

A Novel Header Bits Estimation Scheme for H.264/AVC Standard

ShuQian He^{1, 2,a}, ZhengJie Deng¹ and Chun Shi^{1,b}

¹ School of Information Science and Technology
Hainan Normal University Hainan, China

² School of Information Science and Technology
Sun Yat-Sen University Guangzhou, China

^aheshuqian05@126.com, ^bbyshichun@126.com

Keywords: Rate-Distortion Optimization, Header Bit, Rate Control, H.264/AVC, Motion Vector.

Abstract. Rate estimation is useful for many H.264/AVC applications including rate-distortion optimization (RDO) for fast mode decision and precise rate control. In this paper, we propose a new header rate prediction model and an adaptive algorithm to provide more accurate estimation of the number of total coding bits for rate control compared to previously proposed methods. The header bit rate estimation is modeled by a linear combination of the number of mode block, and the sum of absolute values of all motion vectors for each block. Based on the proposed model, a header rate estimation function is also proposed to give a more accurate rate-distortion rate control. The proposed schemes can achieve better results in rate-distortion and rate control to previously proposed approaches.

Introduction

Unlike previously developed MPEG video coding standards, H.264/AVC additionally provides the following set of submacroblock types for each 8*8 submacroblock of a P-slice macroblock that is coded in INTER-8*8 mode: INTER-8*8, INTER-8*4, INTER-4*8, and INTER-4*4.

The H.264/AVC JM reference encoder adopts the RDO technique for mode-decision and rate-control to achieve the best coding performance. For example, there are 11 candidate modes (including SKIP, 16*16, 16*8, 8*16, 8*8, 8*4, 4*8, 4*4, intra_4*4, intra_8*8 and intra_16*16) for the inter-frame MB coding. The coding mode for each block is determined using the Lagrangian cost function can be expressed as [1] and [2]: $J_{Mode} = D + \lambda R$ (1)

The distortion D is measured as the sum of the squared difference between the current block and the reconstructed block. The rate R is the sum of the bits for mode including the mode information, the motion vectors, the side information and the transformation coefficients after entropy coding. λ is an appropriate Lagrangian multiplier which can be calculated with an empirical formula using the quantization parameter: $\lambda = 0.85 \times 2^{(Qp-12)/3}$ (2)

The cost function (1) can also be expressed as: $J_{Mode} = D + \lambda (R_H + R_{residual})$ (3)

Where R_H is the number of bits for the header information such as MB modes, motion vectors and reference frames; Residual is the bits for DCT coefficients of the residual information.

The RDO mode and rate-control Qp is selected as follows:

$$\min(D(\text{mod } e_i, Qp), \text{ subject to: } R(\text{mod } e_i, Qp) \leq R_T \quad (4)$$

Where e_i is mode type and R_T is the target bit rate.

During the video coding process, all possible inter/intra coding modes are examined and the residual encoding process with the actual encoding and decoding processes are calculated, including DCT/Q, IQ/IDCT, entropy coding and frame reconstruction. Due to the high-complexity RDO approach, some low processing devices for real time video applications cannot afford it[3-6].

In this paper, we propose a joint R-D optimized RC algorithm with an accuracy header bit model by analyzing RDO mode decision. Firstly, we analyze the header bits modeling problem of the header bit characteristics in H.264/AVC video coding which can be approximated by the number of blocks

and the number of MVs. So we propose joint rate distortion models for texture bits and header bits. This joint rate distortion modeling approach outperforms the traditional rate distortion models. Then, based on the joint rate distortion models and a novel target bit allocation approach, we propose joint rate distortion rate control algorithm by considering both the optimization bit allocation and the quality variation. Finally, We conduct extensive tests to evaluate the performance of our algorithm.

Header Rate Modeling

Since the rate and distortion of a macroblock are determined by both texture and header coding, the rate distortion relationship between the texture and header should be considered such that the optimization is possible. In this section, we propose the joint texture-header modeling technique for H.264/AVC video coding which can help us to achieve the joint texture-header optimization.

The bit rate of the video macroblock k consists of the bit rates of both the texture and header information: $R_{mb}^k = R_{texture}^k + R_{header}^k$ (5)

Where R_{mb}^k denotes the combined texture ($R_{texture}^k$) and header (R_{header}^k) bit rates for the macroblock k . The bit rates for header and motion information are involved in R_{mb}^k .

In this section, a joint rate model for both header bits and residual bits are proposed. We combine the bit-rate model with a distortion model to result in a cost function. Besides, an adaptive bit-rate estimation is proposed to improve the prediction accuracy in model-based rate control algorithms.

Generally speaking, there are two types of coded block in H.264/AVC coding: coded block and skipped block. Since the bit of the skipped block equal to 0, the bits of video coding produce by the coded block which can be coded by inter and intra-modes. Therefore, we need to estimate the rate of coded mode through the number of coded block only. The mode rate of coded blocks can be estimated via : $R_{mode} = N_{code} \cdot a_1 + b_1$ (6)

Where a_1 and b_1 are model parameters. N_{code} is the number of the coded block.

Accordingly, the Lagrange cost function can be improved by the texture rate and header rate model:

$$J_{Mode} = D + \lambda (R_H + R_{residual}) \quad (7)$$

The H.264/AVC JM reference encoder adopts the RDO technique for mode-decision to achieve the best coding block. Since the residual signal of a skipped block is not coded, its cost can be computed directly from its residual signal. Therefore, we need to estimate the cost function of coded blocks only. If $J_{code} < D_{skip}$, the coded block can be processed in the normal way, else the block can be skipped.

The rate-distortion cost function can be obtained by combining the rate model in (5) with the distortion model as: $J_{code} = D_{code} + \lambda R_{residual} + \lambda R_{header}$ (8)

Since $\lambda R_{header} > 0$, we can get $J_{code} > D_{code} + \lambda R_{residual} = J'_{code}$.

Where J'_{code} is the estimation value of the cost function, and D, R can be estimated by (5).

$$J'_{code} = R'_{residual} + \lambda D'_{code} \quad (9)$$

In other ways, if $J_{code} < D_{skip}$ and $J_{code} > J'_{code}$, we obtain $J'_{code} < D_{skip}$, and the number of coded

block N_{code} can define by $N_{code_i} = \sum_{i=1}^N C_i \cdot N_{code_{i-1}}$, with $C_i = \begin{cases} 1 & J'_{code_i} < D_{skip_i} \\ 0 & else \end{cases}$. The bits of mode

information can be estimated by the linear regress model with the model parameters a_1 , b_1 .

On the other hand, the motion bits of header bit have a strong relationship with the difference of horizontal/vertical MV elements MV_x , MV_y , and the number of MVs, N_{code} . As the value of MV with horizontal/vertical increase, the bits of header required to encode their values increases. Likewise, as a frame or unit is finely partitioned, there are more code block associated with it, which results in an increase in header bits including mode and motion bit mainly.

From Fig 1, there is linear relation between N_{code} and motion bits. And the information of the motion vectors for the current frame can be summarized as :

$$\bar{R}_{mv} = (R_{mvx} + R_{mvy}) \cdot g \cdot N_{code} \quad (10) \quad R_{mv} = a_2 \bar{R}_{mv} + b_2 \quad (11)$$

Based on the R_{mv} model and R_{mode} model, we propose a enhanced header rate model:

$$R_{header} = R_{mode} + R_{mv} + q \quad (12)$$

$$R_{header} = (N_{code} \cdot a_1 + b_1) + (a_2 \bar{R}_{mv} + b_2) + q$$

$$R_{header} = N_{code} (a_1 + a_2 g \bar{R}_{mv}) + b \quad (13)$$

Where $R_{mv} = R_{mvx} + R_{mvy}$, $b = b_1 + b_2 + q$.

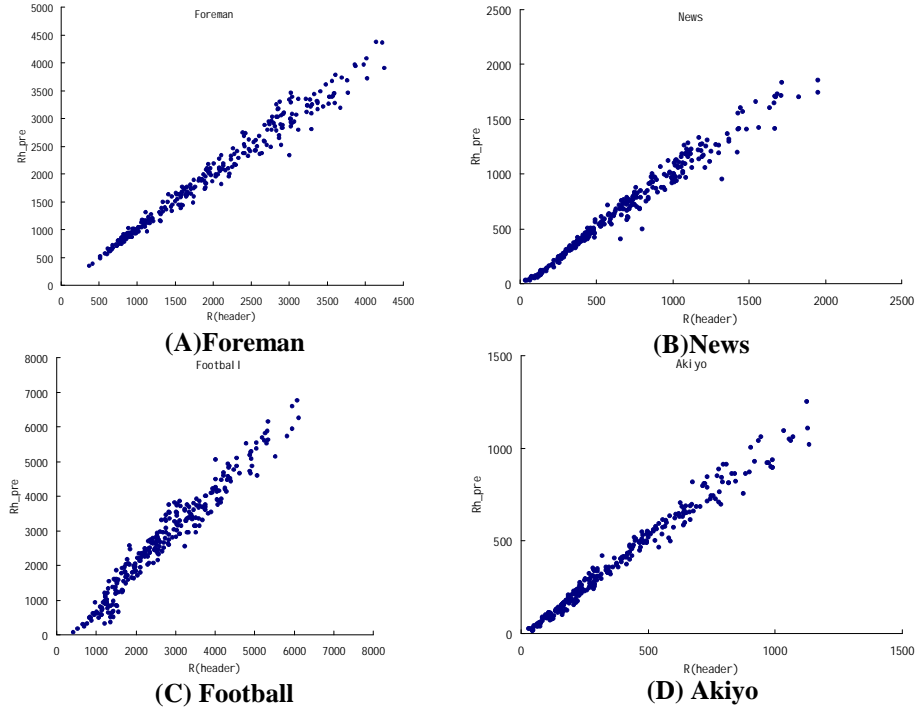


Fig 1: Relationship between R_{header} and $N_{code} (a_1 + g \bar{R}_{mv})$ for test sequences. (A)Foreman (B)News (C)Football (D)Akiyo

Since the rates are determined by both texture and overhead coding, the rate distortion relationship between the texture and overhead should be considered such that the optimization is possible. In this subsection, we propose the joint texture-header modeling technique for H.264/AVC RDO coding which can help us to achieve the joint texture-header optimization. The bit rate of the video frame consists of the bit rates of both the texture and overhead information which indicate above (13). According to (13), joint texture-header R-D model is given by :

$$R_{sum} - [N_{code} (a_1 + a_2 g \bar{R}_{mv}) + b] = Com[i] \left(\frac{a[i]}{Q^2[i]} + \frac{b[i]}{Q[i]} \right) \quad (14)$$

$$D_{residual}[i] = a_2[i]Q + b_2[i]Q^2 \quad (15)$$

Where R_{sum} denotes the combined texture and overhead bit rates.

Experimental Comparison Results

To validate the performance of the proposed model, we apply the estimation functions to the H.264/AVC reference software JM version 10.2. Fast motion estimation is enabled with 1/4 pixel resolution. The number of reference frames is set to 1 and the motion search range is 16. The R-D optimization is enabled in motion estimation and CAVLC entropy coding is enabled. In the simulations, several test video sequences are used to obtain various coding performances. To evaluate

the performance of the proposed rate distortion tradeoff approach, more results are listed in Table I. We achieved a PSNR gain of up to 0.87dB with an average PSNR gain of 0.33dB over the entire test set. We also list the different between target bit and bit rate in Table I to evaluate the robustness of our algorithms.

Table I:Simulation results of our proposed algorithm in teams of PSNR(dB) and Bit rate (kb/s)

Sequence	RC Algorithm	Target Rate	PSNR	Bitrate
Foreman QCIF 30Hz	[3]	65.34	32.67	65.45
	Ours		32.68	65.63
	Gain		0.01	-
News QCIF 30Hz	[3]	36.47	32.51	36.96
	Ours		32.82	36.68
	Gain		0.31	-
Akiyo QCIF 30Hz	[3]	13.96	33.68	14.32
	Ours		34.55	14.13
	Gain		0.87	-
Mobile QCIF 30Hz	[3]	147.09	28.55	147.42
	Ours		28.95	148.12
	Gain		0.4	-
Silent QCIF 30Hz	[3]	38.67	32.23	38.96
	Ours		32.62	38.85
	Gain		0.39	-

Conclusions

In this paper, we first proposed a header bit rate model included mode and motion parameter, and an adaptive header bit rate estimation approach for joint rate distortion model in rate control. The proposed adaptive method utilizes information from the previous coded macroblocks. Experimental results show that the proposed estimation methods are effective for rate distortion model prediction and rate control.

Acknowledgements

This work was financially supported by the Hainan Natural Science Foundation (611128, 612122,613164, 613163) and Natural Science Foundation of China under Grant (61163042).

References

- [1]T.Wiegand,G.J.Sullivan,G.Bjontegaard,and A.Luthra. Overview of the H.264/AVC video coding standard[J].IEEE Transaction on Circuits and Systems for Video Technology, 2003 , 13(7) :560-576.
- [2]T.Wiegand, G.J.Sullivan, and A.Luthra. Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification(ITU-T Rec. H.264---ISO/IEC 14496-10 AVC),
- [3]Z.G.Li,F.Pan,K.P.Lim,G.Feng,X.Lin,and S.Rahardja. Adaptive basic unit layer rate-control for JVT[S],presented at the 7th JVT Meeting,Pattaya II JVT-G012-rl Thailand, 2003.
- [4] Y.-K. Tu, J.-F. Yang, and M.-T. Sun, "Efficient rate-distortion modeling for efficient H.264/AVC encoding," IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 5, pp. 530–543, May 2007.

[5]D.K.Kwon,M.Y.Shen,C.C.Kuo.Rate control for H.264 video with enhanced rate and distortion models. IEEE Transaction on Circuits and Systems for Video Technology,2007,17(5):517-529.

[6]Yang Liu,Zhengguo Li and Yeng Chai Soh. A Novel Rate Control Scheme for Low Delay Video Communication of H.264/AVC Standard[J]. IEEE Transaction on Circuits and Systems for Video Technology,2007,17(1):68-78,January 2007.