Rate Control Algorithm of H.264 on IPP for TETRA

LIU Jian-qing¹, HOU Shu-juan², LI Hai³
Beijing Institute of Technology
Beijing, China
E-mail:¹yr.ljq@hotmail.com, ²shujuanhou@bit.edu.cn, ³haili@bit.edu.cn

Abstract—To overcome the disadvantages of output stream volatility due to the ignorance of the complexity of different frame types by IPP development platform for H.264 in TETRA, a layered rate control algorithm is proposed. According to the channel conditions and buffer state, the proposed algorithm allocates bits to Group of Picture (GOP), frame and MB, respectively. Then, the optimized RD model and allocated bits are used to determine quantization parameter (Qp). Simulation results show the variance of the number of coding bits and coding time per frame decreases respectively by 61.3% and 57.4%, while the average PSNR increases 0.2dB.

Keywords-H.264 video coding; TETRA; rate control; quantization parameter

I. INTRODUCTION

Terrestrial Trunk Radio (TETRA) is a professional mobile radio (PMR) wireless communication system, which is standardized by the European Telecommunications Standards Committee (ETSI) on Radio Equipment and Systems (RES6)¹. In order to meet the demand of customers for multimedia services, how to serve a stable multimedia application has become a research hotspot in TETRA system. Due to the characteristics of high compression efficiency, adaptive coding based on video content and adaptive processing², H.264 has currently become the most widely used video coding standard.

IPP is a set of cross-platform software library launched by Intel Corporation, which provides a favorable video communication development platform for TETRA system³. Rate control algorithm is one of the key factors of video transmission quality, but the rate control algorithm provided by IPP only focus on the selection of initial quantization parameters. It ignores the different coding complexity of different types of video frames, leading to sharp jitter of the output stream. Therefore, the rate control optimization algorithm of H.264 in TETRA system based on IPP is studied in this paper.

II. RELATED RESEARCH

The author in [4] proposes a rate control algorithm which adjusts Qp based on video texture bits. The proposed algorithm can quickly adjust the output bit rate according to the video source information. However, it considers the channel conditions little. The work in [5] adjusts target allocation of frame layer using the complexity of associated frame, which can be satisfactorily applied to wireless channel, but rate control of basic unit is ignored. The algorithm proposed in [6] distributes target bit using linear RD model, achieving good control effects. But the complexity factor increases. In [7], a good balance in actual rate and image quality are obtained. But there are still shortcomings such as inaccurate of MAD prediction and large amount of calculation.

To solve the above problems, simultaneously, consider TETRA is a narrow-band communication system, this paper proposes a layered rate control algorithm based on linear RD model. This algorithm firstly distributes target bits for GOP, frame and MB layer according to the channel conditions, the buffer share and the type of frame. Secondly, this algorithm optimizes the quadratic RD model to linear model according to the actual encoding condition. On the premise of image quality not significantly reduced, it greatly reduces the coding complexity, output data and encoding time.

III. LAYERED RATE CONTROL ALGORITHM BASED ON LINEAR RD MODEL

In video communications, rate control is aimed at a higher quality video transmission at the limited bandwidth, and it is mainly achieved through two parts: the target bit allocation and Qp determination. However, the bit allocation is not applied to determine Qp in the rate control module of IPP library. The initial Qp is calculated by a fixed formula, ignoring the different coding complexity of different frames, which leads to the volatility of output stream, and the video will be skipping or stuck.

A. Bit Allocation

Bit allocation is one of the key factors that determine whether rate control is effective. In this paper, a layered bit allocation scheme of the GOP, frame and MB is proposed.

To adapt to the time-varying and band-limited wireless channel, the output bit rate of video encoder should match the bit rate provided by the transmission channel. Specifically, the ratio of the number of output video sequence goal bits to the actual number of encoded bits should be equal to the ratio of output to input bandwidth, which is also suitable for GOP layer. Therefore the current target GOP layer coded bits can be expressed as...
\[ T_o = \frac{R_o}{R_i} \times T_i \]  
\( T_i \) represents the number of target bits of current GOP, and \( T_i \) is the number of actual encoded bits. \( R_o \) and \( R_i \) are the bandwidth of input and output channel respectively. It can be inferred from (1) that the number of target bits has a linear correlation with bandwidth. Assuming that the initial buffer share is \( B_s / 8 \), where \( B_s \) is the maximum size of the encoder buffer. Thus, considering the changes of buffer size, the target GOP layer coded bits may be modified to
\[ T_i = \frac{u}{F_s} \times N_{\text{GOP}} - (\frac{B_s}{8} - B_c) \]  
where \( T_i \) stands for the number of target bits allocated for current GOP, \( u \) is available bandwidth of channel, \( F_s \) represents the pre-determined frame-rate, \( N_{\text{GOP}} \) is the number of frames in the current GOP, \( B_s \) represents the maximum size of the encoder buffer, and \( B_c \) is the current buffer share.

Target bit allocation for frame layer is implemented after the GOP layer is done. The idea of average bit weight in [7] is borrowed in frame layer. And in the general cases, the weight ratio of frame I and frame P is 1:0.6. Taking into account the IPPP format of video sequence and over-allocated frame I bits, the target bits of frame layer are allocated by
\[ \begin{align*}
B_i &= \frac{T_i}{N_i + 0.6 \times N_p} \\
B_p &= \frac{N_p}{N_i + N_p} \times B_i = \frac{N_p}{N_i + N_p} \times \frac{T_i}{N_i + 0.6 \times N_p}
\end{align*} \]  
where \( B_i \) and \( B_p \) stand for the number of target bits of frame I and frame P, respectively. \( N_i \) and \( N_p \) are the number of frame I and frame P in current GOP.

When obtaining the target bits for frame layer, it comes to the MB layer. Considering the header information of the video frame and the number of MB, the allocated bits for the first MB \( mb_0 \) is
\[ mb_i = \frac{B_i}{N_{\text{MB}}} - h \quad i = I, P, B \]  
where \( B(i = I, P, B) \) is expressed as the number of target bits of corresponding frame, \( h \) represents the average header information of the encoded MBs, \( N_{\text{MB}} \) is the total number of MB in current frame. Since the first two MBs own the same \( Qp \), the number of target bits for second MB is approximately equal to \( mb_0 \). The number of target bits is allocated for the rest MBs are
\[ mb_j = \frac{B_j - mb_j - mb_j}{N_{\text{MB}} - j} - h \quad 2 \leq j < N_{\text{MB}} \]  
where \( j \) stands for the serial number of current MB, \( B_j \) is the number of remaining bits, \( mb_j \) represents the number of target bits of MB. \( B_j \) is updated after finishing the encoding of current MB, that is
\[ B_j = B_j - mb_j' \]  
where \( mb_j' \) represents the number of bits after encoding.

\[ B = \text{MAD} \times \left( \frac{c_1}{Qp} + \frac{c_2}{Qp^2} \right) \]  
where \( B \) is the number of target bits of current frame; \( \text{MAD} \) is the mean absolute distortion of the frame, which can be estimated from the previous frame; \( Qp \) is the quantization parameter of current MB; \( c_i \) and \( c_2 \) are the adjustment parameters. The relationship between \( Qp \) and \( B \) is shown in Fig. 1.

![Figure 1. The relationship between Qp and B](image-url)
\[ Qp = \frac{coef \times MAD}{B} \]  \hspace{1cm} (8)

where \( coef \) is the updated parameter after each MB is encoded according to the complexity of neighboring MB. Obviously, the computational complexity is greatly reduced, but the "egg paradox" still exists. To solve this problem, \( MAD \) or a linear coefficient \( k = coef \times MAD \) is needed.

Since \( k \) of real-time captured sequence is almost unchanged, \( k \) is utilized to calculate \( Qp \), instead of predicting \( MAD \). Assuming the initial values are \( k_0 = coef_0 \times MAD_0 \) and \( k_1 = coef_1 \times MAD_1 \). Where \( MAD_0 \) and \( MAD_1 \), \( coef_0 \) and \( coef_1 \) are actual errors and parameters of the first two MBs of current frame, respectively. For the rest of the MB, the initial value of \( k \) is obtained on the basis of left and top MB, that is

\[ k = \frac{k_l + k_t}{2} \]  \hspace{1cm} (9)

To increase precision, \( k \) should be updated after encoding current MB, that is

\[ k_j = coef_j \times MAD_j \]  \hspace{1cm} (10)

On the premise of \( k \) and target bits of MB are known, \( Qp \) can be calculated by linear RD model. The initial \( Qp_0 \) is needed before calculating the other \( Qp \). \( Qp_0 \) is

\[ Qp_0 = \frac{coef_0 \times MAD_{prev}}{mb_0} \]  \hspace{1cm} (11)

where \( MAD_{prev} \) is the average \( MAD \) of all of the previous encoded MBs, \( coef_0 \) is the preset parameter on the basis of prior information. For the remaining MBs, the target bits can be obtained by (5), and \( k \) by (9). Therefore, \( Qp_j \) in current MB is available by

\[ Qp_j = \frac{k_j \times mb_j}{MB_j} \]  \hspace{1cm} (12)

To maintain the continuity of the video quality, the difference of \( Qp \) in two consecutive coding unit does not exceed 3\(^{[7]}\). After obtaining the current \( Qp_j \), through (12), it will be modified by

\[ Qp_j = \begin{cases} Qp_{j-1} + 3 & Qp_j > Qp_{j-1} + 3 \\ Qp_j & Qp_{j-1} + 3 \geq Qp_j \geq Qp_{j-1} - 3 \\ Qp_{j-1} - 3 & Qp_j < Qp_{j-1} - 3 \end{cases} \]  \hspace{1cm} (13)

When completing encoding MB, both \( coef \) and \( B \) should be updated. \( mb_j \) and \( MAD_j \) are available after encoding the MB, thus, \( coef_j \) are updated by

\[ coef_j = \frac{Qp_j \times mb_j}{MAD_j} \]  \hspace{1cm} 0 < j < N_{MB} \hspace{1cm} (14)

In summary, the proposed rate control algorithm is as follows:

1) Allocate target bits for GOP, frame and MB layer by (2), (3) and (4), respectively;
2) Determine the initial \( Qp \) by (11), and calculate the value of \( k \) for the first two encoded MBs;
3) Select \( Qp \) according to linear model RD by (12) and encode the rest of MBs;
4) Undertake RDO and update parameters in MB by (6), (10) and (14);
5) Return to step 3) until all the MB is encoded.

IV. SIMULATION RESULTS

The simulation is carried out on the platform of Intel IPP7.1, and processor is Intel Core i3. Video sequence is real-time in QCIF size and YUV 4:2:0 format. The frame-rate is 10 frame/s and coding structure is IPPP. Each frame is divided into basic units of 360 bytes. Maximum number of reference frames is 4, the intra-coded refresh cycle is 10.

Fig. 2 shows the comparison of the PSNR of the video sequence between the proposed algorithm and the IPP library. In Fig. 2, the fluctuation of the PSNR of the proposed algorithm is less than the IPP library. The average PSNR of the sequence has increased by about 0.2 dB compared with IPP library, while the variance of PSNR has decreased by 52.8%, therefore the video quality of the continuous frames will not change too much, and the proposed algorithm has achieved a good PNSR control.

![Figure 2. Comparison of PSNR](image)
Fig. 4 shows the comparison of encoding time per frame between the proposed algorithm and the IPP library. As shown in Fig. 4, the average time of encoding time per frame has decreased by about 0.2ms, so the entire encoding time is shortened. What’s more, the variance of actual encoding time has decreased 57.4%, which has alleviated the buffer stuck, so the video stream can output smoothly.

Simulation results show that the average PSNR has increased by 0.2dB, and both the number of the encoded bits and encoding time have been reduced, at the same time, the variance have decreased by 61.3% and 57.4% respectively. Therefore, the proposed rate control algorithm can get better rate control effect.

V. CONCLUSION

To adapt the rate control of H.264 in IPP to the narrow-band TETRA system, a layered rate control algorithm based on linear RD model is proposed. Firstly, target bits are allocated according to the channel conditions and buffer state for different layers. Then the quadratic RD model which determines $Q_p$ is optimized to a linear one. Simulation results show that the average PSNR has increased about 0.2dB. Meanwhile, the variance of PSNR, the number of coding bits per frame and encoding time per frame decreases by 52.8%, 61.3% and 57.4%, respectively. So the proposed algorithm has achieved good control effect.

REFERENCES

[1] ETSI EN 300 392-2 v3.4.1. Terrestrial Trunked Radio (TETRA); Voice plus Data (V+D); Part 2: Air Interface (AI)[S]. August 2010.
[2] ITU-T Rec. H.264 advanced video coding for generic audiovisual services[S]. 2011.
[3] Information on http://software.intel.com/zh-cn/node/256490. Jun, 2012.
[4] Yihhan Tan, Chunhao Yeo, Zhengguo Li. Single-pass rate control with texture and non-texture rate-distortion models[J]. IEEE Transactions on Circuits and System for Video Technology, 2012, 22(8): 1236-1245.
[5] Myoung-Jin Kim, Min-Ched Hong. Adaptive rate control in frame-layer for real-time H.264/AVC[C]. Proceedings of APCC, Tokyo, 2008: 1-5.
[6] Zhaohua Lin. A study on the low-complexity rate control algorithm for H.264/AVC[C]. Fourth International Conference on Computational and Information Sciences, Chongqing, 2012: 837-847.
[7] JVT, ITU-T JVT-G012. Adaptive Basic Unit Layer Rate Control for JVT[S]. 2003.
[8] Guang Y. Zhang, et al.. A simplified rate control algorithm for H.264/SVC[C]. SPIE Visual Information Processing, California, 2011, 8056: 1-7.