A NOVEL RATE CONTROL ALGORITHM for H.264/AVC

Zhao Min†, Takeshi Ikenaga, Satoshi Goto
Graduate School of Information, Production and Systems, Waseda University, Japan
†email: minzhao@fuji.waseda.jp

Abstract — Rate control has become an important part in the whole video coding scheme in order to get obtain consecutive and high quality pictures under a certain bandwidth. This paper presents a novel rate control algorithm for H.264/AVC. Instead of the traditional Mean Absolute Difference (MAD) linear prediction model, which is proposed JVT-G012, a new model is proposed including both temporal and spatial information of MAD. Because calculation complexity of the rate control part is not high comparing with other parts, such as motion estimation, this novel algorithm does not focus on reducing the calculation complexity.

Keywords — rate control, H.264/AVC, MAD, rate-distortion optimization

Ⅰ. INTRODUCTION

In order to obtain high quality video pictures in the constant bit rate (CBR) channel, rate control scheme should be used in the encoding process of variable bitrate video bit stream. So the rate control has become a very important part in the video communication application. When designing a rate control algorithm, there are two basic problems to be considered. One is the optimal assignment of available bits, and the other is the adjustment of the encoder parameters such as quantization parameter (QP) to control the bit rate. A number of rate control schemes are proposed in many different standards in order to solve these two problems. The literature [1, 2] proposed the scheme of bit allocation adopted by MPEG-2 TM5, which allocates the constant bit amount for each group of picture (GOP). To realize this bit allocation scheme, literature [3, 4] proposed the rate control schemes based on the using Lagrange rate distortion optimization (RDO), which is adopted by TMN8 of H.263. This scheme uses the square difference as the control parameter and its complexity is very high, so the correlation between of the number of non-zero coefficient through transforming and quantizing and bitrate of video sequence is discovered in literature [5] and a simple linear model is proposed. The literature [6, 7] proposed the quadratic quantization model, which is very simple and useful until now. Up to now, the typical schemes mainly include MPEG-2 TM5, H263 TMN8, MPEG-4 VM8, VM18 ρ-region model and JVT-G012 [8] for H.264/AVC.

The official rate control scheme was presented in JVT-G012 with temporal MAD prediction method and rate-quantization (R-Q) quadratic model. But this scheme has two obvious problems:

♦ *Inaccurate MAD prediction*: MAD is a kind of simple computational measure in the spatial domain. So the temporal model performs poorly in the presence of such sudden changes.

♦ *Inaccurate bit allocation and estimation:* In JVT-G012, the amount of header bits of the current basic unit is predicted from the previous encoded basic units, for example by averaging. However actually header bits of the basic unit are quite different and it should not be regarded as a fixed-size component.

Based on G012 rate control scheme a novel rate control scheme is proposed in this paper for H.264/AVC standard. This rate control scheme is mainly focused on these two problems in the first paragraph of introduction and tries to use new algorithms to improve them. And the rest of this paper is organized as follows: The section II proposed a more accurate MAD prediction model; the section III discusses the quadratic R-D model; the section IV presents the simulation result and conclusion.

II. More Accurate MAD Prediction

The rate control for H.264 is more difficult than those for other standards. Because the quantization parameters are used in both rate control algorithm and rate distortion optimization (RDO), which resulted in a chicken and egg dilemma. That is to say to perform RDO for macroblocks (MBs) in the current frame, a QP should be first determined for each MB by using the MAD of current frame or MB [9]. However, the MAD of current frame or MB is only available after the RDO. To solve this QP dilemma, a simple linear model is proposed in [8] to predict the MAD

$$MAD_{cb} = a_1 \times MAD_{pb} + a_2$$

where a_1 and a_2 are two coefficients of prediction model, MAD_{cb} and MAD_{pb} denote the predicted MAD of current basic unit in the current frame and the actual MAD of basic unit in the co-located position of previous frame respectively.

Unfortunately, this model which only uses previous temporal information to predict the MAD is not accurate enough for MAD prediction especially when MAD changes abruptly due to high motion or scene changes. In order to improve the model and propagate the property of current frame, it is desirable to collect more information and add a spatial model. Let MAD_d be a measure for evaluating the difference between the current original frame and the previous reconstructed frame. It can be calculated after motion determination only with the INTER16×16 and INTRA16×16 modes. Then we can get another formula

$$MAD_{cb,spa} = b_1 \times MAD_d + b_2$$

where $MAD_{cb,spa}$ is the spatial predicted MAD of the current basic unit, b_1 and b_2 are two coefficients of this prediction model.

We can present MAD_{cb} as $MAD_{cb,tem}$ to show the difference between temporal model and spatial model. So the formula should be as below

$$MAD_{cb,tem} = a_1 \times MAD_{pb} + a_2$$

In order to choose the better model, we introduce two similarity measures as indicators of the efficiencies of the temporal MAD prediction model and the spatial MAD prediction model. Since the variance is a common way to capture the data scale or degree of being spread out, we apply in the following formula

$$\begin{aligned}\mu_{tem} &= E((MAD_{cb,tem} - MAD_{actual})^2) \\ &= \sum_K (MAD_{cb.tem} - MAD_{actual})^2\end{aligned}$$

$$\mu_{spa} = E((MAD_{cb,spa} - MAD_{actual})^2) = \sum_K (MAD_{cb.spa} - MAD_{actual})^2$$

where K is the number of MAD samples used to measure μ and MAD_{actual} is the actual MAD of current basic unit.

After comparing the variance of the temporal model and spatial model of the current basic unit, we can decide which model should be chosen when predicting the next basic unit:

if $\mu_{spa} > \mu_{tem}$ **then**

$$MAD_{nb} = MAD_{nb.spa}$$

else

$$MAD_{nb} = MAD_{nb.tem}$$

where MAD_{nb} is the predicted MAD of next basic unit.

That is to say that we choose a better linear MAD prediction model of next basic unit by comparing μ_{spa} and μ_{tem}. This new model replaces the old one in the whole rate control scheme.

III. QUADRATIC R-Q MODEL

The rate distortion optimization problem is to find a mathematical R-Q model to approximate this monotonic increasing relation between SumBit/MAD and $1/Q_p$. In H.264/AVC standard, this R-Q model [10, 11, 12] is more complex than any other standards and employed as

$$R_{sum} = MAD_{cb} \times (\frac{X_1}{Q_p} + \frac{X_2}{Q_p^2})$$

where R_{sum} is the target bit of the basic unit to be encoded, X_1 and X_2 are two parameters of the R-Q model, Q_p is the QP value.

Though the quadratic R-Q model may increase the calculation complexity, it could achieve more accurate approximating result. Comparing with other module such as motion estimation, the rate control module only takes very little computation time and hardware cost, so it is acceptable to keep using the original R-Q model [13] in H.264. So the quadratic R-Q model part in [8] is kept.

According to [8], the whole rate control algorithm scheme has three layers: GOP and frame and basic unit layer rate control. Our new rate control scheme still uses this three-layer algorithm except the different models in these three layers.

IV. EXPERIMENTAL RESULTS AND CONCLUSIONS

The test condition and test sequences are given Table 1. The test result is showed in Table 2.

Hadamard	ON	RD optimization	On
Search Range	±32	GOP structure	IPPP
Reference Frames	1	Symbol Mode	CABAC

Table 1. Test Conditions

Sequence	RC Sch.	Bitrate(kb/s)	PSNR(dB)
FORMAN QCIF@30Hz	G012	19.59	24.43
	Ours	19.56	24.71
FOOTBALL QCIF@30Hz	G012	63.44	22.77
	Ours	63.45	23.06
MOBILE QCIF@30Hz	G012	49.72	23.00
	Ours	49.68	23.37

GRAMMA QCIF@30Hz	G012	32.41	31.96
	Ours	32.41	32.18
NEWS QCIF@30Hz	G012	27.30	27.03
	Ours	27.29	27.24

Table 2. Test Results

Hence an improved MAD linear prediction algorithm for rate control scheme is proposed in this paper. The two linear models are used to predict MAD of current basic unit in the current frame by that of basic unit in the same position of previous frame and in the position of current reconstructed frame, which present temporal and spatial information respectively. The prediction accuracy is increased by introduction of these models. The quadratic R-Q model in [8] is kept using because it does not increase calculation complexity of the whole algorithm comparing with other high complexity modules.

The algorithm uses three layers rate control scheme. And by applying the new models, it makes the sudden change of rate less than the standard algorithm and improves the PSNR of the signal and reduces the bit rate. The PSNR is averagely increased 0.272db comparing with the original G012 while having a lower bit rate.

REFERENCES

[1] J.W.Lee, and Y.S.Ho, "Target Bit Matching for MPEG-2 Video Rate Control," *Proceeding of Tencon* 1998, New Delhi, India, 1998, vol.1, pp: 66-99

[2] "MPEG-2 Video Test Model 5", ISO/IEC/JTC1/SC29/WG11, MPEG93/457, 1993

[3] J.Ribas-Corbera, and S.Lei, "Rate Control in DCT Video Coding for Low-Delay Communications," *IEEE Trans. On Circuits and Systems for Video Technology*, 1999, vol.9 (1), pp: 172-185

[4] "Video Codes Test Model Near-Term," *TMN8.* ITU-T/SG16/VCEG/Q15 A59, Portland, USA, 1997

[5] Zhihai. He, Yong Kwan Kim, and Sanjit K. Mitra, "Low-Delay Rate Control for DCT Video Coding via ρ-Domain Source Modeling," *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, vol.11(8), pp: 928-940

[6] T.Chiang, and Y.Q.Zhang, "A New Rate Control Scheme Using Quadratic Rate Distortion Model," IEEE Transactions on Circuits and Systems for Video Technology, 1997, vol.7 (1), pp: 246-250

[7] Vetro, H.Sun, and Y.Wang, "MPEG-4 Rate Control for Multiple Video Objects," *IEEE Transactions on Circuits and Systems for Video Technology*, 1999, vol.9, pp: 186-199

[8] Z. G. Li, F. Pan, K. P. Lim, G. Feng, X. Lin, and S. Rahardja, "Adaptive basic unit layer rate control for JVT," presented at the 7th JVT Meeting, Pattaya II JVT-G012-rl Thailand, Mar. 2003.

[9] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document JVT-B118R2, 2002-03-25.

[10] H.J.Lee and T.H.Chiang and Y.Q.Zhang. Scalable Rate Control for MPEG-4 Video. IEEE Trans. Circuit Syst. Video Technology, 10: 878-894, 2000.

[11] A.Vetro, H.Sun and Y.Wang. MPEG-4 rate control for multiple video objects. IEEE Trans. Circuit Syst. Video Technology, 9: 186-199, 1999.

[12] Z. G. Li, Lin Xiao, C. Zhu and Pan Feng. A Novel Rate Control Scheme for Video Over the Internet. In Proceedings ICASSP 2002, Florida, USA, 2065--2068, May 13-17, 2002.

[13] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document JVT-B118R2, 2002-03-25.