

ESTIMATING THE VALUE OF θ IN THE INTRA FRAME FOR ρ -DOMAIN RATE CONTROL ALGORITHMS

Zhihang Wang¹, Shengfu Dong², Long Xu¹, Wen Gao^{1,2}, and Qingming Huang³

¹Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China
{zhihangwang, lxu, wgao}@jdl.ac.cn

²Institute of Digital Media, Peking University, Beijing 1000871, China
sfdong@jdl.ac.cn

³Graduate School, Chinese Academy of Sciences, Beijing, 100039, China
qmhuang@jdl.ac.cn

ABSTRACT

The ρ -domain rate control algorithm is very simple and effective. There is only one parameter, θ , used in the model. θ reflects the relationship between the rate and the quantization step. Accurately estimating the θ value of the I (intra) frame is important for an efficient ρ -domain rate control algorithm because the I frame will influence the remaining frames in the GOP (group of pictures) greatly. In this paper, we propose to estimate the θ value based on the energy of the frame. The energy difference of the I frames is used to predict the new θ . A linear relationship between θ and the energy is deduced. Experimental results show that using the energy can more accurately predict the θ value of I frame than only using the θ value of the previous P (predict) frame.

Index Terms— ρ -domain, Rate Control, θ Estimation

1. INTRODUCTION

Rate control plays an important role in video coding. We always want to use the give bit rate efficiently to encode the video sequence so that the video quality is maximized. Many research papers have been published discussing how to accurately and effectively control the bit rate for the video encoder. Various rate distortion models based on the classical $R - D$ functions are discussed. The logarithmic expression is proposed and discussed in [1]. Power model [2], spline model [3], and polynomial model [4] are also proposed in the literature. Although these models can be used to control the bit rate, its complexity is usually large. The ρ -domain rate control algorithm was first proposed in [5]. Despite its accuracy in the rate control algorithm, a big advantage is its simplicity. The rate is modeled with the percentage of zeros (ρ) and their relation is formulated as

$$R = \theta \cdot (1 - \rho) \quad (1)$$

Only one parameter, θ , should be controlled using this model. The accuracy of the θ value controls the performance of the

algorithm.

The difficult problem is how to estimate the θ value accurately especially for the I frame because the I frame is used as the reference frame for the remaining inter frames. This problem includes two parts: one is how to decide the θ value for the first I frame, the other is how to estimate the θ value of the I frame in every GOP. Usually the θ of the first I frame in the video sequence is set to a fixed value. For example, Z. He *et al.* [6] proposed to use 7 as the initial θ value in the first macroblock. It is obvious that this method is not robust because the diversity of the images. For the other I frames, usually the θ value of the previous encoded frame is used to estimate the current I frame θ value. Z. He *et al.* proposed to use

$$\theta^{K-1} = \frac{R^{K-1}}{1 - \rho^{K-1}} \quad (2)$$

to estimate the θ value of the K frame. Although this method make sense in the same video scenes, it still needs to study the method to estimate θ in different scenes which is the main task of this paper.

Based on extensive experiments, we find that the value of θ is related to the energy of the frame. Thus we propose a linear model to model their relationship and give the theoretical analysis of this relationship.

The rest of this paper is organized as follows. In Section 2, we analyze the reason that why the θ value will change. In Section 3, We analyze the relationship between θ and *energy* in the theory and propose an linear model to describe their relationship. Experimental results are also given in this section to justify this model. The algorithm using this model is proposed in Section 4. Experimental results using this model are also given in this section. Section 5 concludes this paper.

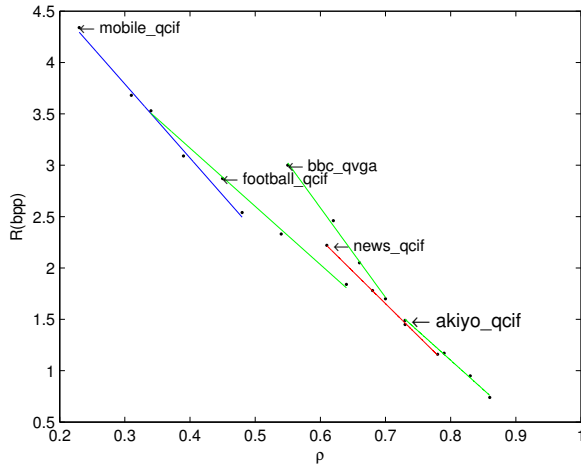
2. WHY θ WILL CHANGE

The ρ -domain rate control algorithm is based on the observation that zeros play an important role in the entropy coding.

Table 1. ONE AVS INTRA VLC TABLE

Run	EOB	level > 0					
		1	2	3	4	5	6
	8	-	-	-	-	-	-
0	-	0	4	15	27	41	55
1	-	2	17	35	-	-	-

Zeros can be efficiently encoded with *run*. So more zeros mean higher coding efficiency. Following this logic, we will get the same R if we have the same percentage of zeros (ρ) in every frame. But just as Fig. 1 shows that different video sequence will have different R with the same ρ .

**Fig. 1.** The relationship between R and ρ

From equation (1), we can know that the θ value should be different. This is due to the entropy coding method. We use AVS (Audio Video Coding Standard) as an example to explain this phenomena. AVS adopts the context-based 2D-VLC entropy coding engine [7]. Zeros and the quantization levels form the *run-level* record. The *run-level* record is encoded by looking up VLC tables and the corresponding table element is encoded by Exponential-Golomb code. Table 1 shows an intra VLC table in AVS. In the VLC table, we can find that for the same run, bigger levels correspond to bigger table elements. Thus it will consume more bits to encode the bigger levels. This is the reason that the R value may be different for the same ρ . So the θ value is different. It also motivates us to find the relationship between θ and the quantization levels.

3. THE RELATIONSHIP BETWEEN θ AND THE QUANTIZATION LEVELS

In this section, we first analyze the relationship between θ and the quantization levels in the theory. Then we design the experiment to find their concrete relationship.

3.1. THEORETICAL ANALYSIS OF THE RELATIONSHIP

It has been well accepted that, after intra prediction, the transformed residuals follow a Laplacian distribution f_L [8]. Normally, it can be supposed as a zero-mean distribution, i.e.

$$f_L(x) = \frac{\lambda_L}{2} e^{-\lambda_L|x|} \quad (3)$$

Note that for Laplacian distribution,

$$\lambda_L = \frac{\sqrt{2}}{\sigma} \quad (4)$$

where σ is the standard deviation of the signal. The differential entropy (H_0) of the Laplacian distribution can be calculated as (5)

$$H_0 = \log_2(\sqrt{2}e\sigma) \quad (5)$$

From (1) and (5), we can get (6)

$$\theta = \frac{\log_2(\sqrt{2}e\sigma_c)}{1 - \rho} \quad (6)$$

where σ_c is the standard deviation of the DCT coefficients. Thus the relation between θ and the quantization levels can be modeled as (7)

$$\theta = \eta \log_2(\sqrt{2}e\sigma_c) + b \quad (7)$$

Equation (7) suggests that it is possible to estimate θ from the DCT coefficients. We use the standard deviation of the quantization levels to estimate the standard deviation of the coefficients. Formula (7) becomes

$$\theta = \eta \log_2(\sqrt{2}e\sigma_l \times \frac{Qstep}{32}) + b \quad (8)$$

where σ_l denotes the standard deviation of the quantization levels. Dividing 32 is only a normalization operation. It doesn't influence the real relationship. Next we design the experiment to investigate their relationship. We use *energy* to denote $\log_2(\sqrt{2}e\sigma_l \times \frac{Qstep}{32})$.

3.2. EXPERIMENTS TO FIND THE CONCRETE RELATIONSHIP

In order to investigate the relationship between θ and the quantization levels, we select 12 QCIF video sequence as

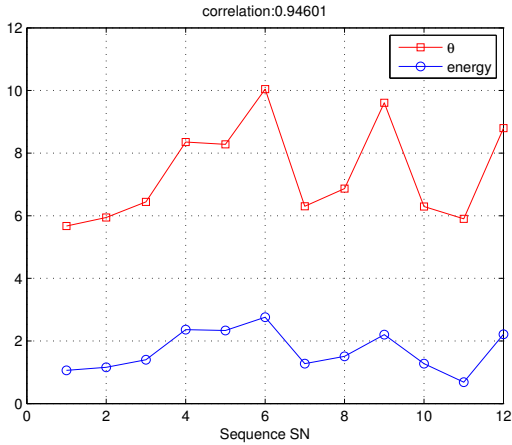
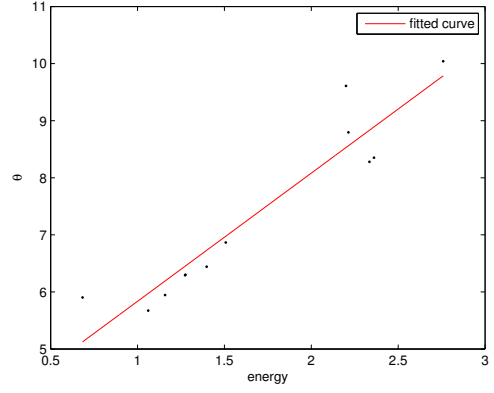
Table 2. QCIF VIDEO SEQUENCES

SN	Sequence	SN	Sequence	SN	Sequence
1	foreman	2	football	3	bus
4	paris	5	tempe	6	mobile
7	container	8	news	9	flower
10	hall	11	soccer	12	stefan

Table 3. CORRELATION COEFFICIENTS OF θ AND ENERGY FOR DIFFERENT FRAMES

QP	1st frame	50th frame	140th frame
20	0.8810	0.8932	0.9266
24	0.8959	0.9000	0.9352
28	0.9067	0.9027	0.9419
32	0.9232	0.9068	0.9460

listed in Table 2. We encode the first 10 frames of every sequence as I frames and set the quantization parameter to 20, 24, 28, and 32. Thus we can get the θ value of every frame according to the residual bits and the percentage of zeros. We also output the *energy* value of every frame. θ and *energy* are averaged for every sequence. It can be seen from Fig. 2 that the correlation coefficient between them is 0.94601. Many other frames are also tested and we get similar results as shown in Table 3. These experimental results show that there exists high correlation between θ and *energy*. We can fitted these points and get the corresponding equation as shown in Fig. 3. Since η and b should be updated for every frame, we only use the changing trend of *energy* to estimate θ in the next section.

**Fig. 2. The changing trend of θ and energy****Fig. 3. The fitted line of θ and energy**

4. PROPOSED ALGORITHM

In this section, first we describe the algorithm using the linear prediction model to predict the θ value. Then we give the experimental results to show its performance.

4.1. PREDICTING θ ALGORITHM

First we do the intra prediction on the original frame. This can utilize the parallel computing capacity of the multi-core server. The QP is set to 32 to quantize the residual of the best mode. The resulting levels are used to calculate *energy*. The initial value of η and b is set to 1.2 and 5.2. The θ value of the next I frame is calculated according to

$$\hat{\theta}^k = \theta^{k-L_{gop}} + \eta^k \log_2(\sqrt{2}e\sigma_l^k \times \frac{Qstep}{32}) + b^k - \eta^{k-L_{gop}} \log_2(\sqrt{2}e\sigma_l^{k-L_{gop}} \times \frac{Qstep}{32}) - b^{k-L_{gop}} \quad (9)$$

where L_{gop} is the GOP length and $\theta^{k-L_{gop}}$ is the true θ value of the previous I frame. Equation (9) can reflect the changing trend of θ . Thus it will be more accurate than using the previous θ^{k-1} of P frame to predict θ^k . In order to avoid the calculation of η and b and keep track of the changing trend of θ , we simplified equation (9) and use ΔE to reflect the change. Thus (9) changes to 10.

$$\hat{\theta}^k = \theta^{k-L_{gop}} + \log_2(\sqrt{2}e\sigma_l^k \times \frac{Qstep}{32}) - \log_2(\sqrt{2}e\sigma_l^{k-L_{gop}} \times \frac{Qstep}{32}) \quad (10)$$

The θ value of the first I frame is predicted using equation (8).

4.2. EXPERIMENTAL RESULTS

We select some QCIF sequences in Table 2 to verify the model and the algorithm. Our optimized AVS RM52j reference soft-

ware is used as the test software. The first 10 frames of every sequence are concatenated to form a 120 frames long sequence with 12 scenes. The first frame of every scene is encoded as I frame and the remaining frames are encoded as P frames. It can be seen from Fig. 4 that the changing trend of θ is almost the same as the energy. The change quality of θ now is approximated by ΔE as (10). It can be seen that this method can predict the θ value more accurate than using the θ of the previous P frame.

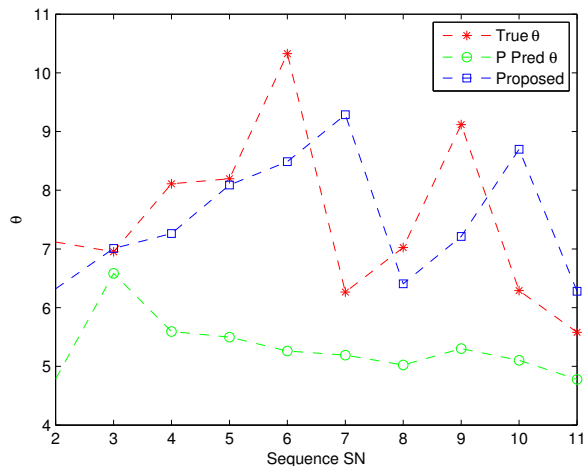


Fig. 4. The prediction result of θ for sequences in Table 2.

5. CONCLUSION AND THE FUTURE WORK

In this paper, we proposed a method to predict the θ value of I frame for different scenes. Because the abrupt change of the scene, it is a very challenging problem. ΔE is proposed to track the change of θ and used as an approximate value of $\Delta\theta$. Traditionally using the θ of the previous P frame to predict the current θ is not reasonable if the two frame belong to two different scenes. Our method can provide better prediction results. Some sequences still can't predict because of the difference between ΔE and $\Delta\theta$. The theoretical basis of our algorithm and more accurate predicting algorithm should be further studied.

6. ACKNOWLEDGMENT

This work is sponsored by National Basic Research Program of China under the contract NO. 2009CB320903 and Co-building Program of Beijing Municipal Education Commission.

7. REFERENCES

- [1] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 9, no. 1, pp. 172–185, 1999.
- [2] W. Ding and B. Liu, "Rate control of MPEG video coding and recording by rate-quantization modeling," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 6, no. 1, pp. 12–20, 1996.
- [3] L.J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 8, no. 4, pp. 446–459, 1998.
- [4] T. Chiang, Y.Q. Zhang, D.S.R. Center, and NJ Princeton, "A new rate control scheme using quadratic rate distortion model," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 7, no. 1, pp. 246–250, 1997.
- [5] Z. He and SK Mitra, "Optimum bit allocation and accurate rate control for video coding via ρ -domain source modeling," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 12, no. 10, pp. 840–849, 2002.
- [6] Z. He, "Zhihai ph.d thesis," Website, 2001, http://vision.ece.ucsb.edu/publications/Zhihai_Thesis.pdf.
- [7] Q. Wang, D.B. Zhao, and W. Gao, "Context-Based 2D-VLC Entropy Coder in AVS Video Coding Standard," *Journal of Computer Science and Technology*, vol. 21, no. 3, pp. 315–322, 2006.
- [8] EY Lam and JW Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *Image Processing, IEEE Transactions on*, vol. 9, no. 10, pp. 1661–1666, 2000.