

It can be observed that the model (4) has only two model parameters, while JVT-G012 has three parameters in the classic quadratic model. To investigate the intra- and inter-frames separately, we employ $a_i$ and $a_p$ to indicate the model parameter for them respectively.

### 2.3 Prediction of Coding Complexity

Although there exists a coding complexity function in expression (3), as a matter of fact, we do not adopt the function to really calculate the coding complexity. It just exists as a symbol in the deduction (4), and no idiotic measures for coding complexity are required, this is our novelty.

The current frame’s coding complexity can be predicted by the real coding complexities of previous coded frames. After the $i$th frame encoding has been finished, the real coding complexity of this frame can be calculated according to Eq. (4):

$$\varphi = R_i \cdot e^{aQ}$$

(5)

To obtain the coding complexity of the current frame, we adopt the harmonic mean (HM) of the previous encoded frames to predict the current one. The harmonic mean $H$ of the positive real numbers $x_1, x_2, \ldots, x_n$ is defined to be:

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^{n} 1/x_i}$$

(6)

Based on the above definitions, at the encoding time $t$, the estimated coding complexity $\tilde{C}_t$ can be derived by:

$$\varphi_t = \left(\frac{1}{n} \cdot \sum_{i=t-n}^{t-1} \frac{1}{\varphi_i}\right)^{-1}$$

(7)

where $n$ is the window size.

In the situation demanding rates and ratios, the harmonic mean provides the truest average. In other words, the harmonic mean of previous coded frame’s coding complexity can reflect the change tendency of the current frame’s coding complexity. Since we require a precise and robust coding complexity of the current frame, the adoption of harmonic mean is reasonable. This can be further demonstrated by our following experiments.

Figure 1 shows the accuracy of the prediction error in sequence News under two different QP values. It can be observed that the prediction error are in the range of [-2%, 4%]. Obviously, this is a comparative low and acceptable error range. Those indicate that the proposed coding complexity prediction and the HM approach are effective.
3. Rate Control Algorithm

With the new prediction measure for coding complexity proposed above, the improved rate control process is developed in this section. Let $CBF$ denote the current buffer fullness, $TBL$ indicate the target buffer level, and $Ratio_t$ stand for the ratio between the bits for intra-frame and the average bits.

3.1 Initiation

Firstly we set the buffer situation: the buffer size $BS$ is set to the half of the bit rate; the under flow level $UFL$ and the over flow level $OFL$ are $1/8$ and $7/8$ of $BS$ respectively; the current buffer fullness $CBF_0$ is set to the under flow level $UFL$ at the beginning of encoding. And we set $TBL$ same as JVT-G012.

Then Intra period $\tau$ is initiated. If $\tau = 0$, which means inter-only encoding, we set $\tau = FrameToBeEncoded$.

Finally the model parameters empirically set as: $a_I = 0.10$, $a_P = 0.15$.

3.2 Bits Allocation

We adopt the average bit allocation approach in our algorithm. The bit allocation procedure is:

When Intra period $\tau = 1$, which indicates intra-only encoding, we allocate the bits to every frame averagely:

$$AB_t = \frac{BitRate}{FrameRate} \cdot \frac{Ratio_t}{Ratio_t + \tau - 1}$$

For inter frames, we allocate bits the same as JVT-G012, which is determined based on the target buffer level, the frame rate, the available channel bandwidth and the actual buffer occupancy [3].

3.3 QP Values Calculation

The QP calculation is a crucial part of rate control, which is determined by the target bits, the current buffer fullness and the target buffer level. In this section, we investigate the QP calculation steps for intra- and inter-frames separately.

We estimate the coding complexity with the Eq. (7). Note, the coding complexity for intra- and inter-frames are only can be computed by the previous encoded intra- or inter-frames separately.

So the QP value can be calculated by model (4).

For intra-frame, the QP is: $Q_I = \left\lfloor \frac{\ln G_I(AB)}{a_I} + 0.5 \right\rfloor$.

For inter-frame, the QP is: $Q_I = \left\lfloor \frac{\ln G_I(AB)}{a_P} + 0.5 \right\rfloor$.

Finally, $Q_t$ is bounded in $\{Q_{t-1} - 2, Q_{t-1} + 2\}$.

3.4 Update

After a frame’s encoding is finished, the model should be updated quickly to adapt the changes of the coding complexity. An appropriate adaptation may reduce the misadjustment of the target bit rate.

To obtain smoother reconstruction visual quality, when the last frame of a GOP has been encoded, we update $Ratio_t$:

$$Ratio_t = \frac{R_{t+1-\tau} \cdot (\tau - 1)}{\sum_{i=0}^{\tau-2} R_{t-i}} \cdot e^{\frac{a_I}{\tau} \left(1 - \frac{1}{\tau} - \frac{\gamma^2}{\tau^2} \cdot R_t\right)}$$

For linear model (2), we update the model parameter a for intra- and inter-frame respectively by linear regression, and the window size is 10.

4. Experimental Evaluation

Numerous experiments have been implemented to evaluate the performance of the proposed scheme. The test software is JM13.2 [8]. To demonstrate the performance of the proposed algorithm, it is compared with JVT-W042 [9], which reorganized the code structure of JVT-G012. All parameters are configured for both two algorithms to be equivalent. “RCUpdateMode” is set to 3. The number of frames to be encoded is 150 and the frame rate is set to 15 frames per second. The RDO is enabled and CABAC is adopted for entropy coding.

Table 2 shows the performance comparison between two algorithms. The $TBR$ and $ABR$ denote the target bit rate and the actual bit rate respectively. The $Init.QP$ is the initial
QP value for the testing sequence. It is obvious that the actual bit rate of the proposed algorithm is closer to the target bit rate than that of JVT-W042. One can see from Table 2 that, compared with JVT-W042, our algorithm significantly reduces frame skipping and gains higher average PSNRs.

It can be further illustrated by Fig. 2. The experimental results of the sequence City at 48 kbps. From Fig. 2(a), one can see that our algorithm can control the buffer better, and the buffer curve is more stable and robust throughout the whole sequence than that of JVT-W042. And it is obvious that JVT-W042 buffer is underflow sometimes in the encoding process. Figure 2(b) reveals the average PSNR of the luminance component for the same sequence. It can be observed that our rate control algorithm achieves better PSNR and the PSNR variations of the proposed algorithm are also smaller than those of JVT-W042.

The performance comparison of Hall at 96 kbps is shown in Fig. 3. JVT-W042 still has buffer underflow frequently while the proposed algorithm is more smooth. In Fig. 3(b), it is noted that for the proposed scheme, the PSNR of nearly throughout the sequence are higher than those of JVT-W042. It represents that the proposed algorithm has effectively improved the subjective picture quality.

5. Conclusion

In this paper, we introduce an efficient coding complexity prediction for H.264/AVC rate control. The proposed hypothetical virtual coding complexity is employed to simulate the real coding complexity which is used in the R-D model to derive the QP value. Experimental results reveal that our proposed coding complexity prediction can accurately measure the real complexity for both intra- and inter-frames. In addition, the combination of our bit allocation approach and model update method improves the overall rate control performance. Compared with JVT-W042, our algorithm successfully improves the average PSNRs, precisely controls the buffer fullness and achieves accurate rate regulation.

References

[1] T. Wiegand, G.J. Sullivan, G. Bjntegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” IEEE Trans. Circuits Syst. Video Technol., vol.13, no.7, pp.560–576, July 2003.
[2] H.-J. Lee, T. Chiang, and Y.-Q. Zhang, “Scalable rate control for MPEG-4 video,” IEEE Trans. Circuits Syst. Video Technol., vol.10, no.6, pp.878–894, Sept. 2000.
[3] Z.G. Li, F. Pan, K.P. Lim, and G.N. Feng, “Adaptive basic unit layer rate control for JVT,” Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-G012, 7th meeting, pp.7–14, Pattaya, Thailand, 2003.
[4] J.C. Dagher, A. Bilgin, and M.W. Marcellin, “Resource-constrained rate control for Motion JPEG2000,” IEEE Trans. Image Process., vol.12, no.12, pp.1522–1529, Dec. 2003.
[5] I.M. Pao and M.T. Sun, “Modeling DCT coefficients for fast video encoding,” IEEE Trans. Circuits Syst. Video Technol., vol.9, no.4, pp.608–616, June 1999.
[6] E. Akyol, D. Mukherjee, and Y. Liu, “Complexity control for real-time video coding,” IEEE Int. Conf. Image Processing, ICIP 2007, pp.I-77–I-80, 2007.
[7] Y. Zhou, Y. Sun, Z. Feng, and S. Sun, “New rate-complexity-quantization modeling and efficient rate control for H.264/AVC,” Multimedia and Expo, 2008 IEEE International Conference on June 23 2008-April 26 2008 Digital Object Identifier 10.1109/ICME.2008, pp.717–720, 2008.
[8] Joint Video Team, “Reference software JM13.2,” http://iphome.hhi.de/suehring/tml
[9] A. Leonaris and A.M. Tourapis, “Rate control reorganization in the JM Reference Software,” Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-W042, San Jose, California, April 2007.