Performance Analysis of Rate Control Scheme in H.264 Video Coding

T. Bernatin^{1*}, G. Sundari², B. Gayethri³ and Swarna Ramesh³

ISSN (Print): 0974-6846

ISSN (Online): 0974-5645

¹Sathyabama University, Chennai - 600119, Tamil Nadu, India; bernatin@gmail.com ²ECE Department, Sathyabama University, Chennai - 600119, Tamil Nadu, India; sundarig2014@gmail.com ³Robotics Division, CVRDE, Avadi, Chennai – 600054, Tamil Nadu, India; gayathri.b@cvrde.drdo.in, swarnaramesh@cvrde.drdo.in

Abstract

Background/Objectives: Digital Video streams represent huge amount of data at high definition resolution. The video in its original size needs more time and large storage space, which necessitates video compression. To surpass the challenges in video compression those appear in terms of latency, the rate control scheme is the best method. To provide an efficient video coding based on H.264/AVC at substantially low bit rate. This ensures higher performance in terms of compression ratio, lower complexity and video reconstruction. Methods/Statistical Analysis: The cost estimation technique is complex and takes considerable time for computation. In an attempt to make it simple, rate distortion baseline profile encoder is parallelized, which makes the R-D cost calculation feasible. Findings: In this work, the performance analysis of Rate control scheme with optimized Quantization parameter value is carried out. This rate-control scheme in the macro-block layer of H.264 baseline encoder with bit-stream calculation and distortion evaluate can potentially contribute to efficient video transcoding systems. In this proposed work, the rate control for the prediction frame done after encoding the I-frames. Here, QP estimation happens between the encoder interface and the user interface. The rate distortion model is attached to the P-frame to perform the rate control for prediction and encoding the P-frames by estimating the QP values, thus making the algorithm less complex. This method achieved the better quality with the optimized quantization parameter. Applications/Improvements: The Quantization parameter is varied to regulate the coded bit streams to achieve good perceptual quality that is suitable for surveillance applications.

Keywords: Bit Rate, Cost Estimation, Compression Ratio, Rate Control Scheme, Rate Distortion Modeling

1. Introduction

H.264/AVC is block-oriented video compensated based codec standard by ITU-T video coding expert team and ISO/IEC moving picture expert team. Higher coding efficiency as compared to older standards, flawless incorporation of video codec standard for different architecture, error resilience tools are some of the key features of interest. High-end coding tools provide superior rate distortion efficiency and improved video compression performance. In video coding, the rate distortion plays an important role for many functions such as rate-control and mode-decision. Earlier research works have shown significant contributions in this domain. Lagrangian Rate

Distortion Optimization based on the Laplace distribution of transformed residuals is capable to adaptively optimize the input sequences which results in improved overall coding efficiency. Initially, based on Laplace distribution of transformed residuals, accurate rate and distortion models are developed¹. A Rate-Distortion Optimization (RDO) accelerator using transform-domain based on fast sum of squared difference and VLC-based rate estimation algorithm is used to minimize the complexity of RDO technique, a new FSSD algorithm with an iterative lookup table quantization process². It is also being suggested to have modified rate- control³ for both intra and inter -modes, wherein an adaptive bit-rate calculation is proposed to increase the estimate accuracy. The

^{*}Author for correspondence

RDO algorithm calculates the rate-distortion cost(R-D) of each coding blocks with respect to each macro blocks and chooses the coding mode with the lowest rate distortion cost. Though it is a time taking process, it is a reliable technique for rate control scheme.

2. Video Sequence for Compression

The perception system used for operator's visual feedback is housed in the teleoperated Unmanned Ground Vehicle (UGV). It uses a combination of cameras (CCD, Thermal Imager) for efficient visual feedback. The input to the encoder is the acquired raw video from the CCD camera mounted in the teleoperated vehicle, the bit stream is transmitted over the digital wireless communications network for decoding and display in the Operator console resided in the base station. The output of the decoder is a reconstructed video that is displayed on the operator's console for navigation and control of unmanned vehicle as illustrated in Figure 1. The acquired video series

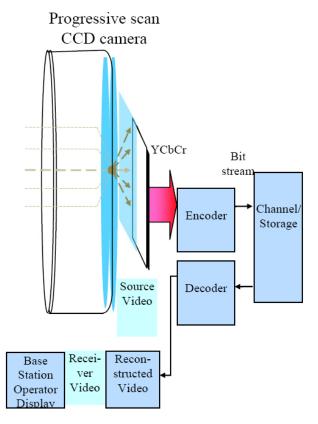


Figure 1. Block diagram of video sequence for compression.

is spliced into frames. The first frame, at the start of the video series is coded as "Intra" mode (Intra frame); the information from this frame is being used and no prediction reference exists to other preceding frame in the bit stream. Each section of a block in such an image is predicted by exploiting the spatial redundancy. For the left behind the frames of a video series, "inter" (Inter frame) coding is utilized.

The successful deployment of encoder for the application of this nature requires the insight of encoder control. The enhanced performance of H.264 compression is attributed to two key factors: (a) Motion compensation (b) De-blocking filter.

Due to temporal redundancy, the consecutive frames are typically very similar and the minor shifts in the consecutive frames are attributed to the vehicle motion. This redundancy is exploited by recording the difference or error between a frame and a previous or a next frame and is corrected by the amount of shift. These small differences referred as motion vector, when encoded gives us a fewer bit than the original pixel values and this is the prominent advantage of Motion compensation^{4,5}.

The Motion estimation technique exploits by searching the exact block in the neighboring frame. Either one frame or more than one frame can be used at a time for comparison. The latency parameter for video compression basically depends on the motion estimation time. The difference between two frames gives the motion vector which is indicated in the reference frame and can be used to predict the current frame.

The primary role of the rate distortion optimization technique is to estimate the prediction for video coding. There are basically two types of prediction techniques viz. intra prediction and inter prediction. The Intra (4X4) prediction is based on predicting each 4X4 luma samples separately. On the other hand⁶, the Intra (16X16) prediction performs prediction of the whole sample and is more suited for coding of very smooth places of a picture. In this work, 4x4 prediction has been implemented. After the prediction of the first frame called I-frame, next subsequent frames are predicted i.e. P-frames called as an Inter Frame prediction. This prediction tries to achieve a higher compression rate from the temporal redundancy between neighboring frames³. The rate-control and mode-decision are used mainly to improve the video quality in video compression. Here Rate-Distortion Optimization (RDO) optimizes the amount of distortion against the amount of data required to encode the video7. The RDO algorithm

estimates the Rate-Distortion(R-D) cost of each coding block with respect to each macro block. This algorithm chooses the coding mode with the lowest rate distortion cost8.

H.264 Encoder

In H.264, each macroblock of fixed size consists of three components, i.e. Y, C_r and C_b Y is the luma component and C_r C_h are chroma components. For 4:2:2 chroma sampling, each macroblock consists of one block of 16x16 pixels for the luminance component and two blocks of 8x8 pixels for the color components illustrated in Figure 2.

Slice is a sequence of macroblocks processed in raster scan order. Each slice is predicted using motion compensation depending on their macroblock using inter and intra prediction9, video codec with Rate Distortion model. After all the luma and chroma components are predicted either spatially or temporally, the resultant of the prediction i.e. the residual, is encoded using transform encoding process.

For this each sample is subdivided into 4X4 samples, then Integer transform is applied and entropy encoding technique is used to encode the data4. The encoded data (bit stream) is transmitted to the base station through wireless communication link for decoding the recon-

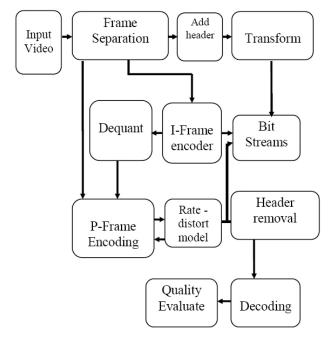


Figure 2. Block diagram for H.264/AVC.

structed and undistorted video sequence. The transform block improves the compression efficiency and also the use of 4x4 transform block size leads to significant reduction in ringing artifacts.

4. Rate Distortion Modeling

The R-D enhanced mode decision technique and motion estimation have various inter and intra prediction modes and various reference frames, is best in coding efficiency among several new coding techniques. The cost estimation technique generally takes time and it is complex so, to make it simple we parallelize rate distortion baseline profile encoder and thus the R-D cost can be calculated¹⁰. The rate distortion model is attached to the P-frame to perform the rate control for prediction and encoding the P-frames by estimating the QP values, thus making the algorithm less complex.

In the Lagrangian multiplication (λ) is calculated with an observed formula using the Quantizer Parameter (QP= 27) for every macroblock.

$$\lambda = 0.85 \times 2^{(27-12)/3}$$
 (1)

During the compression, the Rates (R) and the Distortions (D) are calculated for all coding modes of every macroblock. The mode that has the least value of *I* is selected as the best mode for every macroblock using the below formula.

$$I = D + \lambda \times R$$

In this technique, one stable QP should be chosen for every macroblock and then the best macroblock mode is defined and the bit-rate estimation R is

$$R = aN_{onz} + bE_{nzq} + cN_{onzzr} + dN_{oz} + eN_{oto}$$

where, a to e are the model parameters, Nonzer denotes the number of times non-zero zero-run-length. After modeling the bit rate for the video codec, the rate calculation can be determine classified as

$$RDcost = D + \lambda mode \times R$$

In simple function, it can be written as

$$\begin{split} J &= D + \lambda R \\ &= (S_{STD,z} + c_{TD} \cdot N_{onz} \cdot q^2) + \lambda (aN_{onz} + bE_{nzq} + cN_{onzrr} + dN_{oz} + eN_{oto}). \end{split}$$

where, λ mode = 0.85 × 2^{\(\Lambda\)} (QP-12)/3 and R refers to the required number of bits for coding, S_{STD} is the sum of the squared transformed difference. The term D is the difference between the sum of squared reconstructed pixel and the original pixels in the macroblock. Nonz is the no of non-zero, Enzq is the encoding zero coefficient, QP is the quantization parameter N_{onzzr} denotes the number of times non-zero zero-run-length, N_{oto} is the number trailing ones.

5. Proposed Rate Control Scheme

In this proposed work, the rate control for the prediction frame done after encoding the I-frames using the base line profile encoding i.e. CAVLC encoding and taking the reference data from it. The rate control block diagram Figure 3 explains how the QP parameter is estimated. This happens between the encoder interface and the user interface. The complexity of the video is estimated and after that frame separation is done. The residuals

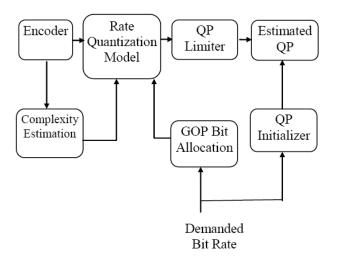


Figure 3. Rate control by estimating the quantization parameters.

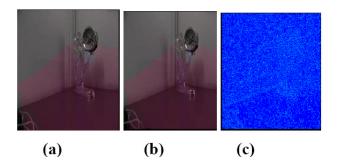


Figure 4. (a). Test image (b) QP = 10 (c) SNR Y= 52.2 U= 51.7 V= 52.0.

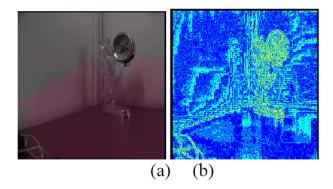


Figure 5. (a). QP = 27 (b) SNR Y = 42.8 U = 46.1 V = 52.0.

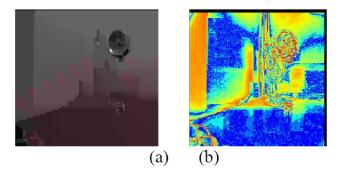


Figure 6. (a) QP=50 (b) SNR Y= 54.1 U = 30.0 V = 44.0.

between the current and reference frame is estimated. Mean Absolute Deviation (MAD) values are obtained by summing the residuals. The value of QP is initialized to a range manually. QP limiter is used to limit the demanded QP to a range and the parameters are estimated for the procedure

By using these parameters, rate control is performed on the input video to get an output video. This procedure ensures controlled bit-rate and better compression ratio.

Performance metrics are used to assess the quality of the obtained video. The test video sequence for 1 sec has been captured from Basler Scout scA 780-54 gm/gc camera by setting the frame rate at 24 per second. Figures 4, 5 and 6 shows the frames observed at different QP values.

The performance metrics on which the video quality adjudged are Compression ratio, Mean Square Error, Peak Signal to Noise Ratio, correlation coefficient, PRD and SSIM.

where,

- (i) Compression ratio: Input file size/Compressed file size.
- (ii) MSE (Mean Square Error):

$$\text{MSE} = (\frac{1}{MxN})\sum_{i=1}^{M}\sum_{j=1}^{N}(aij-bij)^2$$

Table 1. Comparative statement of QP, CR, PSNR(dB), MSE, PRD, Corr-Coeff, SSIM

QP	CR	PS NR	MSE	PRD	Cor- co-ff	SSIM
15	13.8	51.28	0.48	0.04	0.99	0.99
20	25.5	46.99	1.29	0.07	0.99	0.98
27	34.5	42.92	3.31	0.11	0.99	0.96
30	39.9	39.17	7.870	0.17	0.99	0.91
40	43.7	34.98	20.64	0.32	0.99	0.78

- (iii) PSNR (Peak Signal to Noise Ratio):
 PSNR= 10log₁₀ (255²/ MSE)
 M,N are Number of Rows and Columns , aij Input frame, bij Compressed frame.
- (iv) Correlation Coefficient: It is used to find the similarity between two different images with their intensities. It can be defined by, Cor_coef = [sum(sum(u1.* u2))] / [sqrt (sum

(sum (u1.* u1))*sum(sum(u2.* u2)))]

where, $u1 = a_{ij}$ – mean of a_{ij} , $u2 = b_{ij}$ – mean of b_{ij}

(v) Percentage Residual Difference(PRD):

It is used to measure the difference between original Image and Restored Image.

 $PRD = sqrt(sum((F1 - F2).^2)/sum(F1))$

where, F1 – Original frame and F2 – Compressed frame (vi) Structural Similarity Index (SSIM):

It is used to find the similarity between two different images. It will be described by,

$$SSIM = [(2u_x u_y) + c1)(2S_{xy} + c2)] / [(u_x^2 + u_y^2 + c1) (S_x^2 + S_y^2 + c2)] and c1 = (K_1L)^2, c2 = (K_2L)^2$$

where,

x, y - Original and compressed images,

 $u_x u_y$ - Mean of x and y. S_{xy} - Covariance of x and y.

 S_x^2 - Variance of x. S_y^2 - Variance of y.

L - Dynamic range of intensity values

 K_1 , K_2 – 0.01 and 0.03 (Default constant values)

Comparative analysis of the results using the parameters is represented in Table 1.

6. Conclusion

The inferences drawn using the results of this work are Rate control scheme for H.264 baseline encoder implemented by estimating the QP value using less complex algorithm. The rate distortion optimization has been used to reduce the coding complexity thereby improving the coding efficiency. The Quantization Parameter (QP) is varied to regulate the coded bit streams, thus achieves good perceptual video quality. When QP is 27, the quality of video is good because signal to noise ratio is less. It is observed that the quality of output video is found to be better in terms of performance metrics.

7. Acknowledgement

The Rate-Control Scheme mentioned in this paper is a part of the research project awarded to Sathyabama University as a consultancy programme funded by Combat Vehicles Research Development and Establishment (CVRDE), DRDO, Min. of Defense, Avadi Chennai-600054 India during the year 2014. The authors wish to extend thanks to Mr. P. Sivakumar, Distinguished Scientist, Director, CVRDE and all other members for entrusting and providing us this opportunity for contributing to eminent defense organization in India.

8. References

- Xiang Li, Norbert Oertel. Laplace distribution based Lagrangian rate distortion optimization for Hybrid Video Coding. IEEE Transaction on Circuits and Systems. 2009 Feb; 19(2):193–205.
- Sarwer MG, Po LM, Guo K, Wu QMJ. Transform-domain rate-distortion optimization accelerator for H.264/AVC video encoding. International Journal of Information and Communication Engineering. 2009; 3(11):1–11.
- Kwon D-K, Shen M-Y, Kuo CJ. Rate control for H.264 transactions on circuits Video with enhanced rate and distortion models. IEEE and Systems for Video Coding Technology. 2009 May; 17(5):372–81.
- James C, Alex MM. An efficient video codec based on H.264/AVC standard. IRACST. International Journal of Computer Networks and Wireless Communication. 2014 Apr; 4(2):1–4.
- Sathya S, Ramu JG. Performance evaluation of basic compression methods for different satellite imagery. Indian Journal of Science and Technology. 2015 Aug; 8(19):1–8. DOI: 10.17485/ijst/2015/v8i19/61658.
- Yu H, Lin Z. An improved rate control algorithm for H.264.
 2005 IEEE International Symposium on Circuits and Systems, 2005 May 23–26. 2005 May; 1:312–15.
- Wang J-X, Chiu Y-C, Su AWY, Shieh C-K. On parallelizing H.264/AVC rate-distortion optimization baseline profile encoder. Journal of Information Science and Engineering. 2010; 26:409–26.

- 8. Huang C-C. Perceptual rate distortion optimization for block mode selection in hybrid video coding. IEEE International Symposium on Circuits and Systems; 2013 May. p. 489-92.
- 9. Bernatin T, Sundari G. Video compression based on Hybrid transform and quantization with Huffman coding for video code. International Conference on Control,
- Instrumentation, Communication and Computational Technologies (ICCICCT). 2014 Jul; 4(7):476-80.
- 10. Liao K-Y, Yang J-F, Sun M-T. Rate-Distortion cost estimation for H.264/AVC. IEEE Transactions on Circuits and System Video Coding Technology. 2010 Jan; 20(1):38-49.