A LAGRANGIAN OPTIMIZED RATE CONTROL ALGORITHM FOR THE H.264/AVC ENCODER

M. Mahdi Ghandi and Mohammed Ghanbari

Department of Electronic Systems Engineering, University of Essex
Wivenhoe Park, Colchester CO4 3SQ, UK, Emails: {Mahdi, Ghan}@essex.ac.uk

ABSTRACT

The H.264/AVC video coding standard has been recently proposed by the Joint Video Team of ITU-T and MPEG experts. The AVC encoder achieves video bitstreams at the same quality as the previous standard coders, while requiring typically 50% of the bit rate. This paper proposes a Lagrangian optimized rate control algorithm for the H.264/AVC video encoder. It controls the bit rate by adjusting the Lagrangian multiplier for every picture and specifying the quantizer parameter for every Macroblock. In the proposed method, the accuracy of desired bit rate is tunable by allocating more searching time to the encoder. Experimental results show that the proposed method gives a quality improvement of about 0.5dB when compared to the rate control method utilized in the JM7.4 test model, with an acceptable encoder complexity.

1. INTRODUCTION

The H.264/AVC video coding standard [1] has recently been proposed by the Joint Video Team (JVT) of ITU-T and MPEG experts. It is based on a motion compensated hybrid DCT coding method similar to H.263 [2] and MPEG4 [3] standards, but it achieves a significant improvement in the Rate-Distortion (R-D) efficiency relative to the existing standards [4, 5]. The H.264 standard, similar to the previous ones, specifies only the bitstream syntax and the decoding process while the encoding process is not specified in the standard. This makes the encoders flexible in implementation and different in rate-distortion efficiency. There are a number of parameters that can be selected for every Macroblock (MB) which determine the overall rate and the distortion of the coded bitstreams. Rate control is to select these parameters to achieve a determined target bit. However every MB coding parameters should be selected very carefully to achieve more R-D efficient encodings.

In [5], the Lagrangian optimized mode decision and Motion Estimation (ME) method adapted to the AVC test model software are discussed. In this method, the Lagrangian multiplier (λ) is first calculated with an empirical formula using the selected Quantizer Parameter (QP) for every MB:

$$\lambda = 0.85 \times 2^{\frac{QP-12}{3}} \tag{1}$$

During the encoding process, all coding modes of every MB are examined and the resulting rates (R) and the distortions (D) are calculated. The mode that has the minimum J is selected as the optimum mode for every MB:

$$J = D + \lambda \times R \tag{2}$$

In this method, one fixed QP should first be selected for every MB and then the optimum MB mode is defined. Thus, the number of bits of every picture is not controlled.

In a JVT contribution [6], a rate control based on an adaptive linear model is proposed. In this method, the distortion of every basic unit is predicted using an adaptive linear model. Every basic unit can be a frame, a slice or a MB. To control the number of bits, the QP of every basic unit is calculated using a quadratic R-D Model. Using the selected QP, all MBs of the basic unit are coded in the modes selected by R-D optimization of [5]. This method has been adapted and implemented in the JM7.4 test model software of JVT.

A Lagrangian optimized rate control algorithm for the AVC video encoder is proposed in this paper. It controls the encoded video bit rate by adjusting the λ value for every picture. The encoder examines different combinations of MB mode and QP for every MB and finds the optimum setting that meets the rate constrains. To achieve the acceptable encoding delay, the number of QPs to search is minimized. Additionally, the λ values and the QP starting points are adaptively tuned. In the proposed method, the accuracy of the achieved bit rate is adjustable by allocating more searching time to the encoder.

The paper is organized as follows. Section 2 gives the basic principle of the Lagrangian optimization method used in the proposed rate control. Details of the proposed method are described in Section 3. The simulation results are given in Section 4 followed by the conclusions in Section 5.

2. RATE CONTROL OPTIMIZATION

In the Lagrangian optimization method, small λ values correspond to high rates and PSNRs; the converse is true for large λ values [7, 8]. It implies that changes in the λ value influence the rate and PSNR of the resulting coded picture. This is the basic assumption for the proposed rate control method. The coder operates by adjusting the λ value to control the bit rate (or PSNR) of every picture.

Figure 1 shows the R-D curves of the second frame of the Foreman test sequence. The picture is coded using three different QPs and a range of λ values. The diagrams show how the λ value affects the picture rate and PSNR. However, in order to maintain the R-D efficiency of the coder, a proper QP value should be selected as well as an appropriate λ value. In the proposed method the QP value is adaptively selected for every MB to achieve better R-D efficiency. In addition, it has simplifications to meet an acceptable encoding speed.

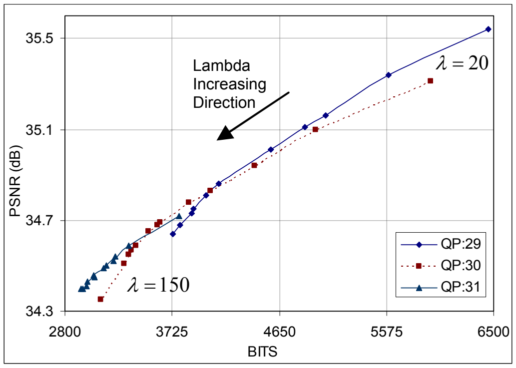

Figure 1: R-D diagrams, the 2nd picture of Foreman QCIF coded by JM7.3 encoder with different fixed QPs and a range of λ values, first picture is intra coded with QP equal to 30

3. THE PROPOSED ALGORITHM

The block diagram of the proposed rate control algorithm is illustrated in Figure 2. For encoding of a picture the λ value and the starting middle value of QP (mQP) are first selected. Subsequently, R-D optimization process selects the modes and the QPs of all MBs using the specified λ. The picture is then coded using the selected settings and the resulting rate and distortion are specified. If the resulting rate or distortion is acceptable, the encoding of the picture is completed, otherwise the picture is coded again using another λ and mQP.

The encoder delay is very sensitive to two particular points: the number of QP values that are examined during the R-D optimization and the proper mQP and λ selection. The latter one is more critical because improper selection of λ encumbers the whole process with more picture coding phases. In the following sections these issues are addressed.

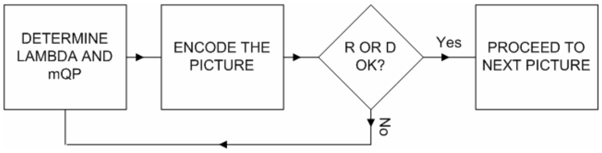

Figure 2: The picture encoding process in the proposed method

3.1. Selection of λ and mQP

For coding every picture a value for λ and a starting point for mQP should be selected. The encoder receives a QP value for the starting point as an input parameter (iQP). The first time in the video sequence that a picture of a specific type is coded (e.g. first I or P pictures), the mQP is set equal to the iQP. The value of λ is then calculated from (1).

After coding every picture, whether it is accepted for writing or not, the average value of the selected QPs is calculated and stored. This is used for selecting the next mQP for the same picture type. The latest value of λ is also stored for every picture type and then the next λ value (of the picture with the same type) is adjusted equal to this value.

If the resulting coded picture rate (or PSNR) is not acceptable, the λ value has to be adjusted. To increase the rate (or PSNR), the λ value should be decreased and vice-versa. We increase or decrease the λ value by 50% of the last λ value. The value of λ that has the nearest resulting rate below the target rate, is called $\lambda_{\min}$, and the one which has the nearest upper rate is $\lambda_{\max}$. In a particular λ adjustment stage, if both $\lambda_{\min}$ and $\lambda_{\max}$ are known, the next λ value is adjusted between $\lambda_{\min}$ and $\lambda_{\max}$ values. Note that in every λ adjustment stage, the value of mQP is set to the last average QPs.

If a combination of λ and mQP has an acceptable resulting rate, it is likely that this combination also has an acceptable rate for the next picture with the same type. Thus the number of λ adjustment stages which is very critical to the encoding speed may be high for the first picture, but it is low for the following pictures in the video sequence resulting a faster procedure. It should be noted that to achieve more accurate rate control, more precise λ adjustment is needed and the speed of the encoder would be slower. We will show in the simulation results that an accurate rate control is achievable in an acceptable encoding delay using the proposed method.

3.2. Detection of the Optimum MB Settings

To find the optimum combination of MB modes and QPs, Rate-Distortion (R-D) optimization is performed in the encoding process. For every MB, three different QPs are examined with all possible MB modes. The examined QP values are mQP, mQP+1 and mQP-1. Since the value of λ is constant during the R-D optimization, Motion Estimation (ME) which is a time-consuming process, is performed only once for every MB.

The mQP for encoding the next MB is set to the selected QP of the current MB. The optimization process starts with mQP because it is likely to be selected as an optimum value. It will be shown that examining three values of QP is almost enough to find the optimum value. The reason is that if the difference between two consecutive MB QPs (delta-QP) is larger, the number of bits needed to send delta-QP (in the MB header) will be bigger. Thus, high delta-QPs are rarely selected as optimum values. To verify this, we have performed encoding simulations at the same parameters but with different number of examined QPs. Table 1 shows the bit rate, PSNR and the number of iterations (NI) for encoding the Foreman test sequence at three different numbers of QP search windows. From the table, it is confirmed that by increasing the QP search window only the encoding time is increased without any significant gain in the performance.

Table 1: Foreman QCIF@10Hz, 100 frames coded using the proposed method with different QP search windows

Number of Examined QPs	Bit rate (kb/s)	PSNR-Y (dB)	Number of Iterations (NI)
3	98.7	37.26	240
5	98.5	37.24	370
7	98.9	37.28	500

After encoding every picture, the encoder decides whether to accept the coded picture along with the coding parameter or code it again with an adjusted λ value. This decision is made using the resulting bits and PSNR of the coded pictures. It is possible to adjust either the rate or PSNR of pictures. If a Constant Bit-Rate (CBR) bitstream is needed, the rate of the pictures should be within a range specified by a given target rate and threshold. The same decision is made for every picture PSNR when the overall PSNR of the video bitstream should be controlled. Furthermore, the accuracy of the rate-control which directly influences the encoder delay is tuned in this stage.

4. SIMULATION RESULTS

We have implemented our proposed rate control scheme on the JM7.3 AVC test model software. As a reference for the comparisons, the rate control of JM7.4 test model software has been selected. In all simulation tests, the encoding parameters have been set exactly the same for both JM7.4 and the proposed encoders in order to have fair comparisons. The first pictures are intra coded and the remaining pictures are P-TYPE. For all tests, the CABAC mode is enabled and ME search window is set to 16. All other parameters such as de-blocking filter, context initialization and file mode have been carefully selected equivalent.

The resulting R-D performances are strongly dependent on the temporal bit allocation. On the other hand, if different numbers of bits are allocated to the first I pictures, different performances are achieved. The JM7.4 rate control is not able to tune the first picture rates and it just uses a fixed QP for the first (I and P) frames. Thus, for all tests first the sequences have been encoded using the JM7.4 encoder. Secondly, the proposed encoder has been forced to adjust the first picture rate exactly to the same amount generated by the JM7.4 encoder. Thus, we can assure that all comparisons are absolutely fair.

Table 2 shows the resulting bit rates, PSNRs and NIs for the Carphone test sequence in five different bit rates. Note that in our implementation we have not optimized the source code highest speed and the demonstrated NIs just give an approximate estimation of the algorithm delay. From Table 2 it can be seen that to have the same bit rate accuracy, our implementation of the proposed method, needs almost twice the encoding time. However the output quality is about 0.5 dB better than the JM7.4 encoder.

Table 2: The bit rate, PSNR and NIs of JM.7.4 encoder and the proposed method for Carphone test sequence

Target Bit rate (Kb/s)	JM7.4			Proposed Method		
	Bit rate (kb/s)	PSNR Y(dB)	NI	Bit rate (kb/s)	PSNR Y(dB)	NI
100	98.8	38.63	130	99.8	39.10	240
80	79.7	37.41	120	78.9	37.79	250
60	59.4	35.86	120	58.8	36.28	270
40	39.5	33.84	110	39.0	34.18	250
20	19.8	30.58	100	19.5	31.00	270

Figures 3, 4 and 5 show the R-D performance of JM7.4 and the proposed encoders for three different test sequences. From the figures it can be seen that our proposed method has better resulting PSNR compared to the JM7.4 encoder.

Figure 6 shows the first 33 pictures rates of the Foreman test sequence in two different rate accuracies. It illustrates that when more encoding time is allocated to the encoder, a more accurate rate regulation is achieved.

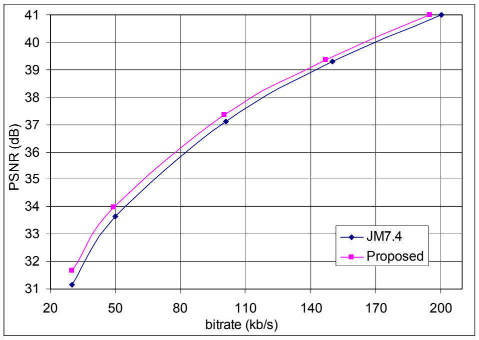

Figure 3: The R-D performance of JM7.4 and proposed rate control, Foreman QCIF@10Hz, 100 frames coded

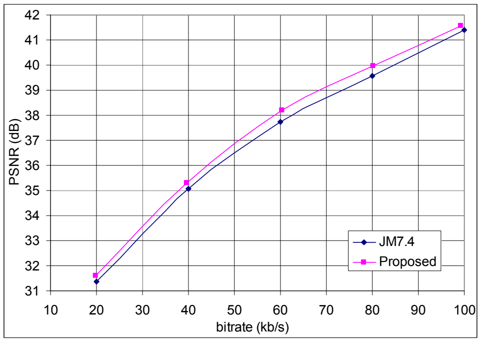

Figure 4: The R-D performance of JM7.4 and proposed rate control, Silent QCIF@10Hz, 100 frames coded

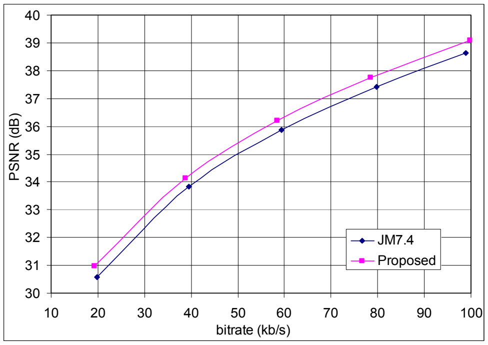

Figure 5: The R-D performance of JM7.4 and proposed rate control, Carphone QCIF@10Hz, 100 frames coded

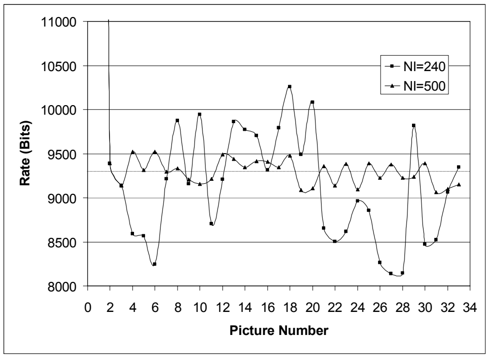

Figure 6: Picture by picture PSNR of the Foreman test sequence with the proposed method in two different rate accuracy settings

5. CONCLUSIONS

A Lagrangian optimized rate control for the H.264/AVC video encoder has been proposed in this paper. It has the ability to control every picture rate by adjusting the λ value of the R-D optimization process. The QP for every macroblock is adjusted separately and the middle point QP values for search in the R-D process are adaptively tuned. Simulation results show that our proposed method generates bitstreams with higher PSNR levels compared to the JM7.4 rate control. Furthermore, the accuracy of achieving bit rate is tunable by allocating more searching time to the encoder.

6. ACKNOWLEDGEMENTS

This work is supported by the Engineering and Physical Sciences Research Council (EPSRC) of the UK. The authors would also like to thank Dr. Ernesto Luiz Andrade Neto for his valuable contributions.

7. REFERENCES

[1] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, “Final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC),” JVT-G050, 2003

[2] International Telecommunication Union (ITU), “Recommendation H.263: video coding for low bit-rate communications,” Feb 1998

[3] ISO/IEC MPEG, “Coding of audio-visual objects – Part 2: Visual,” ISO/IEC 14 496-2 (MPEG-4 Visual Version 1), Apr. 1999

[4] Thomas Wiegand et al, “Overview of the H.264/AVC video coding standard,” IEEE Trans. on circuits and systems for video technology, Vol. 13, No. 7, pp. 560-576, July 2003

[5] Thomas Wiegand et al, “Rate-Constrained Coder Control and Comparison of Video Coding Standards,” IEEE Trans. on circuits and systems for video technology, Vol. 13, No. 7, pp. 688-703, July 2003

[6] Zhengguo Li et al, “Adaptive Basic Unit Layer Rate Control for JVT,” JVT of ISO/IEC MPEG & ITU-T VCEG, Pattaya, Thailand, 7-14 March, 2003

[7] Gary J. Sullivan and Thomas Wiegand, “Rate-Distortion Optimization for Video Compression,” IEEE Signal Processing Magazine, pp. 74-90, November, 1998

[8] Antonio Ortega and Kannan Ramchandran, “Rate-Distortion Methods for Image and Video Compression,” IEEE Signal Processing Magazine, pp. 23-50, November, 1998